

# Bruk av chatbot i praksis

*En kvalitativ studie av utfordringer ved bruk  
av chatbot i offentlige tjenester*

Tina Steinstø



Oppgave for graden  
Master i Informatikk: design, bruk, interaksjon  
60 studiepoeng

Institutt for informatikk  
Det matematisk-naturvitenskapelige fakultet

UNIVERSITETET I OSLO

Juni 2020



# **Bruk av chatbot i praksis**

*En kvalitativ studie av utfordringer ved bruk av  
chatbot i offentlige tjenester*

Tina Steinstø

© 2020 Tina Steinstø

Bruk av chatbot i praksis

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

# Sammendrag

Denne masteroppgaven er en kvalitativ studie av utfordringer som kan oppstå ved bruk av chatbot i offentlig sektor. Masteroppgaven beskriver en case-studie av bruk av chatbot i NAV. NAV sin chatbot, Frida, er en del av NAV sin førstelinjetjeneste og kan gi svar på generelle spørsmål om NAV sine tjenester og ytelser. Brukere som ønsker å kommunisere med NAV via chat vil først bli satt i kontakt med Frida før de kan bli satt over til en menneskelig veileder. Studien vil gjennom intervjuer, observasjoner og analyse av chatbotlogger gjøre rede for utfordringene som oppstår når NAV sine brukere interagerer med Frida, samt hvordan overgangen mellom chatbot og veileder fungerer. Analysen vil bygge på Lucy Suchman sitt analytiske rammeverk og chatloggene vil bli kodet ut ifra en hermeneutisk tilnærming. Loggene viser at de mest fremtredende utfordringene er tilknyttet Språk, Chatbotforståelse, Domenekunnskap og Intensjoner. Språk dreier seg om *skrivefeil, feilidentifisert språk og stavekontroll*. Chatbotforståelse omfatter *personlige spørsmål og oppfølgingsspørsmål*. Intensjoner dreier seg om utfordringer knyttet til *feil eller manglende intensjon, presisjon og kontekstuell relevant informasjon*. Domenekunnskap omfatter bruk av *feil begrep og puttekassen*. Andre utfordringer er *tillit, ignorerte ord og brukere som ikke ser at de snakker med en chatbot*. Analysen viser at flere av utfordringene oppstår på grunn av funksjoner som er skjult for brukeren, samt forventninger fra brukeren som ikke stemmer overens med chatbotens faktiske funksjoner. Studien konkluderer med at å gjøre chatbotens funksjoner tydeligere, samt mer bruk av avklarings spørsmål og knapper, vil gjøre at svarene oppleves som mer presise for brukerne. I tillegg vil formidling av domenekunnskap til brukere med manglende forståelse for NAV sitt begrepsapparat, gjøre at brukerne kan føle mer tillit til svarene de får av chatboten.

**Nøkkelord:** chatbot, digitalisering i offentlig sektor, HCI



# Forord

Det er en rekke personer som har bidratt til muliggjøring av denne masteroppgaven. Først vil jeg takke min hovedveileder Tone Bratteteig for faglig oppfølging og hjelpsomme tilbakemeldinger. Takk også til Guri Verne, som jeg også hadde veiledningstimer med i høst, for oppklarende og nyttige diskusjoner. Jeg vil også takke Linett Simonsen for et godt samarbeid med datainnsamling og analyse. Det å ha noen å diskutere med underveis og samarbeide med under analysen har vært til stor nytte. Jeg vil også takke informanter i NAV Kontaktsenter Oslo og Vest-Viken og NAV Kontaktsenter Vestland for deres tid og for den nyttige informasjonen dere bidro med. En spesiell takk til Terese i NAV Kontaktsenter Oslo og Vest-Viken for tilrettelegging og hjelp underveis. Jeg vil også takke informanter i Skatteetaten og NetNordic for at dere satte av tid til oss og delte deres erfaringer med oss. Til slutt vil jeg gjerne takke familie og venner som har hjulpet til med korrekturlesing og kommet med nyttige tilbakemeldinger på oppgaven underveis.

**Tina Steinstø**

Universitetet i Oslo, juni 2020





# Innholdsfortegnelse

1	Introduksjon .....	2
1.1	Innledning.....	2
1.2	Forskningsspørsmål.....	4
1.3	Oppgavens struktur.....	4
2	Bakgrunn.....	6
2.1	NAV.....	6
2.1.1	Om NAV .....	6
2.1.2	NAV Kontaktsenter.....	7
2.1.3	Digitalisering i NAV .....	8
2.2	Kunstig intelligens.....	8
2.2.1	Kunstig intelligens.....	8
2.2.2	Chatboter .....	11
2.2.3	Maskinlæring.....	12
2.2.4	Dype nevralt nettverk .....	13
2.2.5	Chatbot som informasjonsformidler.....	15
2.3	Tidligere forskning .....	16
2.3.1	Bruk av kunstig intelligens i dagliglivet .....	16
2.3.2	Automatisering av offentlige tjenester .....	18
2.3.3	Bruk av chatbot i offentlig sektor.....	20
2.4	Oppsummering .....	21
3	Chatboten Frida .....	23
3.1	Hvordan Frida fungerer .....	23
3.1.1	Intensjonstrær .....	25
3.1.2	Hvordan Frida analyserer en setning.....	27
3.2	Trening.....	31
3.3	Utviklingen av Frida.....	32
4	Metodologi .....	36
4.1	Datainnsamlingsmetoder .....	36
4.2	Oversikt over datainnsamlingsaktiviteter .....	40
4.2.1	Tidslinje over arbeid med masteroppgaven i 2019 .....	41
4.3	Analyse .....	43

4.3.1	Koding chatbotlogger .....	43
4.3.2	Gjengivelse av chatlogger .....	45
4.3.3	Analyse med Suchman sitt ramme verk .....	51
4.4	Formelle tillatelse for studien .....	57
5	Kategori: Vellykkede samtaler .....	59
5.1	Riktig svar .....	60
5.2	Riktig henvisning til skjema/nettside .....	61
5.3	Riktig henvisning til veileder .....	62
5.4	Oppsummering .....	64
6	Når samtale settes over til veileder .....	66
6.1	Frida foreslår å sette over .....	66
6.2	Bruker spør om å bli satt over .....	71
6.3	Brukere som har problemer med å bli satt over .....	76
6.4	Oppsummering .....	82
7	Utfordringer .....	84
8	Språk .....	85
8.1	Utfordringer knyttet til språk .....	85
8.1.1	Skrivefeil .....	85
8.1.2	Feilidentifisert språk .....	87
8.1.3	Stavekontroll .....	89
8.2	Språk som utfordring .....	91
8.3	Oppsummering .....	92
9	Chatbotforståelse .....	93
9.1	Utfordringer knyttet til chatbotforståelse .....	93
9.1.1	Personlig spørsmål .....	93
9.1.2	Kontekst/oppfølgingsspørsmål .....	95
9.2	Chatbotforståelse som utfordring .....	98
9.3	Oppsummering .....	98
10	Domenekunnskap .....	100
10.1	Utfordringer knyttet til domenekunnskap .....	100
10.1.1	Feil begrep .....	100
10.1.2	Puttekassen .....	102
10.2	Domenekunnskap som utfordring .....	107

10.3	Oppsummering.....	108
11	Intensjoner.....	110
11.1	Utfordringer knyttet til intensjoner .....	110
11.1.1	Feil/manglende intensjon .....	110
11.1.2	Presisjon .....	113
11.1.3	Kontekstuell relevant informasjon .....	116
11.2	Intensjon som utfordring.....	117
11.3	Oppsummering.....	118
12	Andre utfordringer.....	120
12.1	Tillit.....	120
12.2	Brukere som ikke ser at de snakker med en chatbot.....	124
12.3	Ignorerte ord/Stoppord.....	125
12.4	Oppsummering.....	129
13	Diskusjon.....	130
13.1	Skjulte funksjoner .....	130
13.2	Kategoriseringen av utfordringene .....	132
13.3	Domenekunnskap.....	134
13.4	Chatbotforståelse.....	140
13.5	Brukere som blir satt over.....	141
13.6	Riktige forventinger .....	143
13.7	Sammensatte problemer .....	144
13.8	Designimplikasjoner .....	146
14	Konklusjon .....	149
	Litteraturliste .....	1
	Vedlegg .....	5

## Liste over figurer

Figur 1: Brukergrupper som trenger ekstra informasjonsbistand (NAV, 2018) .....	6
Figur 2: Målbilde for NAV Kontaktsenter (NAV, 2018).....	7
Figur 3: Illustrasjon av samtale med ELIZA(botwiki, 2018).....	12
Figur 4: Illustrasjon av dype nevrale nettverk (Datatilsynet, 2018).....	14
Figur 5: Illustrasjon av nevralt nettverk for å identifisere en hund (Teknologirådet, 2018) ....	14
Figur 6: Skjermbilder fra åpningsfrasene til henholdsvis Kommune-Kari, DNB og Skatteetaten sin chatbot.....	15

Figur 7: Illustrasjon av de åtte utfordringene (Verne, 2015; Verne & Bratteteig, 2016).....	20
Figur 8: Intensjonstre .....	26
Figur 9: Intensjonstre av Frida sin åpningsfrase.....	27
Figur 10: Illustrasjon av Frida sin analyse .....	28
Figur 11: Illustrasjon hentet fra Thakur(2018).....	29
Figur 12: Illustrasjon hentet fra Thakur(2018).....	30
Figur 13: Illustrasjon hentet fra Boost AI(2019f) .....	30
Figur 14: Illustrasjon hentet fra Thakur(2018).....	31
Figur 15: Skjerm bilde av hvordan Frida pleide å se ut (Simonsen, 2019).....	33
Figur 16: Tidslinje .....	41
Figur 17: Koding av første iterasjon.....	44
Figur 18: Koding av andre iterasjon.....	45
Figur 19: Skjerm bilde av chat ifra mai 2019.....	46
Figur 20: Skjerm bilde av en chat i en PDF-fil .....	47
Figur 21: Skjerm bilde av chatter i Excel-fil .....	47
Figur 22: Excel-fil bestående av 18 kolonner med informasjon koblet til hvert utsagn .....	48
Figur 23: Skjerm bilde av utdrag fra chat hvor intensjonen 'Om NAV' forekommer (ID: 125235).....	49
Figur 24: Skjerm bilde av intensjonstre 'Om NAV' .....	50
Figur 25: Illustrasjon av Suchman sitt analytiske rammeverk basert på (Suchman, 2007) s.123 .....	52
Figur 26: Oversikt over utfordringene presentert i Linett sin oppgave (Simonsen, 2019).....	132
Figur 27: Min oversikt over utfordringer .....	133
Figur 28: Skjerm bilde av Frida sin avatar .....	144

## Liste over tabeller

Tabell 1: Datainnsamlingsaktiviteter.....	40
Tabell 2: Chat i Suchman sitt rammeverk .....	57

## Akronoymer

AI	Artificial intelligence/Kunstig intelligens
NAV	Arbeids- og velferdsforvantingen
ML	Maskinl�ring
NSD	Norsk senter for forskningsdata
HCI	Menneske-maskin interaksjon
NKS	NAV Kontaktsenter
DL	Dyp l�ring
NN	Nevralt Nett
DNN	Dypt Nevralt Nett
CA	Conversational Agents/talebaserte assistenter





# 1 Introduksjon

Denne masteroppgaven vil omhandle digitalisering i offentlig sektor og hvilke utfordringer som oppstår ved bruk av chatbot. Oppgaven tar utgangspunkt i en case studie av bruk av chatbot i NAV.

## 1.1 Innledning

Samfunnet vårt blir stadig mer digitalisert. Både måten vi samhandler på og måten informasjon formidles på, overlates stadig mer til automatiserte løsninger. Tidligere var det programmerte, regelstyrte systemer som var normen, men utover 2000-tallet tok maskinlæring over som den rådende tilnærmingen. Ved hjelp av maskinlæring kan datamaskiner lære seg å se sammenhenger og regler ut ifra innsamlet data. Dette gjør at datamaskiner i større grad kan utføre oppgaver på egenhånd og at komplekse oppgaver kan gjøres raskere og rimeligere. (Teknologirådet, 2018)

Fremskritt innen kunstig intelligens og maskinlæring gjør at offentlig tjenester i større grad går over til digitale løsninger og nettbaserte tjenester. Søknadsprosesser og kommunikasjonskanaler automatiseres, og erstatter i større grad de prosessene som vanligvis ville bestått av menneskelig interaksjon. Dette kan føre til mer personlig tilpassede tjenester, men kan også føre til at brukergrupper med lavere teknisk kompetanse vil føle seg ekskludert. (Grönlund, 2007)

Denne masteroppgaven vil fokusere på de utfordringene som oppstår når chatbot tas i bruk som en del av førstelinjetjenesten i en stor offentlig organisasjon. Mange brukere tar kontakt med Arbeids- og velferdsforvaltningen (NAV) fordi de er usikre på om de har krav på en ytelse eller om de har søkt på riktig måte (Ringnes, 2018). NAV er nå i gang med å endre nettsidene sine fra å være et informasjonsnettsted, til å bli en tjenestekanal (Larsen, 2016). Innbyggere som tidligere ble direkte satt i kontakt med en veileder når de kontaktet NAV på chat, blir nå først satt i kontakt med en chatbot. Denne chatboten kan svare på generelle spørsmål og blir stadig bedre til å gi riktige svar, ved å lære ut ifra treningsdata basert på tidligere samtaler. Denne digitaliseringen av kommunikasjonen med NAV, kombinert med økt automatisering av NAV sine tjenester, kan gjøre at enkelte brukere opplever møte med NAV som utfordrende. Hvordan fungerer det når digitale løsninger overtar for menneskelig



## INTRODUKSJON

kontakt? Og hvilke utfordringer oppstår når du skal samhandle med en chatbot i stedet for et menneske? Dette er spørsmål jeg vil undersøke nærmere i denne oppgaven.

### **Arbeidet med oppgaven**

Jeg begynte arbeidet med masteroppgaven min våren 2019 hvor jeg først satt meg inn i den tidligere forskningen gjort rundt automatisering i offentlig sektor og bruk av chatbot, før jeg etter hvert fikk tilgang hos NAV og kunne begynne arbeidet med å gå gjennom chatlogger. Høsten 2019 samarbeidet jeg med Linett Simonsen som også var interessert i hvordan bruken av chatbot i NAV fungerte og hvilke utfordringer som oppstod. I denne perioden gjorde vi intervjuer og observasjoner sammen, og diskuterte det vi oppdaget i chatloggene. Linett skrev kort oppgave og leverte oppgaven sin ‘Når brukerdialogen automatiseres – hva blir vanskelig? En kvalitativ studie av sekvensen av handlinger mellom menneske og chatbot’(Simonsen, 2019) i desember. Utfordringene vi kom frem til her i fellesskap, vil jeg også presentere i denne masteroppgaven, men jeg vil gå nærmere inn på hver av utfordringene og årsakene til disse. I tillegg har jeg også sett nærmere på hvordan det fungerer for brukere som ønsker å bli satt over til en veileder. Forskjellene mellom oppgavene våre vil bli gått nærmere inn på i diskusjonen.

I mars 2020 skrev jeg også en artikkel sammen med Linett Simonsen, Guri Verne og Tone Bratteteig, hvor vi tok for oss utfordringene knyttet til domenekunnskap, som var et av temaene Linett og jeg hadde sett på. I artikkelen “‘I’m disabled and married to a foreign single mother’ Public service chatbot’s advice on citizens’ complex lives’ presenterte vi tre eksempler og skrev om problemene som oppstod på grunn av mangel på domenekunnskap hos brukerne(Simonsen, Steinstø, Verne & Bratteteig, 2020). Jeg har jobbet videre med kategoriseringen av utfordringene og vil i denne oppgaven presentere den kategorisering jeg har kommet frem til, samt diskutere sammensatte problemer og årsakene til utfordringene som blir presentert.

## 1.2 Forskningsspørsmål

Forskningsspørsmålene som undersøkes i denne oppgaven er som følger:

- Hvilke utfordringer oppstår ved bruk av chatbot i NAV?
- Hva gjør at samtaler blir satt over til veileder?

Bruker som vil chatte med NAV blir først henvist til en chatbot, før de får muligheten til å snakke med en veileder. Hvilke utfordringer oppstår i disse chattene med chatbot? Hva er årsakene til dette? Og i hvilke tilfeller blir brukere satt over til veiledere? Studien vil gjennom en analyse av et utvalg av chatlogger ta for seg utfordringene som oppstår og årsakene til dette, med et fokus på de delene av funksjonene til Frida som er skjult for bruker.

## 1.3 Oppgavens struktur

Masteroppgaven er delt inn i 14 kapitler hvor dette introduksjonskapittelet utgjør første kapittel.

I kapittel 2 vil jeg presentere relevant bakgrunnsinformasjon. Dette kapittelet vil presentere konteksten for oppgaven og presentere viktige begreper som vil bli brukt i resten av oppgaven. Dette er delt inn i tre deler; NAV, kunstig intelligens og tidligere forskning.

I kapittel 3 vil jeg redegjøre for metodene som har blitt brukt for å samle inn og analysere data, samt de formelle og etiske aspektene ved studien.

I kapittel 4 vil jeg presentere casen for denne studien. Her vil jeg forklare hvordan NAV sin chatbot Frida fungerer, samt formålet med og utviklingen av denne.

I kapittel 5 vil jeg forklare hvordan vellykkede samtaler med chatboten pleier å foregå, samt hva som gjør at disse samtalene klassifiseres som vellykket.

## INTRODUKSJON

I kapittel 6 vil jeg redegjøre for hva som skjer når samtalene settes over til veileder, hva som er grunnene til dette, og i hvilke tilfeller dette er problematisk.

I kapittel 7 vil jeg presentere oversikten over kategorier av utfordringer jeg har kommet frem til.

I kapittel 8-11 vil jeg presentere utfordringene som oppstår ved bruk av chatbot i NAV. Jeg har delt utfordringene inn i fire kategorier som presenteres i hvert sitt kapittel. Kapittel 8 vil ta for seg utfordringer knyttet til Språk, herunder *feilidentifisert språk* og *stavekontroll*. Kapittel 9 tar for seg kategorien Chatbotforståelse, som omfatter *personlige spørsmål* og *oppfølgingsspørsmål*. Kapittel 10 handler om Intensjonene og viser hvordan det oppstår problemer ved *feil eller manglende intensjon*, *presisjon* og *kontekstuell relevant informasjon*. Kapittel 11 handler om Domenekunnskap som omfatter bruk av *feil begrep* og *puttekassen*. Og kapittel 12 tar for seg andre utfordringer, slik som *tillit*, *ignorerte ord* og *brukere som ikke ser at de snakker med en chatbot*.

I kapittel 13 vil funnene og analysen fra de foregående kapitlene bli diskutert og satt i sammenheng med tidligere relevant forskning. I dette kapittelet vil også alternative designgrep bli presentert.

I kapittel 14 vil jeg oppsummere og konkludere, og komme med forslag til fremtidig forskning.

## 2 Bakgrunn

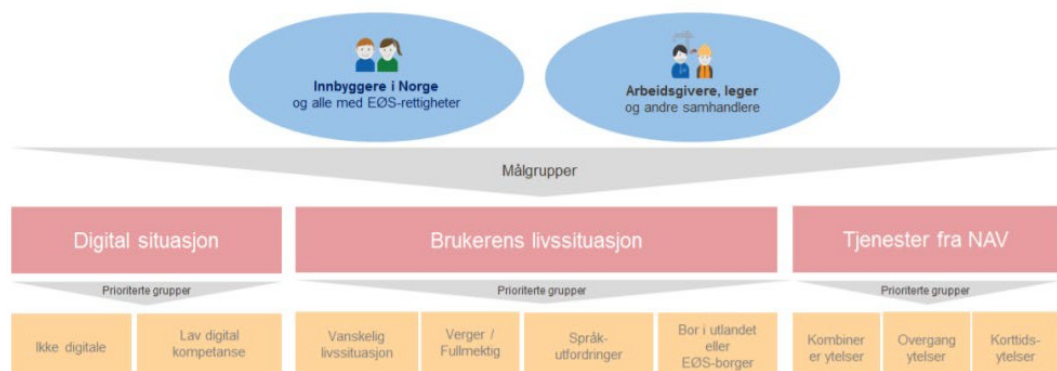
I dette kapitlet vil nøkkelbegreper og bakgrunnsinformasjon som er relevante for oppgaven bli presentert. Kapitlet er inndelt i tre deler. Del én vil handle om NAV og ta for seg NAV og digitalisering i NAV. Del to vil handle om kunstig intelligens, herunder maskinlæring, nevralt nettverk og chatbot-teknologi. Del tre vil dreie seg om tidligere forskning på automatisering i offentlig sektor og relevante studier som tar for seg bruk av chatboter.

### 2.1 NAV

#### 2.1.1 Om NAV

NAV (Arbeids- og velferdsforvaltningen) er en landsdekkende offentlig virksomhet som består av både kommunale og statlige tjenester. NAV har rundt 19 000 medarbeidere, flere av dem er på ulike spesialenheter i tillegg til 456 vanlige NAV-kontor. Målet til NAV er å få flere i arbeid og aktivitet, gi riktige støtteordninger til riktige tider og gi god service tilpasset brukeren sine forutsetninger og behov. (NAV, 2020) NAV har 127 millioner henvendelser hvert år, forvalter 60 ulike ytelser og har over 300 ulike IT-systemer. (Vågeng, 2018) NAV sine brukere er mennesker som har rett til ulike former for støtteordninger. Mange av brukerne er også i nye eller vanskelige livssituasjoner. NAV sine tjenester og kommunikasjonsformer må derfor være tilpasset flere ulike brukergrupper med forskjellige forutsetninger.

NAV gir også ekstra informasjonsbistand til prioriterte grupper.



Figur 1: Brukergrupper som trenger ekstra informasjonsbistand (NAV, 2018)

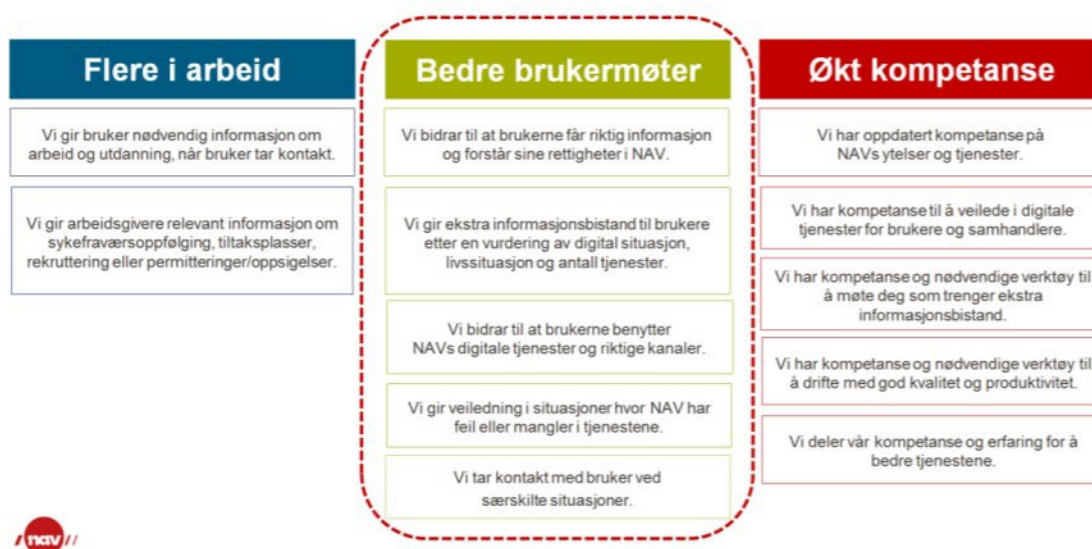
## BAKGRUNN

Disse gruppene er klassifisert av NAV som grupper som trenger ekstra informasjonsbistand. Dette kan være lav digital kompetanse, vanskelige livssituasjoner, språkutfordringer eller type ytelse.

### 2.1.2 NAV Kontaktsenter

NAV Kontaktsenter(NKS) er en førstelinjetjeneste samlet under arbeids- og tjenestelinjen. Førstelinjen i NAV består av kanalene nav.no, telefon, chat, spørsmål-svartjeneste, sosiale medier og personlig oppmøte. Her kan brukerne få informasjon, status eller veiledning om ytelser og tjenester knyttet til arbeids- og tjenestelinjen, ytelseslinjen eller økonomilinen. Kontaktsentrene startet som en ren telefonitjeneste, men de har gradvis utvidet sitt tjenestetilbud, både av faglig spesialisering og tjenester. I tillegg til telefon har de nå flere kanaler hvor de kan nå ut til brukerne, slik som innlogget spørsmål/svar-tjeneste, chat og sosiale medier. NKS består av 18 kontaktsentre, NAV Servicesenter og en styringsenhet. I dag har de ca 1000 medarbeidere fordelt på 20 enheter, hvor enkelte også har flere lokasjoner.(NAV, 2018)

Den høyeste prioriterte oppgaven i NAV kontaktsenter er bedre brukermøter.



Figur 2: Målbilde for NAV Kontaktsenter (NAV, 2018)

Oversikt som viser målsetningene til NAV kontaktsenter. De skal bidra til at brukerne får riktig informasjon, ekstra informasjonsbistand, bidra til at brukerne benytter digitale tjenester

og riktige kanaler, veiledning ved feil eller mangler i tjenestene og ta kontakt ved særskilte endringer i brukers situasjon.

### **2.1.3 Digitalisering i NAV**

NAV har de siste årene hatt en større satsing på digitalisering. Dette medfører større effektivitet, men større grad av selvbetjening kan også gjøre det mer krevende å finne ut hva man har krav på fra NAV, særlig når man allerede er i en ny eller utfordrende livssituasjon. I en stortingsmelding fra 2016 ble det vedtatt at digital samhandling skal være hovedregelen for kontakt mellom bruker og forvaltning. (Regjeringen, 2016) Det pågår flere IT-prosjekter i NAV hvor de automatiserer ulike velferdsordninger. I 2018 ble digital sykemelding lansert og det pågår stadig prosjekter for mer digitalisering og automatisering (Vågeng, 2018). IT-utviklingen hos NAV er et kontinuerlig prosjekt hvor ytelser digitaliseres gradvis og flyttes bort ifra den gamle løsningen. Det handler om å effektivisere hverdagen for brukerne, både for arbeidstakere og arbeidsgivere. (Vågeng, 2018)

## **2.2 Kunstig intelligens**

### **2.2.1 Kunstig intelligens**

Kunstig intelligens vil si at en datamaskin er i stand til å løse oppgaver uten å få instruksjoner fra et menneske på hvordan det skal gjøres (Tidemann, 2020). Begrepet 'kunstig intelligens' har vært en del av datateknikken siden 1950-årene (Elster & Tidemann, 2019), først med Alan Turing sin artikkel om intelligent adferd.

#### **The Turing test**

Starten på diskusjonen rundt intelligente maskiner og bruk av kunstig intelligens begynte med utviklingen av 'The Turing test' av Alan Turing. Dette var en test som ble beskrevet i artikkelen 'Computer Machinery and Intelligence' som skulle teste om datamaskiner kunne utøve intelligent adferd. Artikkelen tok for seg spørsmålet 'kan maskiner tenke?', og beskrev følgende eksperiment'; Hvis et menneske ikke klarer å si hvorvidt den man snakker med er et menneske eller en datamaskin, fordi maskinen klarer å overbevise om at den er et menneske, så er maskinen intelligent (Turing, A, 1950)

### **Utbredelse av kunstig intelligens**

Det er særlig siden 2010-tallet at teknologiske fremskritt har ført til at denne type teknologi har fått en større utbredelse.

Det var et vendepunkt i 2012, da Alex Krizhevsky utviklet en metode som senket feilraten på bildekjenning ved hjelp av en teknikk basert på nevralt nettverk. Etter dette ble nevralt nettverk også brukt til tolkning av sekvenser for lyd og tekst. Nevrale nettverk hadde eksistert i flere tiår, men det var først i 2012 at man hadde maskiner som var så kraftige at de kunne trene opp store modeller relativt raskt, og at man hadde store nok datamengder for å trene opp slike modeller. (Tidemann, 2020)

Nå er kunstig intelligens blitt en sentral del av utviklingen innen informasjonsteknologi og introduseres stadig i nye felt. Eksempler på fagfelt hvor kunstig intelligens er i fokus er talegjenkjenning, bildegjenkjenning, brukerinteraksjon og styring av fysiske prosesser. (Elster & Tidemann, 2019) Dette er fagfelt hvor kunstig intelligens kan bidra til en mer persontilpasset opplevelse for brukeren, ved å for eksempel tilpasse seg til brukerens måte å snakke på, eller sortere og tagge bilder.

Ved å samle inn data om hver enkelt innbygger, kan man også gjøre tjenester mer tilpasset den enkelte innbyggers behov. Dette gjør at man kan tilby tjenester som persontilpasset behandling, tilpasset undervisning, skreddersydde tilbud og individuelle treningsprogram. (Teknologirådet, 2018) Men denne måten å tilpasse tjenester basert på innsamlet data, innebærer også en rekke etiske dilemmaer. Store mengder av personopplysninger kan lett misbrukes om formålet med innsamlingen av opplysningene ikke er tydeliggjort. Historiske data og forutinntatte treningsdata kan også føre til at beslutninger tas ut ifra fordommer, heller enn et representativt datasett.

### **Representative data**

Dataen som brukes til å trene opp maskinen, vil styre hva maskinen lærer seg. Er treningsdataen forutinntatt, vil maskinen også bli det. Det er derfor ekstremt viktig med en stor nok datamengde med riktige og representative data.

Microsoft sin chatbot Tay, er et eksempel på hvor galt det kan gå, når en chatbot lærer basert på treningsdata som ikke er kvalitetssikret. Den skulle lære av brukerne den snakket med, og

## KAPITTEL 2

endte opp med å bli rasistisk og fiendtlig (Hunt, 2016). Et annet eksempel er et system for utmåling av straff og kausjonsbetingelser i USA som diskriminerte afroamerikanere og beregnet det som dobbelt så stor sjanse for at en afroamerikaner ville begå et lovbrudd. Slike systemer kan få store konsekvenser ved at strengere straffer gis på bakgrunn av etnisitet i stedet for reelle faktorer som er knyttet til gjentagende kriminalitet, noe som kan føre til at farlige kriminelle blir sluppet ut tidligere, og at de som i realiteten har lavere sjanse for gjentagelse får lengre straffer (Angvin m.fl., 2016) Andre situasjoner hvor kunstig intelligens kan føre til problemer, er annonser ifra Google. Noen jobban annonser vises hovedsaklig til menn fordi de har en større sannsynlighet for å søke på jobben, ut ifra tidligere data. (Gibbs, 2015) Kvinner får derfor ofte ikke opp annonsene og får heller ikke søkt, noe som igjen fører til flere mannlige søkere og mer data som kalkulerer menn som riktige mottagere for disse annonsene. Disse tilfellene viser derfor at vi må tenke nøye gjennom hva vi inkluderer i treningsdata.

### **Personvern**

Bruk av kunstig intelligens i slike tjenester stiller også spørsmål til personvern. Ved at maskinen «lærer» informasjon om brukeren for en brukertilpasset opplevelse, så lagrer den også veldig mye informasjon som kan være nyttig til andre kommersielle tjenester slik som personlig reklame. Men dette kan også føre til større grad av overvåking og lagring av informasjon på tvers av tjenester.(Barland, 2016)

Behandling av personopplysninger skal gjøres i henhold til gjeldende lovverk. I mai 2018 trådte regelverket GDPR(General Data Protection Regulation) i kraft(Regjeringen, 2018). Kort oppsummert så handler dette om at behandling av personopplysningene er gjennomsiktig og korrekt, kun brukes til angitte formål og dataminimering(kun samle inn det som er relevant og nødvendig for formålet). En personopplysning er alt som kan knyttes til deg som enkeltperson(Datatilsynet, 2019). For at behandlingen av dataopplysningene skal være i henhold til lovverket, må brukerne bli gjort oppmerksomme på hvordan dataene behandles, hva den skal brukes til, hvor lenge dataene vil være lagret og hvordan de kan selv kan få innsikt i og korrigerer egne opplysninger og tilgang til å slette dem.(Datatilsynet, 2018) Samtykket til lagring av personopplysninger må være frivillig, spesifikk, informert og utvetydig.



## BAKGRUNN

Men særlig prinsippene om rettferdighet, formålsbegrensning, dataminimering og gjennomsiktighet blir utfordret ved bruk av kunstig intelligens (Datatilsynet, 2018). Ved bruk av kunstig intelligens kreves store mengder data og tidligere data som er samlet inn kan være nyttig til flere formål enn hva det opprinnelig ble samlet inn til. Det kan derfor være vanskelig å definere akkurat hvilke opplysninger som vil være nødvendige og hvilke formål man vil trenge de til. Dette har blant annet ført til utfordringer ved behandling av helseopplysninger. Helseopplysninger går under sensitive data, noe som gjør at disse behandles under særskilte regler. (Datatilsynet, 2019) Teknologien vil stadig utvikle seg og kunstig intelligens kan brukes til stadig nye formål, da er det en utfordring at dataen som er nødvendig ikke alltid er tilgjengelig fordi regelverket gjør det umulig å ta i bruk disse opplysningene.

Prinsippet om gjennomsiktighet kan være vanskelig å oppfyllet i forbindelse med kunstig intelligens. Avanserte former for kunstig intelligens kan være vanskelig både å forstå og forklare, og det kan derfor være vanskelig å få formidlet hvordan opplysninger blir koblet og vedlagt i en spesifikk behandling. Kunstige intelligente systemer som selv identifiserer nye mønstre og sammenhenger, kan også være lite gjennomsiktige, som igjen kan gjøre ansvaret utydelig. Når maskinlæringsalgoritmer i økende grad tar beslutninger på områder som har betydning for innbyggernes livskvalitet kan det bli problematisk om de ansvarlig ikke kan forklare hvordan eller hvorfor en beslutning er tatt. (Teknologirådet, 2018)

### 2.2.2 Chatboter

En chatbot er et program som imiterer menneskelig kommunikasjon og gjør det mulig å kommunisere via naturlig språk. Ordet chatbot er satt sammen av to ord. 'Chat' som i nettpat og 'bot' som i robot. (SINTEF, 2019) Dette er altså en form for kommunikasjon som foregår enten skriftlig eller muntlig ved bruk av kunstig intelligens, hvor man fører en samtale med en maskin.

Det er to måter å lage en chatbot på; som en regelbasert chatbot hvor hardkodete kommandoer kobles til ulike nøkkelord, og maskinlæringsbaserte chatboter hvor chatbotene analyserer utsagnet fra brukeren og lærer over tid. (Tidemann, 2020)

## KAPITTEL 2

‘Eliza’ regnes som den første chatboten og ble utviklet av Joseph Weizenbaum ved MIT i perioden 1964-1966(Ireland, 2012). Dette var en av de første maskinene som kunne forsøke ‘the Turing test’. Eliza var en regelbasert chatbot, som svarte med fraser ut ifra nøkkelord som brukeren kom med(Weizenbaum, 1966).

```
Welcome to
          EEEEE LL      IIII ZZZZZZ  AAAA
          EE      LL      II      ZZ  AA  AA
          EEEEE LL      II      ZZ  AAAAAA
          EE      LL      II      ZZ  AA  AA
          EEEEE LLLLL IIII ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:   █
```

Figur 3: Illustrasjon av samtale med ELIZA(botwiki, 2018)

Weizenbaum utviklet Eliza for å demonstrere hvor overfladisk samtaler mellom maskiner og mennesker er. Eliza skulle etterligne en psykoterapeut og stilte spørsmål som matchet med brukerens utsagn. (Ireland, 2012). Dette ga en illusjon av at programmet forstod hva som ble sagt selv om den ikke kunne sette ting inn i kontekst eller forstå sammenhengen. Mange brukere følte allikevel at Eliza kunne forstå dem og ble følelsesmessig engasjert i samtalen. (Weizenbaum, 1966)

### 2.2.3 Maskinlæring

Maskinlæring vil si at statistiske metoder er brukt for å finne mønstre i store datamengder. Maskinen trenes opp ved å bruke et stort datasett som er delt opp i et treningssett og et testsett. Treningssettet er reglene for hvordan maskinen skal respondere. Testsettet brukes for å sjekke at maskinen har «lært» det den skal, og at den klarer å sortere ut riktige responser selv om testsettene varierer fra treningsdataen. (Elster & Tidemann, 2019) Ved at maskinen «lærer» selv på denne måten, klarer den selv å finne mønstre på hvordan for eksempel en samtale skal foregå, uten at man trenger å hardkode hver respons til et spesifikt brukerutsagn.

Et eksempel på dette vil være om en bruker spør om lokasjon ved å skrive 'Hvor ligger kontoret deres?'» I et program som er hardkodet vil det være programmert at utsagnet 'hvor ligger kontoret deres' skal kobles opp mot 'dette er adressen'. Ved maskinlæring, derimot, så kan maskinen lære seg at forskjellige former som 'hva er adressen' og 'hvor holder dere til' og skal kobles opp mot 'dette er adressen'.

### **2.2.4 Dype nevralt nettverk**

Nevralt nettverk er «en samlebetegnelse for datastrukturer, med tilhørende algoritmer, som er inspirert av måten nerveceller i en hjerne er organisert på» (Tidemann, 2019). Dette kan brukes til å identifisere sammenhenger og påvirkningskraften forskjellige faktorer vil ha i en avgjørelse. Datastrukturer som dette kan derfor brukes til å klassifisere mennesker ut ifra handlingsmønstre, for å for eksempel anslå hvordan en reklame vil slå an hos en bruker eller til bruk ved etterretningsvirksomhet.(Tidemann, 2019)

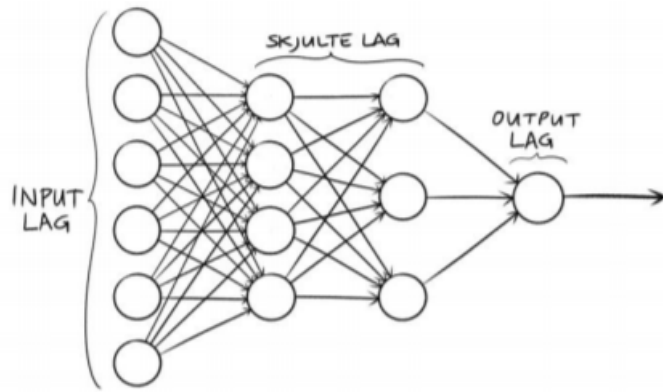
### **Dyp læring**

Dyp læring vil si å trene opp såkalte 'dype kunstige nevralt nettverk'. Læremekanismen i dyp læring er intuitiv, det vil si at maskinen skal kunne resonnerer seg frem til løsninger på nye problemer basert på eksisterende data. For at maskinen skal klare å resonnerer seg frem til et riktig svar, må maskinen derfor gjennom mange eksempler som fungerer som treningsdata.(Tidemann, 2018)

Ved dyp læring er det mange lag av nevroner og synapser. Nevroner påvirker hverandre via synapser. Den synaptiske vekten styrer hvor mye et nevron påvirker et annet. Læreprosessen foregår ved at man sender store datamengder gjennom det nevralt nettverket og justerer vektene underveis for å få et mer presist resultat. (Tidemann, 2018)

Et nevralt nettverk vil derfor bestå av tre deler; et input lag, et eller flere skjulte lag, og et output lag.(Datatilsynet, 2018)

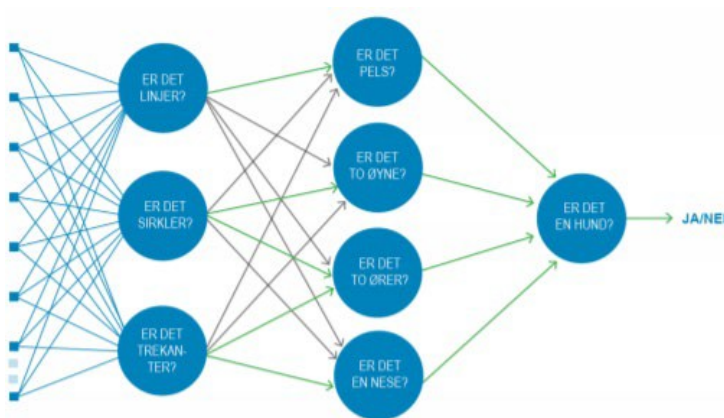
## KAPITTEL 2



Figur 4: Illustrasjon av dype nevralt nettverk (Datatilsynet, 2018)

Her kan man se hvordan de ulike lagene henger sammen. Hvis det er mer enn ett skjult lag, regnes det som dyp læring.

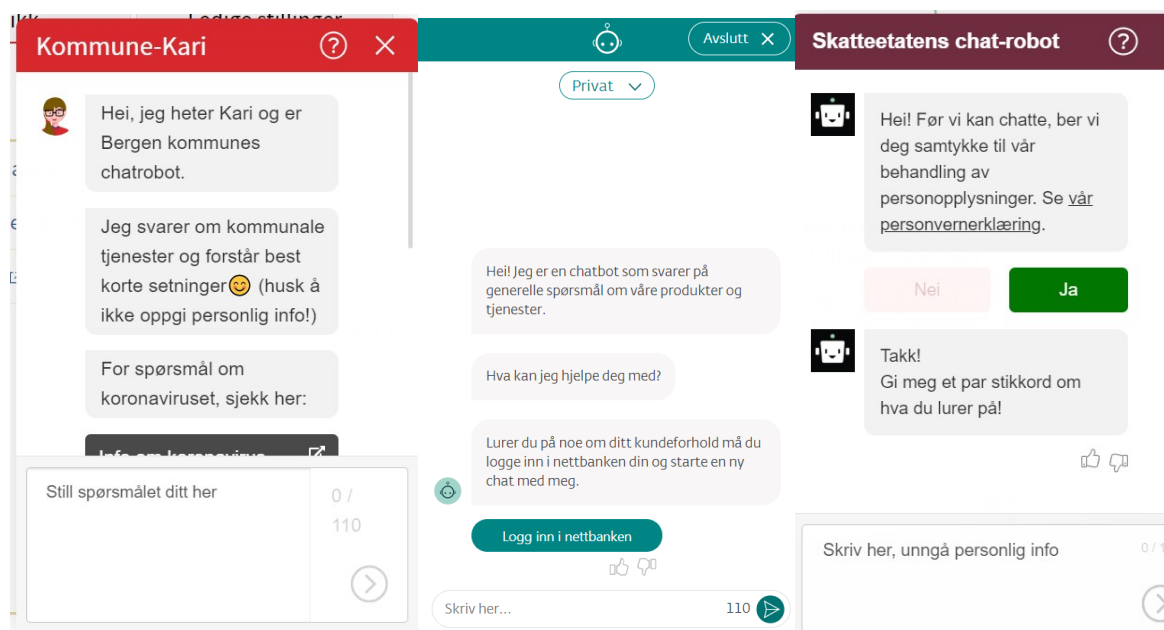
Hvis en maskin for eksempel skal klare å klassifisere hva som er i et bilde, vil dataen bestå av bilder hvor et menneske har lagt inn data om hva som er på bildene. Deretter kan man presentere bilder til maskinen, testdata, og se om den klassifiserer det riktig ut ifra treningsdataen. Hvis datamaskinen for eksempel har tilgang på tusenvis av bilder av dyrearter med informasjon om hvilken art det er, vil den etter hvert klare å generalisere slik at den kan si hvilken dyreart som befinner seg i nye bilder den bli presentert for. På denne måten «lærer» den seg å klassifisere dyrearter. Men for å få til dette krever det store mengder treningsdata. Det holder ikke bare med å memorere det, den må «forstå» hva som utgjør de ulike dyreartene.



Figur 5: Illustrasjon av nevralt nettverk for å identifisere en hund (Teknologirådet, 2018)

## 2.2.5 Chatbot som informasjonsformidler

Chatboter kan brukes til flere formål, slik som underholdning og sosiale og relasjonelle faktorer, men de brukes også av offentlige tjenester til å formidle informasjon og gjøre det lettere å finne frem til den tjenesten man ønsker. (Brandtzæg & Følstad, 2017) Stadig flere offentlige tjenester tar i bruk chatboter til dette formålet for å øke brukervennlighet, blant disse er NAV og Skatteetaten. Også banker som DNB og Nordea tilbyr kundene mulighet til å snakke med en chatbot. Flere kommuner har også tatt i bruk en chatbot kalt Kommune-Kari. Disse skal alle hjelpe brukerne med å finne frem, men har også alle til felles en utfordring i at chatboten skal kunne brukes av en stor brukergruppe med varierende domenekunnskap og chatbotforståelse.



Figur 6: Skjermbilder fra åpningsfrasene til henholdsvis Kommune-Kari, DNB og Skatteetaten sin chatbot.

Kommune-Kari er en chatbot som skal kunne svare på spørsmål om kommunale tjenester og brukes i flere av landets kommuner. Hun kan svare på korte fraser og oppgi valgalternativer for brukeren om hun er usikker på hva hun blir spurt om. Men denne chatboten har fått mye kritikk. I flere tilfeller misforstår hun hva brukeren mener på grunn av nøkkelord og kommer med villedende informasjon. Professor Kai A. Olsen trekker frem flere eksempler på dette: Da Kommune-Kari ble spurt om røykvarsler anbefalte hun røykeavvenning, og da hun ble spurt om hvordan man kan få slått plenen, henvendte hun brukeren til voldsavdelingen hos politiet

## KAPITTEL 2

på grunn av ordet 'slå'. (Olsen, 2018) Men de eksemplene som kommer frem her ser ut til å ha blitt rettet opp i etterkant. Kommune-Kari lærer av å prate med folk og utvikler seg stadig.

Chatboten til DNB er laget for å kunne svare på henvendelser fra DNB sine kunder. Den kan snakke med 3000 brukere samtidig og det er derfor langt flere som kan få svar på sine spørsmål til enhver tid. I mange tilfeller må den sette over til en rådgiver, men den klarer ifølge DNB's talsperson å svare på over halvparten av henvendelsene på egenhånd. Det jobbes kontinuerlig med å forbedre chatboten, ved å se på hvilke tilfeller chatboten ikke ga godt nok svar og trene den på hva den skal svare i slike tilfeller. (Strzelecki, 2018) Chatboten har kun et ikon, i stedet for en avatar, og den har heller ikke noe navn.

Skatteetaten lanserte sin chatbot høsten 2019. Den skulle i første omgang dekke domeneene 'Skatt for person' og 'Folkeregister'. Denne chatboten har i likhet med DNB sin, ikke noe navn eller avatar. Siden den trengte en del trening før den kunne fungere optimalt, ble det foretatt en såkalt stille lansering. Funksjonen for å spørre en chatbot ble lagt til som et alternativ på undersiden for Skatt og Folkeregister. Når man bruker denne chatboten, må man godta en personvernerklæring før man kan starte samtalen.

Disse tre chatbotene skal kunne gjøre det lettere for kunder å finne informasjon, men det er flere situasjoner som setter dem fast og gjør at brukeren enten blir henvendt feil sted, eller videresendt til et menneske i stedet. Alle tre er knyttet til et spesifikt domene; henholdsvis en kommune, DNB og Skatteetaten. Hver av chatbotene kommer med alternativer om de er usikre på hva brukeren spør om, og prøver å veilede brukeren frem til en spesifikk tjeneste eller sted på nettsiden.

### **2.3 Tidligere forskning**

Jeg har ikke funnet noe forskning om bruk av chatboter i praksis som undersøker de tingene jeg ser på i denne studien. Jeg vil derfor i stedet presentere studier om chatboter og talebaserte assistenter i andre situasjoner og studier som fokuserer på kontaktsenter hvor relevante begreper og fenomener blir diskutert.

#### **2.3.1 Bruk av kunstig intelligens i dagliglivet**

## BAKGRUNN

Det har i løpet av de siste ti årene vært en økning i bruk av såkalte conversational agents/talebaserte assistenter(CA) i dagliglivet. Denne formen for menneske-maskin interaksjon(HCI) fungerer som et dialogsystem med tilnærmet menneskelig egenskaper. Apple sin CA Siri, som kom i 2011, og Google Now som kom i 2012, viser at konversasjon stadig blir en større del av menneske-maskin interaksjon.(Luger and Sellen, 2016) Siri kan aktiveres ifra en apple-enhet og kan brukes som et grensesnitt for informasjonsinnhenting. Alle henvendelser til Apple er prosessert gjennom Apple sine servere slik at Apple har store mengder data som kan brukes til å forbedre tjenesten.(Aron, 2011) Google Home, som er en del av Google Assistant, aktiviseres ved bruk av nøkkelord og kan brukes som et kontrollsenter for hjemmet ditt, i tillegg å spille av musikk og videoinnhold.(Tillman, 2019)

En studie fra 2016 tok for seg hvordan forventingene brukerne har til denne type teknologi påvirker bruken og i hvor stor grad brukerne er fornøyde med denne måten å interagere på. Ofte er det et skille mellom forventingene brukerne har til CA og dens faktiske egenskaper og mål. Mange brukere hadde mentale modeller som ikke stemte med systemets funksjoner. Studien viste at når systemet ikke gjorde det brukeren forventet, endte ofte brukeren opp med å gi opp det de prøvde å gjøre. Mangel på tilbakemeldinger fra systemet gjorde det også vanskelig å forstå hvorfor det ikke fungerte og hva de måtte gjøre for at det skulle fungere. Brukere med lav teknisk kompetanse ville derfor tidlig gi opp bruken av talebaserte assistenter. Brukere med mer kunnskap om CA, derimot, var mer tilbøyelige til å prøve igjen, med bruk av kortere setninger og bruk av nøkkelord, ettersom at de forstod mer om hvordan systemet tolket det de sa. Det er derfor viktig at brukerne vet hva systemet er i stand til og hvordan det fungerer, ved å vise systemets egenskaper og bruk av feedback. (Luger&Sellen, 2016)

Det ble gjort en studie av chatboten Cleverbot som sammenlignet samtaler mellom chatboter og mennesker med samtaler mellom mennesker. Denne studien viste at det var flere forskjeller i hvordan folk kommuniserte med andre mennesker i forhold til med chatboter ved å sammenligne variabler som antall ord per samtale og antall meldinger per samtale. Brukerne i studien sendte dobbelt så mange meldinger i samtale med chatboter, sammenlignet med chattene med mennesker. Studien viste at brukerne ikke var mindre komfortable med chatboter, sammenlignet med mennesker. (Hill, Ford & Farreras, 2015). En annen studie viste også at brukerne var mindre redde for å stille spørsmål til chatboter og mindre bekymret for hvordan det de skrev ble oppfattet(Følstad, Nordheim & Bjørkli, 2018).

### 2.3.2 Automatisering av offentlige tjenester

I 2007 ble det gjort en studie rundt hvordan elektroniske tjenester påvirker kravene som ble stilt til innbyggerne. Bruk av offentlige tjenester vil føre til at visse krav stilles til innbyggerne i form av kunnskap og ferdigheter. Når metodene automatiseres, vil disse kravene endres. Bruk av elektroniske tjenester og automatiserte løsninger vil derfor føles mer inkluderende for dem som opplevde den opprinnelige, mer manuelle prosessen, som komplisert og krevende, men det vil samtidig føles ekskluderende for de med lavere teknisk kompetanse. (Grönlund, 2007)

#### Bruk av kanaler

Dette gjelder også de kommunikasjonskanalen man kan bruke ved kontakt med offentlige tjenester. Skaarup sin studie fra Danmark viser hvordan de forskjellige kanalene påvirker følelsen av kontroll, og følelsen av å bli hørt og bli tatt på alvor. Ansikt-til-ansikt gjorde at brukerne følte større grad av kontroll, og større mulighet for feedback underveis. Men nettsider kan fungere bra når du har konkrete spørsmål og ikke er i tvil om du har forstått noe riktig. (Skaarup, 2012) Dette viser at det kan være mange faktorer som påvirker valg av kanal, og at forskjellige kanaler egner seg til forskjellige type problemer. Å avklare misforståelser er også ofte lettere å gjøre gjennom dialog, enten over telefon eller ansikt-til-ansikt. Da kan brukeren fortelle om sin tolkning og få svar på om det er riktig, sammenlignet med om brukeren bare kan trykke på 'jeg forstår' på nettsiden, uten nødvendigvis å ha forstått det som står der. (Ebbers, 2008)

De forskjellige formene for henvendelser kan ifølge Ebbers(2008) deles inn i fire typer: allocution, registration, consultation og conversation. 'Allocution' vil si å gi ut informasjon til innbyggerne. 'Registration' vil si at innbygger sender informasjon til organisasjonen, fordi organisasjonen krever det. 'Consultation' er en handling hvor en innbygger henter informasjonen selv fra nettsiden. I motsetning til en interaksjon så er det her innbyggeren selv som finner frem til informasjonen. Og til sist 'Conversation' som vil si at en bruker etterspør informasjon og får spesialtilpasset informasjon ut ifra sine behov. Som for eksempel ved å ringe og stille et spørsmål.

Valget av kanal vil derfor avhenge av både hva slags type henvendelse det er, og hva slags behov brukeren har. Også kompleksiteten til problemet og situasjonen rundt vil ha en



påvirkning på dette valget, ved siden av ferdighetene og kunnskapen til bruker.

Situasjonsbetingede faktorer kan være: klokkeslettet(om det er innenfor åpningstiden), hvor man befinner seg(hvilke kanaler man har lettest tilgang til) og behovet for bekreftelse for å føle seg ferdig med temaet.

### **Utfordringer ved automatisering**

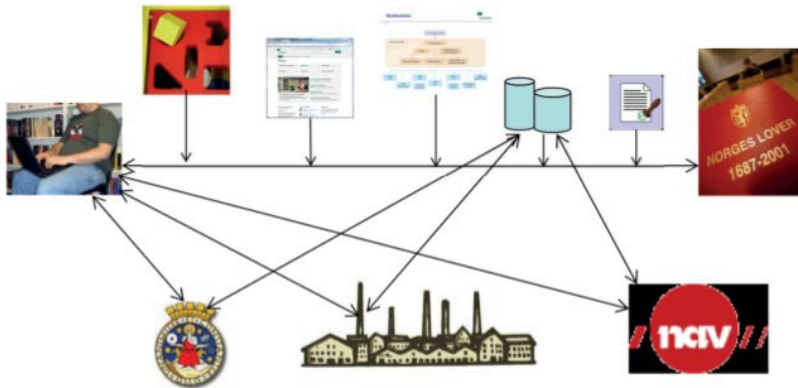
Guri Verne undersøkte i sin doktoravhandling forholdet mellom mennesker og maskiner og automatisering av offentlige tjenester. (Verne, 2015) Denne studien viste at unge skattebetalere, som har lite erfaring med og kjennskap til prosessen rundt skattemeldingen, ofte ringer til Skatteetaten for å få hjelp. Dette kan skyldes at denne brukergruppen kan ha vanskeligheter med å forstå oppbyggingen av det digitale skjemaet, ettersom at de ikke har erfaring med å fylle ut skattemeldingen fra før den ble automatisert. (Verne, 2015; Verne & Bratteteig, 2016.) I artikkelen Do-it-yourself Services and Work-like Chores: On Civic Duties and Digital Public Services (Verne & Bratteteig, 2016) så argumenteres det også for at muligheten til læring og forståelse burde inkluderes i designet av automatiske løsninger.

Verne presenterte i sin studie åtte ulike utfordringer som kan oppstå når innbyggerne skal forholde seg til økende grad av automatisering. Disse åtte var:

1. Omstendighetene i livet til en innbygger: En innbygger kan ha en krevende livssituasjon som kan ha en påvirkning på skatten, men som også kan påvirke innbyggernes evne til å fylle ut skattemeldingen.
2. «Puttekassen»: Innbyggeren har problemer med å tolke og klassifisere egne situasjoner og omstendigheter.
3. Bruk av nettbaserte løsninger: Innringerne finner ikke informasjonen og opplever tjenestene som vanskelige å bruke.
4. Interne strukturer: Noen ganger kan forståelse for indre strukturer hos skattemyndighetene være nødvendig.
5. Tekniske utfordringer og anomalier: Både brukerfeil og tekniske feil kan skje, og det kan være vanskelig å skille mellom dem.
6. Manuelle oppgaver og dokumentasjon: Noen ting må dokumenteres manuelt, og er ikke inkludert i den automatiserte løsningen.
7. Lover og regelverk: Lover og regelverk trenger forklaring, tolkning og tilpasning til

innringerens situasjon.

8. Interaksjoner med tredjeparter: Interaksjoner med blant annet andre offentlige etater kan komplisere situasjonen for innbyggeren.



Figur 7: Illustrasjon av de åtte utfordringene (Verne, 2015; Verne & Bratteteig, 2016).

### 2.3.3 Bruk av chatbot i offentlig sektor

I 2018 ble det gjort en studie av to chatboter tilknyttet Skattetaten (Arlov, 2018). Her ble det sett på hva slags spørsmål det lønner seg å overlate til en chatbot og hvor mye av trafikken chatboten kan håndtere. Ved hjelp av brukertester og veilederintervjuer, kom de fram til at disse punktene var viktige for en vellykket chatbot-samtale:

1. Man må skrive kort. Et chatvindu har begrenset med plass og det er ikke alltid brukerne er klar over at de må scrolle opp for å se hele teksten. Lange svar må fordeles over flere interaktive vekslinger.
2. Ikke bruk fagtrykk. En chatbot skal kunne brukes av brukere uavhengig av domeneskunnskap om dette aktuelle fagfeltet.
3. Unngå generelle svar. Det er viktig at brukeren føler at spørsmålet har blitt riktig oppfattet. Om chatboten bare kommer med generelle svar kan brukeren bli usikker på om dette gjelder brukerens tilfelle eller bli forvirret over for mange valgalternativer.
4. Interaktive samtaler med bruk av knapper/lenker gjør at brukerne føler seg fornøyde og trygge.

## BAKGRUNN

Artikkelen til Amershi, Weld, Vorvoreanu, Fourney, Nushi, Collisson, Suh, Iqbal, Bennet, Inkpen, Teevan, Kikin-Gil og Horwitz, (2019) presenterte også en rekke retningslinjer for bruk av systemer med kunstig intelligens. Denne artikkelen presenterte totalt 18 retningslinjer og mange av dem dreier seg om systemer man bruker over tid, men jeg vil gå gjennom de som er mest relevante for denne masteroppgaven her.

### **G1: Gjør det klart hva systemet kan gjøre**

Dette handler om å gjøre det tydelig hva systemet er i stand til. Funksjonene til systemet skal være synlige for bruker, slik at brukeren vet hva man kan bruke systemet til.

### **G2: Gjør det klart hvor bra systemet kan gjøre det det gjør.**

Gjør det klart for bruker hvor bra systemet kan gjennomføre oppgaver, og hvor stor sjanse det er for mulige feil.

### **G4: Vise kontekstuell relevant informasjon**

Vis informasjon som er relevant for tidspunkt, lokasjon og den nåværende handlingen til brukeren.

### **G11: Gjør det klart hvorfor systemet gjorde det det gjorde**

Gi brukeren forklaring på hvorfor systemet oppførte seg slik som det gjorde, sånn at brukeren skjønner hvorfor systemet endte opp med et gitt resultat.

## **2.4 Oppsummering**

I dette kapittelet har jeg gjort rede for konteksten for casen denne studien dreier som og viktige begreper som vil bli brukt i denne oppgaven. Ettersom at det er gjort lite forskning på bruk av chatboter i praksis har jeg presentert studier fra kontaktsentre som diskuterer relevante begreper og fenomener.

## KAPITTEL 2

## 3 Chatboten Frida

I dette kapitlet vil jeg presentere casen denne studien fokuserer på. Jeg vil gjøre rede for formålet ved bruk av chatbot hos NAV, hvordan chatboten fungerer og hvordan det jobbes med og utvikle den. Starten av dette kapitlet er skrevet delvis sammen Linett Simonsen.

### Formål med chatbot

Ifølge chatbot-trenerne er formålet med chatboten å øke brukeropplevelse og å effektivisere. Siden en chatbot kan gi raske svar på enkle spørsmål, vil brukere med slike spørsmål kunne høste gevinst av å konversere med en chatbot fremfor en veileder. Gevinsten vil eksempelvis være i form av spart tid, da bruker slipper å sitte i telefonkø. En annen fordel med chatbot er at hjelp vil være tilgjengelig hele døgnet. Ifølge chatbot-trenerne vil en chatbot også kunne gjøre oppgavene til veilederne mer spennende. Ved innføring av chatbot vil disse kunne slippe å svare på enkle og repetitive spørsmål.

### Bakgrunn

Chatboten startet som et prosjekt sommeren 2017. En gruppe studenter jobbet denne sommeren med å utvikle fundamentet til det som i dag er chatboten Frida. Selve utviklingen begynte i januar 2018, og fra februar til mars samme år gjennomførte NAV brukertesting via NAV LAB.<sup>1</sup> I april 2018 ble den første versjonen av Frida lansert på NAV sine nettsider. Utviklingen av Frida kan sees i sammenheng med Foreldrepenge-prosjektet. Dette er et prosjekt som startet i 2016 og som dreide seg om utvikling og innføring av nye IT-løsninger og arbeidsprosesser for foreldrepenger, engangsstønad og svangerskapspenger. (Mugås, 2019) Chatboten til NAV er bygget på chatbot-plattform som er utviklet av selskapet Boost AI. NetNordic leverer en komplett pakke med en kommunikasjonsløsning fra Puzzel og chatbotplattform fra Boost, til NAV.

### 3.1 Hvordan Frida fungerer

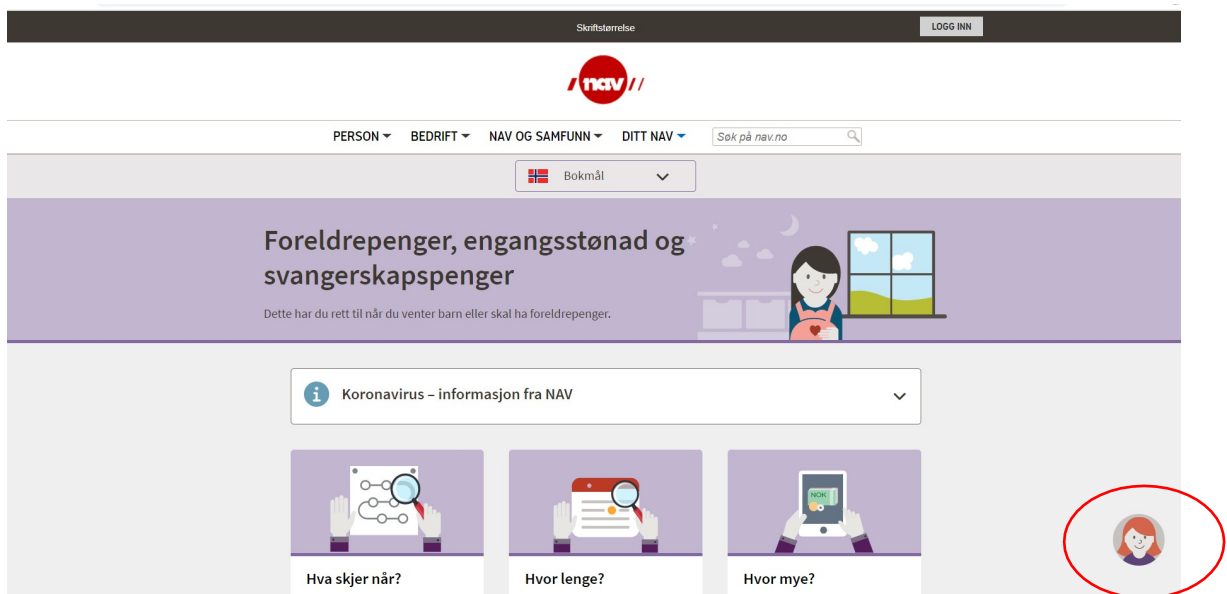
Frida er et chat-tilbud som befinner seg på seg på nettsiden til NAV. For øyeblikket svarer hun på spørsmål om støtteordninger for familie og barn, arbeidsavklaringspenger, dagpenger,

---

<sup>1</sup> <https://navlab.no/tester/bli-med-a-laere-opp-frida-navs-nye-digitale-ekspert-pa-foreldrepenger>

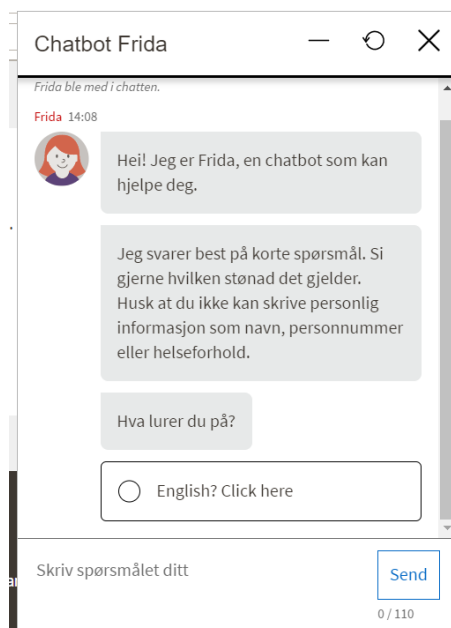
## KAPITTEL 3

sykepenger og sosiale tjenester, men datasettet til Frida blir stadig utvidet slik at hun kan svare på spørsmål om flere av NAV sine tjenester og støtteordninger. Da jeg startet å gjennomgå chatt-logger våren 2019 var det kun spørsmål knyttet til foreldrepenger, barnetrygd og foreldrepermisjon hun svarte på. Derfor vil også mange av eksemplene i denne oppgaven ifra mai, også dreie seg om disse temaene. På dette tidspunktet måtte man først gå inn på det aktuelle temaet man lurte på, på NAV sin nettside, før man senere fikk opp et valgalternativ om chat. Hvis man da befant deg på et område som Frida kunne svare på ville man bli koblet direkte opp til Frida. Nå som Frida er utvidet, får man også raskere opp Frida som alternativ. Hun vil da dukke opp som et ikon på høyre side av skjermen.



Når du trykker på ikonet får du opp en tekstboble på høyre side av skjermen. Den ser slik ut:

## CHATBOTEN FRIDA



Hun forklarer her at du ikke skal oppgi personlig informasjon og du får også valgalternativene 'Engelsk' som vil sette deg direkte over til chat med veileder, hvis du er innenfor tidsrommet hvor veiledere er tilgjengelig.

### Bruk av både knapper og tekstsvaer

Frida sine svar inneholder ofte knapper brukeren kan trykke på. Disse knappene kan være oppfølgingsspørsmål eller klargjøringer/spesifiseringer. Brukeren kan velge mellom å trykke på knappene eller å skrive inn spørsmålene sine selv. Om brukeren velger å trykke på knappene øker sjansen for at vedkommende vil bli koblet opp mot riktig intensjon. Skatteetaten sin rapport viser at bruk av knapper øker brukervennligheten ved at brukerne føler seg fornøyde og trygge (Arlov, 2018).

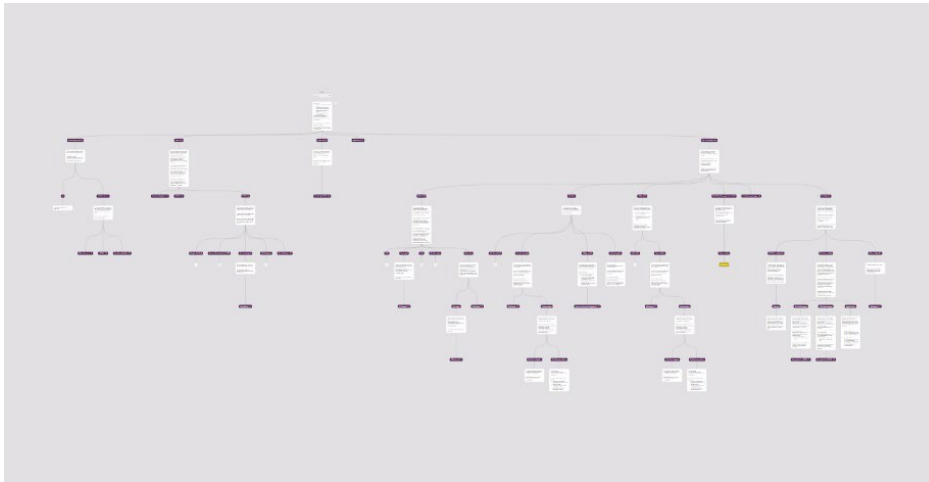
### Intensjoner

Frida vil koble det brukeren sier opp mot intensjoner fra intensjonstrær. Intensjoner er forskjellige emner/temaer som brukeren kan stille spørsmål om. Hver intensjon har sin egen treningsdata, som representerer de forskjellige måtene en bruker kan stille spørsmål om det gitte temaet. Slik som foreldrepenger eller permisjon. Alle intensjonene er organisert i et hierarki hvor intensjonsgrener er koblet sammen. (Boost AI, 2019a)

#### 3.1.1 Intensjonstrær

## KAPITTEL 3

Hver rotintensjon har forskjellige grener som er bygget opp i et hierarki.

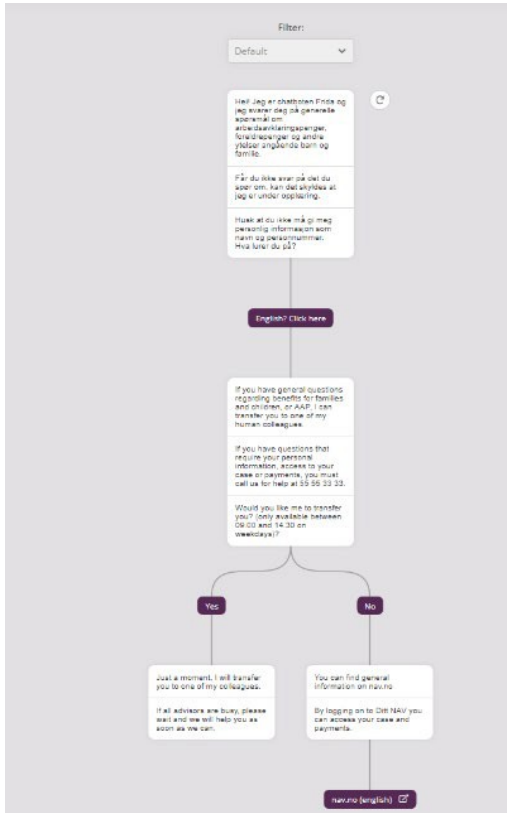


Figur 8: Intensjonstre

De hvite boksene er tekstfelt som Frida svarer ut ifra hvilken intensjonsgrens som svarer best med brukerens utsagn. De lilla feltene representerer knapper/valgalternativer som automatisk vil lede videre til neste gren om brukeren trykker på dem.

Slik ser intensjonstreet til Frida sin åpningsfrase ut:





Figur 9: Intensjonstre av Frida sin åpningsfrase

Her kommer Frida først med generell informasjon. Brukeren kan deretter velge å trykke på knappen for 'English' eller skrive spørsmålet sitt. Om brukeren trykker på knappen vil Frida gå videre til neste gren i treet som er å fortelle om muligheten til å bli satt over til veileder, som igjen gir brukere mulighet til å trykke på 'Yes' eller 'No'. Når Frida sine svar inneholder knapper, kan bruker også velge å skrive inn svarene for hånd i stedet for å trykke på knappene og allikevel fortsette samtalen til neste gren. Men om brukeren her skriver lange setninger, kan Frida ende opp med å hoppe over til en annen intensjon eller ende opp med å ikke klare å komme med et svar. Jeg vil gå nærmere inn på hvordan dette fungerer i neste del.

### 3.1.2 Hvordan Frida analyserer en setning

Frida kobler brukerens nøkkelord opp mot intensjon ut ifra sannsynlighetsberegning basert på tidligere treningsdata. Dette foregår som regel i tre steg, hvor analysen vil stoppe ved siste gyldige steg, men det finnes også unntak til dette.

Det er fire grunnleggende prinsipper/regler som bestemmer om sannsynligheten for den aktuelle intensjonen er gyldig.

## KAPITTEL 3

- Et steg vil være ugyldig dersom den foreslåtte/bregnede intensjonen hører til en annen intensjonsgren enn den forrige beregnede intensjonen. Dette inkluderer tilfeller hvor et steg beregnes til et intensjonstre fra den forrige beregnede intensjonsgrenen.
- Hvert steg i analysen må være på minst 45% for å være gyldig. Det vil si at dataen Frida sin analyse baserer seg på tilsier at det er minst 45% sannsynlighet for at den aktuelle intensjonen stemmer med brukerens utsagn.
- De to intensjonene som har størst sannsynlighet for å være riktig i hvert steg må ha en differanse på minst 15% for at steget skal være gyldig.
- Hvis brukerutsagnet passer perfekt med treningsdataen, vil alle steg automatisk bli 100%.

Analysen for et brukerutsagn vil da se slik ut:

Translated w/synonyms are there date invoic payment		Processed are there date invoic payment			
Step 1		Step 2		Step 3	
1. Payment & billing	83.65%	1. Invoice	59.07%	1. Help with invoice	36.50%
2. General Questions	14.54%	2. Payment	14.06%	2. Pay bill	12.74%
3. Boost.ai	1.71%	3. Payment & billing	9.54%	3. Payment & billing	9.54%

Figur 10: Illustrasjon av Frida sin analyse

Her kan man se at intensjonen 'Payment & billing' beregnes til å være 83,65%. Den er da over 45% og har minst 15% større sannsynlighet enn General Questions som står på andreplass i steg 1.

Analysen går da videre til steg 2. Her er intensjonen 'Invoice' beregnet til å være på 59.07% sannsynlighet og Payment beregnet til 14.06%.

Analysen vil derfor gå videre til steg 3. Her er intensjonen 'Help with invoice' beregnet til 36.5%. Dette er under 45% og derfor ugyldig.

Frida vil derfor stoppe ved steg 2 og gjengi svare for intensjonen Invoice. Nå befinner Frida seg på intensjonsgrenen 'Invoice' tilknyttet intensjonstreet 'Payment & billing'. Brukerens neste utsagn vil derfor også bli forsøkt analysert ifra dette intensjonstreet. (Boost AI, 2019b)

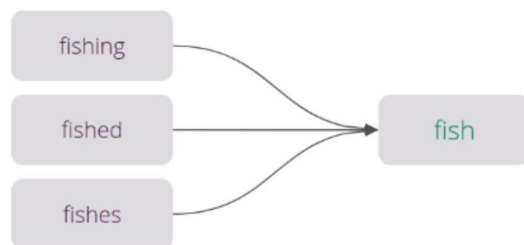
## CHATBOTEN FRIDA

Det er også andre faktorer som påvirker Frida sine analyse(Boost AI, 2019c):

- Hvis kun første steg er gyldig, vil dette bare gjengis om brukerutsagnet inneholder et nøkkelord fra intensjonstreet.
- Det hender at steg 3 kan bli gyldig, selv om steg 2 ikke er det, men dette er sjeldent. Om to steg på rad er ugyldige krever det en forskjell på minst 25% for at neste steg skal være gyldig.
- Hvis sannsynligheten for intensjonen har gått ned med mer enn 50% fra forrige steg, vil steget bli automatisk ugyldig, selv om det er over 45%.
- Hvis utsagnet inneholder for mange unknown kan intensjonen bli ugyldig selv om den opprinnelig var over 45%. (Boost AI; 2019g)

### Stemming

Kun stammen av ordet blir brukt. De forskjellige formene av ordet vil derfor omgjøres slik:



Figur 11: Illustrasjon hentet fra Thakur(2018)

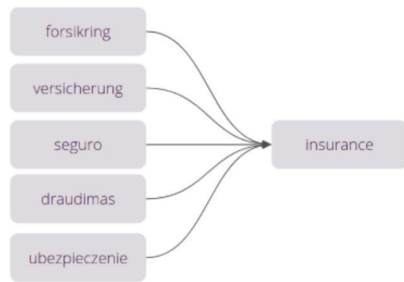
Her ser vi at de engelske variantene av fiske, fisket og fisker blir forkortet til ordet 'fisk'.

Alle utsagn som nevner en av disse variantene vil dermed bli koblet til samme intensjon.

### Synonymer

Chatbot-trenerne vil også stadig oppdatere synonym-listen, med hvilke ord anses som synonymer og dermed kobles opp mot samme intensjon.

## KAPITTEL 3

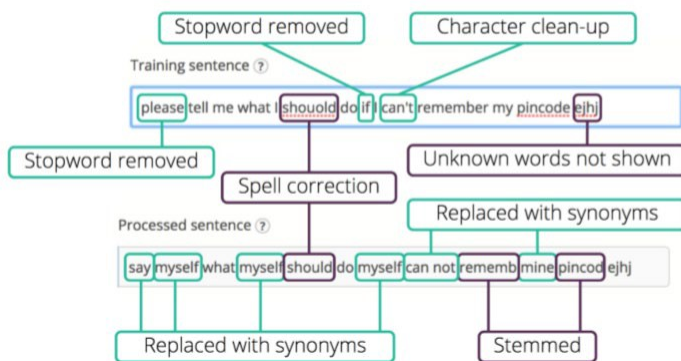


Figur 12: Illustrasjon hentet fra Thakur(2018)

Illustrasjon av hvordan ordene for forsikring på forskjellige språk kobles opp mot samme ord. Det fungerer på samme måte med synonymer innenfor samme språk.

### Stoppord

Stoppord er ord som ikke anses som meningsbærende og blir ignorert av Frida. Disse ordene blir derfor fjernet når setningen analyseres.

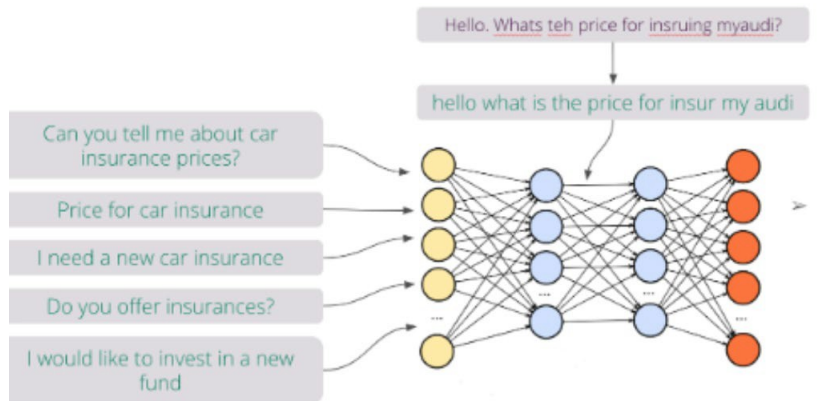


Figur 13: Illustrasjon hentet fra Boost AI(2019f)

Over er en illustrasjon av hvordan et brukerutsagn behandles. Stoppord som 'please' og 'if' fjernes. 'Should' rettes til 'should' av stavekontroll. 'I' og 'me' byttes ut med synonymet 'myself'. Og 'remember' og 'pincode' endres til kun stammene av ordene, henholdsvis 'rememb' og 'pincod'. Den prosesserte setningen blir deretter forsøkt koblet opp mot intensjoner.

Maskinlæringsalgoritmer blir så brukt til å bygge nevralt nettverk:

## CHATBOTEN FRIDA



Figur 14: Illustrasjon hentet fra Thakur(2018)

Ved å bygge opp samtalen på denne måten kan Frida lettere treffe presist med svarene sine, hvis hun får avklart hvilke områder innenfor et tema det er brukeren lurer på. Men om Frida ikke tar utgangspunkt i riktig nøkkelord kan dette også føre til misforståelser.

Mange brukere vet ikke hvordan de skal kommunisere med en chatbot, og skriver derfor lange fraser slik som de ville gjort til et menneske. Det vil da ofte være flere nøkkelord i en setning. Og om Frida da først er inne på feil tre, så kan det være vanskelig å få svar på det man egentlig lurer på.

### 3.2 Trening

Chatbot-trenerne jobber stadig med å forbedre Frida. Dette gjøres ved å skrive treningsdata og testdata og gjennomføre treningsrunder. Grunnlaget for endringene som gjøres baseres på rapporter, gjennomgang av samtaler og tilbakemeldinger fra veiledere.

Frida genererer flere rapporter, blant disse er Unknown og Sentiment. Unknown er samtaler der chatboten har fått spørsmål den ikke har funnet et svar på, eller der det har oppstått en konflikt med chatbotens øvrige innhold. Sentiment er samtaler der bruker har gitt uttrykk av å være fornøyd eller misfornøyd. Chatbot-trenerne fokuserer spesielt på samtaler hvor brukeren har uttrykt at den er misfornøyd. Ved å se på disse rapportene kan de kartlegge hvordan chatboten jobber og se på hvilke tiltak eller endringer som eventuelt må gjøres.

I tillegg til disse rapportene, går chatbot-trenerne også gjennom samtaler. Utfordringen er at

## KAPITTEL 3

en chatbot kan aldri vite om den har svart riktig. Selv om Frida har koblet et brukerutsagn opp mot intensjon og gitt brukeren et svar, så kan man ikke vite om det faktisk er riktig før man har lest samtalen. Derfor vil det alltid være nødvendig å gå gjennom en viss andel av samtaleene for å kartlegge hvor treffsikker chatboten egentlig er. Som regel går de gjennom rundt 2% av samtaleene, men det kan endre seg om volumet blir større. Chatbot-trenere får også tilbakemeldinger fra veiledere på ting som burde fikses.

For at chatboten skal lære seg den riktige måten disse spørsmålene skal besvares, skrives det treningsdata. Treningsdata er eksempler på forskjellige måter en bruker kan spørre om et tema. Her tar man med alle de forskjellige formene for spørsmål som brukerne kan stille og som Frida skal kunne svare på. Frida bruker treningsdata for å forstå hva som skiller en intensjon fra en annen intensjon. Vanligvis skriver man minst 20 testsetninger for hver intensjon. (Boost AI, 2019d) Treningsdata skal være setninger som er presise og med godt språk.

Testdata lages så til hver intensjon for å evaluere hvor bra Frida kan beregne hva brukerne spør om. Vanligvis vil testdata bestå av 10-15 setninger for hver intensjon. Testdata må være varierte måter å spørre om den samme intensjonen på. Her kan man gjerne inkludere skrivefeil for å se om dette fortsatt kobles opp mot riktig intensjon. Det er viktig at testdata ikke kopieres fra treningsdata. Helst skal ikke samme chatbottrener skrive både testdata og treningsdata. (Boost AI, 2019e) I Bergen er de to stykker som jobber med dette. Da kan en skrive treningsdataen og den andre testdataen.

Deretter gjennomføres en treningsrunde hvor chatbot-trenerne ser på hvor mange prosent av testdataene som kom frem til riktig intensjon. Her kan de se effekten en endringene som har blitt gjort, om de for eksempel har lagt til et nytt synonym eller gjort om på en svarstruktur. De kan også se om det er konflikter i forhold til andre intensjoner. Chatbottrenerene må også lage synonymlister, ettersom at NAV sin chatbot operer i et særegent domene.

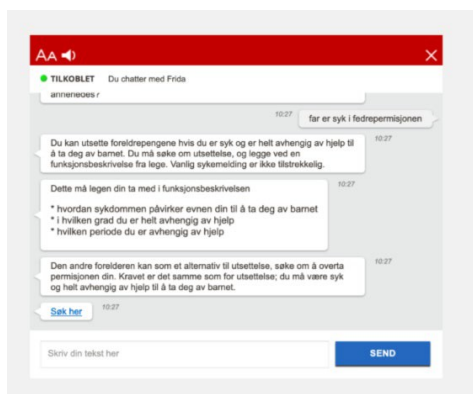
### **3.3 Utviklingen av Frida**

Frida er en chatbot som stadig er i utvikling. Dette er en utfordring når man skal studere de problemene som kan oppstå når bruker samhandler med Frida, ettersom at dette endrer seg fra dag til dag. Heldigvis fikk vi fått tilgang til loggene slik at vi kunne gå tilbake i tid å se

## CHATBOTEN FRIDA

endringene som har blitt gjort, men dette førte også til utfordringer ved at det var vanskelig å peke på konkrete problemer i chatten ettersom det hele tiden ble gjort endringer.

Når man startet en chat med Frida ble chatten åpnet i et eget vindu i nettleseren. Men chattevindu var lite og det var vanskelig å lese all teksten siden man måtte scrolle oppover for å få med seg alt Frida skrev. Et svar kunne gjerne være på 2-3 avsnitt og man havnet automatisk ved bunnen av svaret. Mange fikk derfor ikke med seg at det var mer tekst og at de måtte scrolle oppover for å lese den. I chatloggene jeg leste ifra mai, kunne jeg se at mange brukere uttrykte irritasjon over at de ikke rakk å lese alt Frida skrev og at svarene hennes virket ufullstendige. Mange brukere stilte spørsmål som Frida allerede hadde besvart tidligere i chatten.



Figur 15: Skjerm bilde av hvordan Frida pleide å se ut (Simonsen, 2019)

Denne formateringen gjorde det vanskelig å bruke Frida. Det var liten tekst, mye informasjon på en gang og mange brukere slet med å få med seg alt Frida sa.

### Frida sin åpningsfrase

En annen stor forskjell her, er at i åpningsfrasen til Frida så sier hun at hun er under opplæring. Dette har to mulige konsekvenser. Den ene er at brukerne blir mer tålmodig med Frida, omformulerer spørsmålet sitt og spør på nytt og viser større forståelse for feil. Det andre er at brukeren blir mer skeptiske til svar fra Frida, og i større grad ber om å bli satt over til veileder fordi de ikke stoler på at Frida svarer riktig ettersom at hun er under opplæring.

Dette er variantene av Frida sin åpningsfrase som har vært brukt i denne perioden:

## KAPITTEL 3

Mai 2019:

<p>Jeg heter Frida og er en chatbot.</p>

<p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p>

<p>Hva lurer du på?</p>

For english chat, click here

I mai informerte kun Frida om at hun svarte best på korte spørsmål.

Oktober 2019:

<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva lurer du på?</p>

<p>Får du ikke svar på det du spør om, **kan det skyldes at jeg er under opplæring.**</p>

<p>**Husk at du ikke må gi meg personlig informasjon som navn og personnummer.**</p>

English? Click here

I oktober opplyste hun om at hun var under opplæring og at man ikke må oppgi personlig informasjon om navn og personnummer.

Desember 2019:

<p>Hei! Jeg er Frida, en chatbot som kan hjelpe deg.</p>

<p>Jeg svarer best på korte spørsmål. Si gjerne hvilken stønad det gjelder. **Husk at du ikke kan skrive personlig informasjon som navn, personnummer eller helseforhold.**</p>

<p>Hva lurer du på?</p>

English? Click here

I desember var også helseforhold lagt til i listen over det brukeren ikke skal oppgi. I tillegg var setningen om opplæring fjernet.

Måten brukerne formulerer seg på i chatten kan dermed ha en sammenheng med den åpningsfrasen de ble møtt av.

### **Bruk av vitser og trivia**

Da jeg først begynte å undersøke Frida hadde hun intensjoner som gjorde at hun kunne fortelle vitser og morsomme fakta som var irrelevant for domene. Slike funksjoner kan bidra



## CHATBOTEN FRIDA

til menneskeliggjøring av Frida og påvirke forventingene brukerne har. I tillegg kan det ha en innvirkning på hvor profesjonell chatboten fremstår. Det kan også ha konsekvenser om disse intensjonene blir koblet opp mot andre spørsmål når brukerne egentlig stiller seriøse spørsmål. En chatbot-trener vi snakket med uttrykte skepsis til denne type intensjoner, og dette ser nå ut til å være deaktivert(mai 2020).

## 4 Metodologi

Dette kapittelet vil presentere metodene som har blitt brukt for å samle inn og analysere data, samt etiske aspekter ved studien. Studien er gjennomført ved bruk av kvalitative forskningsmetoder for å undersøke bruk av chatbot i NAV. Samtlige intervjuer og observasjoner har blitt gjennomført i samarbeid med Linett Simonsen.

### Metodologi og paradigme

Denne oppgaven tar for seg en instrumentell case studie av bruk av chatbot i NAV. Utgangspunktet vil være én enkelt case, men jeg ønsker med denne studien å også få en bredere forståelse for hvordan kommunikasjonen mellom bruker og chatbot fungerer.

Jeg anser denne studien som interpretivistisk. Interpretivistiske studier forsøker å forstå fenomener gjennom meninger som mennesker tildeler det (Myers, 1997). I denne studien har jeg forsøkt å forstå hva som kan oppleves utfordrende for brukerne ved å se på det som står skrevet i chatloggene.

### 4.1 Datainnsamlingsmetoder

#### Chat-logger

Hoveddelen av datagrunnlaget for denne masteroppgaven er et utvalg av chatbotlogger. Flere av intervjuene er gjort på bakgrunn av disse loggene. Loggene som er brukt til denne studien er hentet ifra perioden mai 2019 – desember 2019. Jeg brukte da logger ifra to uker i mai, to uker i september, to dager i oktober og to dager i desember. Til sammen utgjorde dette rundt ti tusen chatlogger. I starten brukte jeg mye tid på å lese gjennom logger for å se hva slags problemer som oppstod og hva som så ut til å gjenta seg. Etter hvert gjorde jeg søk på intensjoner og formuleringer for å se etter eksempler med lignende utfordringer. Jeg har derfor ikke lest gjennom alle disse loggene, men de har vært utgangspunkt for søkene jeg har gjort.

Det ble gjort mye endringer av Frida underveis i denne perioden og i oktober ble også et nytt design lansert. Chatloggene som blir presentert i denne oppgaven vil derfor bære preg av

## METODOLOGI

perioden de er hentet fra, både på grunn av endringene som er gjort og på grunn av spørsmålene som varierer etter tiden på året.

Jeg har også fått tilgang til forskjellige dokumenter i forbindelse med masteroppgaven min, blant annet en rapport fra Skatteetaten og presentasjoner ifra NAV.

### **Observasjon av veiledere**

I den innledende fasen av datainnsamlingen observerte vi en veileder på jobb på NAV Kontaktsenter i Oslo. Observasjonen varte i 30 min. Veilederen forklarte underveis hvordan systemet de brukte fungerte, og hvordan arbeidsoppgavene vanligvis ville bli gjort. Vi hadde også anledning til å stille spørsmål underveis. Dette var en nyttig måte å få innblikk i hvordan veilederne jobbet og vi fikk se hvordan det så ut når chattene ble satt over fra chatbot til mennesker.

I oktober hadde vi en ny observasjon av veileder, denne gangen en veileder på NAV Kontaktsenter Vestland i Bergen. Denne observasjonen varte i cirka en og en halv time med en pause underveis. Under denne observasjonen fokuserte vi mer på hvordan veilederen besvarte spørsmålene chatboten ikke hadde klart å svare på, og hvordan veilederen tilpasset spørsmålene ut ifra brukeren. Vanligvis ville en veileder også svare på telefoner, men i anledning observasjonene vi gjorde, satt veilederne kun på chat.

### **Intervju med veiledere**

Vi gjennomførte to 30 minutters semi-strukturerte intervjuer med to veiledere i NAV. Intervjuobjektene våre ble plukket ut av NAV. Disse jobbet til vanlig med kontakt med brukere på både chat og telefon, og hadde god kjennskap til hvordan kontakt med brukere hadde fungert både før og etter innføringen av Frida. Vi startet intervjuene med utdeling av samtykkeskjema hvor intervjuobjektene fikk lese gjennom skjema i ro før de skrev under. De fikk også en kopi av skjema som de kunne beholde. Underveis i intervjuet tok vi notater for hånd, i tillegg til lydopptak. Etterpå hørte vi gjennom lydopptakene og transkriberte intervjuene.

Etter å ha analysert chatloggene tok vi igjen et intervju med en veileder og diskuterte konkrete utfordringer vi hadde funnet i chattene. Vi tok med oss chatlogg-eksempler og forhørte oss

## KAPITTEL 4

med veilederen om hvordan en veileder ville svart på de spørsmålene brukerne hadde stilt Frida.

### **Intervju med chatbot-trenere**

Vi intervjuet to chatbot-trenere som jobber i NAV Kontaktsenter Vestland i Bergen. Vi intervjuet dem først hver for seg. Det første intervjuet varte en time og 40 minutter. Det andre varte i en og en halv time. Vi tok lydopptak av intervjuene samtidig som vi tok notater for hånd. Intervjuene ble gjort på et møterom hvor det også var mulig for intervjuobjektene å koble en skjerm opp til pc-ene sine for å forklare ting når de svarte på spørsmålene våre. Intervjuene var semistrukturerte ved at vi hadde laget en intervjuguide på forhånd, men også stilte oppfølgingsspørsmål underveis.

Etterpå hadde vi et intervju med begge chatbot-trenerne. Dette ble gjort for å gå gjennom det vi hadde snakket om tidligere for å forsikre oss om at det ikke var noen misforståelser, og for å klargjøre noen punkter fra de to første intervjuene. Dette intervjuet varte i cirka to timer. Under dette intervjuet viste også en av intervjuobjektene oss lysbilder ifra en presentasjon han hadde holdt om brukerkontakt.

### **Observasjon av chatbot-trenere på jobb**

Etter de to første intervjuene med chatbot-trenerne så hadde vi en observasjon av en chatbot-trener i en vanlig arbeidssituasjon. Vi fikk her se hvordan chatbot-treneren laget treningssetninger og testdata, og hvordan en treningsrunde ble satt i gang. Under observasjonen satt vi sånn at vi kunne følge med på pc-skjermen mens vedkommende jobbet, samtidig som vi tok notater for hånd. Dette gjorde at vi fikk større innblikk i hvordan treningen foregikk og hvordan chatbot-trenere vanligvis jobbet.

### **Intervju med digital arkitekt i NetNordic**

Vi gjorde et intervju med en ansatt fra NetNordic. Til dette intervjuet hadde vi også forberedt en intervjuguide, men vi gikk i større grad bort ifra intervjuguiden i dette intervjuet ettersom at intervjuobjektet også hadde forberedt en del punkter selv som han ville fortelle oss om. Dette intervjuet ga oss et godt innblikk i hvordan NetNordic, Boost og NAV Kontaktsenter samarbeider.

### **Besøk hos Skatteetaten**

Vi var på besøk hos Skatteetaten hvor de holdt en presentasjon for oss om utviklingen og bruken av deres chatbot. Etterpå fikk vi anledning til å stille spørsmål. Skatteetaten sin chatbot er utviklet av Boost, som også står bak NAV sin chatbot. Dette var derfor en nyttig erfaring for å se hvordan det jobbes med utvikling av chatbot i andre etater og hva Skatteetaten har erfart ved bruk av chatbot.

### **Rekruttering**

Deltagerne til studien har blitt rekruttert ved hjelp av Snøballutvelgelse. NAV valgte ut intervjuobjektene vi skulle forholde oss til, og de vi intervjuet anbefalte andre vi også burde snakke med. Det positive med dette, er at de vi snakket med, var blant de som hadde jobbet der lengst og hadde mest kunnskap om hvordan arbeidsoppgavene fungerte før og etter innføring av Frida.

### **Begrensninger ved studien**

I denne studien fikk vi ikke anledning til å intervju brukerne som benytter seg av NAV sin chatbot. Det er derfor ikke sikkert at våre antagelser om brukerne stemmer. I analysen av chattene har vi kun hatt det brukeren skriver som utgangspunkt for våre antagelser om hvorvidt Frida sine svar er tilstrekkelige for å besvare det brukeren lurte på. Tilfeller der hvor brukerne egentlig lurte på noe mer eller følte at Frida sitt svar ikke passet med sin situasjon, er derfor vanskelig å avdekke, fordi vi ikke kan vite sikkert hva brukerens formål med interaksjonen var.

Under datainnsamling til denne studien, førte også innføringen av GDPR til utfordringer. Chatloggene som ligger til grunn for analysen er hentet ifra perioden mai-desember i 2019. Etter denne perioden gjorde oppdateringer av personvernreglementet hos NAV det vanskelig å få tilgang til loggene. Det gjorde det også vanskelig å få gjort noen flere intervjuer og observasjoner hos NAV. Dette i kombinasjon med Korona-situasjonen, som påvirket store deler av vårsemesteret hvor jeg jobbet med ferdigstillingen av denne masteroppgaven, gjorde at videre datainnsamling ikke ble mulig.

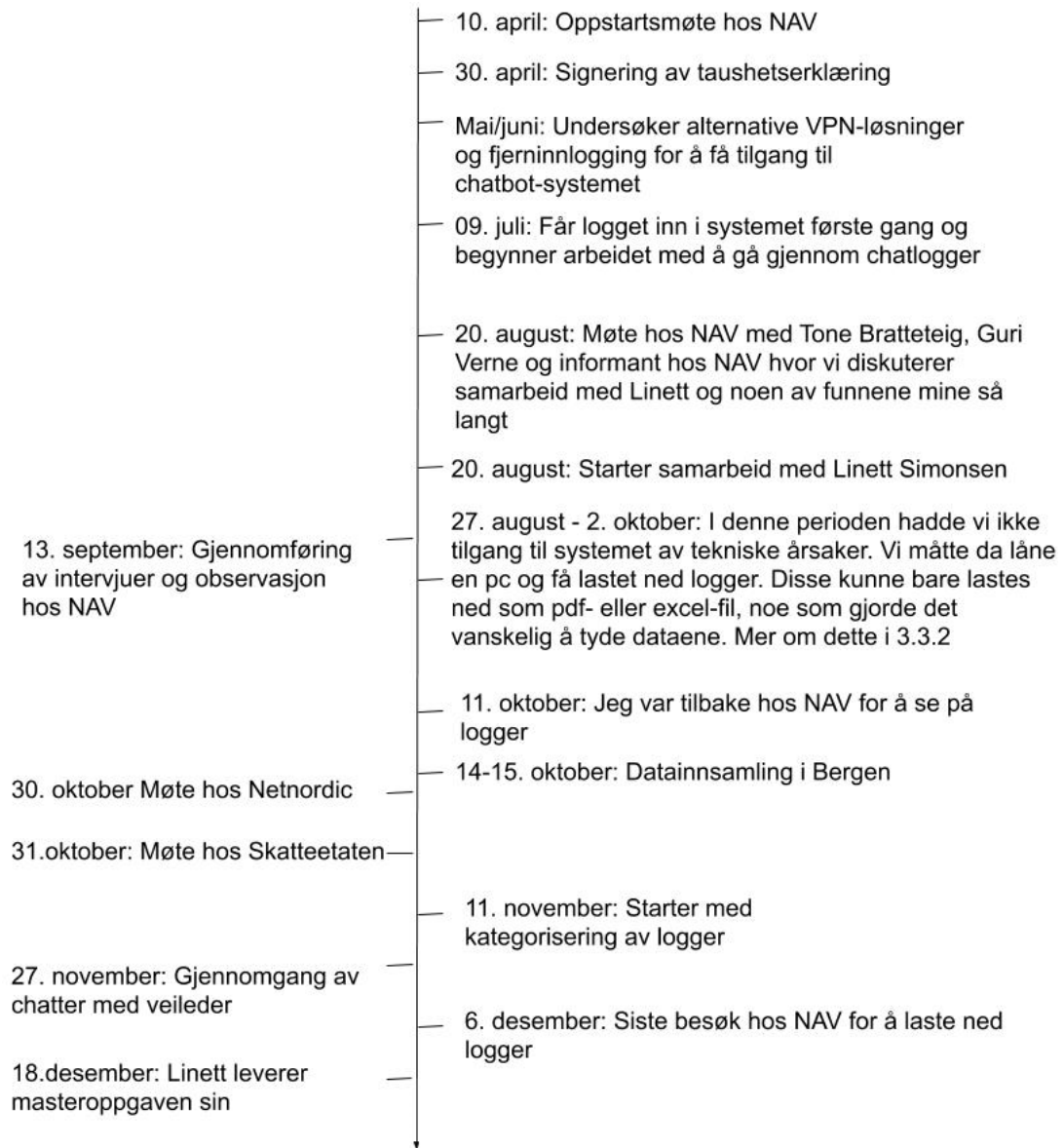
## 4.2 Oversikt over datainnsamlingsaktiviteter

Ukenummer	Aktivitet	Lokasjon	Informant
Uke 15	Oppstartsmøte	NKS Oslo og Vest-Viken	Kontaktperson i NAV
Uke 34	Planleggingsmøte	NKS Oslo og Vest-Viken	Kontaktperson i NAV
Uke 37	Intervju og observasjon	NKS Oslo og Vest-Viken	Veileder 1
Uke 37	Intervju	NKS Oslo og Vest-Viken	Veileder 2
Uke 41	Intervju og demonstrasjon	NKS Vestland	Chatbottrener 1
Uke 41	Intervju og observasjon	NKS Vestland	Chatbottrener 2
Uke 41	Intervju og observasjon	NKS Vestland	Veileder 3
Uke 44	Intervju	NetNordic, Lysaker	Digital arkitekt i NetNordic
Uke 44	Demonstrasjon	Skatteetaten, Helsfyr	Seniorrådgiver i Skatteetaten
Uke 48	Gjennomgang av kategorier og logger med domeneekspert	NKS Oslo og Vest-Viken	Veileder 1

Tabell 1: Datainnsamlingsaktiviteter

### 4.2.1 Tidslinje over arbeid med masteroppgaven i 2019

Jeg førte ikke logg underveis, men jeg har forsøkt å lage en tidslinje for hvordan jeg arbeidet med denne oppgaven.



Figur 16: Tidslinje

I perioden 9. juli til 20. august gikk jeg gjennom loggene ifra 1.mai til og med 15.mai. Dette utgjorde totalt 5321 chatter. Ut fra denne gjennomgangen identifiserte jeg noen problemer som gikk igjen. Dette var blant annet problemer med ordet 'ikke', brukere som ikke så at de

## KAPITTEL 4

snakket med en chatbot, og lange responser fra Frida som brukerne tilsynelatende ikke leste gjennom. Jeg hadde deretter et møte med NAV 20. august. NAV sin informant kunne da fortelle at problemene med at brukere ikke så at Frida var en chatbot og ikke leste gjennom svarene, var noe de også hadde merket selv og jobbet med å forbedre, ved å endre åpningsfrasen til Frida slik at det skulle bli tydeligere at hun var en chatbot, og å dele opp svarene til Frida i flere responser sånn at det skulle være lettere å lese. Med det nye designet i oktober ble det også lettere å lese svarene til Frida.

20. august begynte jeg samarbeidet med Linett Simonsen. Før hun kunne få tilgang til loggene måtte hun først signere taushetserklæringen, og jeg var derfor hos NAV alene i august for å se nærmere på noen av de utfordringene jeg hadde oppdaget så langt. I denne runden med datainnsamling søkte jeg spesifikt etter de tre utfordringene jeg hadde oppdaget og tok skjermbilder av noen av setningsanalysene i det innloggede systemet, siden det var lettere å tyde analysene i systemet enn i Excel.

Den 27. august mistet vi tilgangen til systemet fra våre kontoer på grunn av tekniske problemer hos NAV. Siden Linett ikke hadde fått lastet ned noen logger enda, så sendte jeg henne de loggene jeg hadde fått lastet ned fra mai og jeg fortalte henne om de utfordringene jeg hadde sett så langt. Da det ble klart at det ville ta noen uker å få fikset problemet, fikk vi låne en annen pc for å få lastet ned flere chatter. Linett og jeg satt da sammen ved en pc og leste gjennom logger. Vi så også på intensjonstrærne sammen og satte oss inn i hvordan dette fungerte.

14. og 15.oktober var vi i Bergen. Der så vi på noen av chattene som kom inn i disse dagene sammen med chatbottrenerne. For å senere finne igjen eksemplene vi hadde snakket om, lastet jeg ned og gikk gjennom alle chattene fra disse to dagene. Både Linett og jeg syntes det var viktig å se på chatter fra flere tidsperioder enn bare mai, ettersom at det ble gjort endringer underveis, og at spørsmålene fra brukerne varierte etter tiden på året. Vi lastet derfor også ned en del chatter fra september. Da jeg så på chattene fra september startet jeg først med å lese gjennom en del chatter, deretter søkte jeg etter spesifikke ting slik som intensjonen 'er du en robot'. Jeg forsøkte også å søke på språk for å se hva som skjedde i tilfeller hvor Frida klassifiserte brukerutsagn som andre språk enn norsk og om dette førte til problemer, slik som man ser i 7.1.2. Dette begynte jeg å søke etter da vi hadde vært i Bergen, ettersom at en chatbottrener kommenterte at feilidentifisering av språk kan skje, da vi så på en annen chat.



I november jobbet vi med å analysere og kategorisere loggene. Vi hadde notert noen utfordringer underveis som vi tok utgangspunkt i. Ved å gruppere loggene kom vi etter hvert frem til kategorier og underkategorier av utfordringer. Dette vil utdypes i 3.3.1.

Høsten 2019 tok jeg også et emne som omhandlet bruk av chatboter, parallelt med arbeidet med masteroppgaven. Utfordringene med chatboter som ble diskutert i forelesninger og i pensumtekster har nok også påvirket hvilke ting jeg har tenkt over når jeg har lest gjennom loggene, og hvilke ting jeg har funnet det interessant å søke etter. Jeg prøvde å fokusere på hva som kommer frem i chatter hvor brukere ikke får svar på det de lurer på, og ta utgangspunkt i dette, men artikler jeg leste underveis, blant annet Hill (2015) og Amershi m.fl. (2019) har også påvirket hvilke ting jeg har lagt merke til.

Etter at Linett leverte sin masteroppgave i desember, fokuserte jeg på å gå gjennom chattene fra desember med bakgrunn i kategoriene våre. Jeg fant da ut at jeg ville gjøre noen endringer på kategoriene. Dette forklarer jeg nærmere i 12.2. Totalt i løpet av denne perioden ble det rundt 10 000 chatter jeg benyttet til datainnsamling og analyse. Disse var hentet fra mai, september, oktober og desember.

Jeg hadde også planlagt noen flere datasamlingsaktiviteter våren 2020, men disse ble ikke mulig å gjennomføre på grunn av en kombinasjon av oppdatering av personvernsreglementet hos NAV og Korona-situasjonen. Vårsemesteret 2020 ble derfor brukt til videre analysering av chatter og tilpasning av kategoriene, skriving av artikkel, samt skrivingen av masteroppgaven.

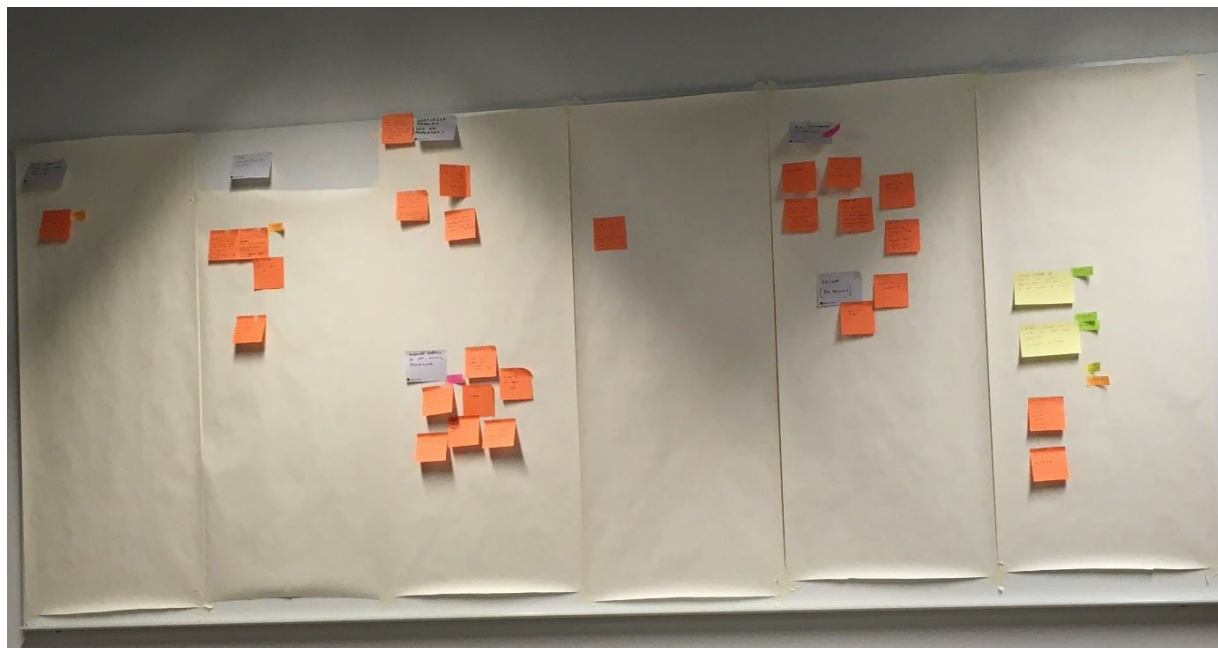
## **4.3 Analyse**

### **4.3.1 Koding chatbotlogger**

Sammen med Linett Simonsen har jeg analysert og kategorisert et utvalg av chatbot-loggene. Vi hadde en hermeneutisk tilnærming hvor vi jobbet med loggene gjennom flere runder med koding. Underveis så vi flere utfordringer som gikk igjen ved å først ha en induktiv tilnærming til loggene. Vi tok også utgangspunkt i noen av kategoriene Guri Verne brukte i sin doktorgradsavhandling til å klassifisere utfordringer i Skatteetaten. Disse samtalen foregikk over telefon. Dette betyr at Verne hadde mer informasjon om brukerne enn oss, og

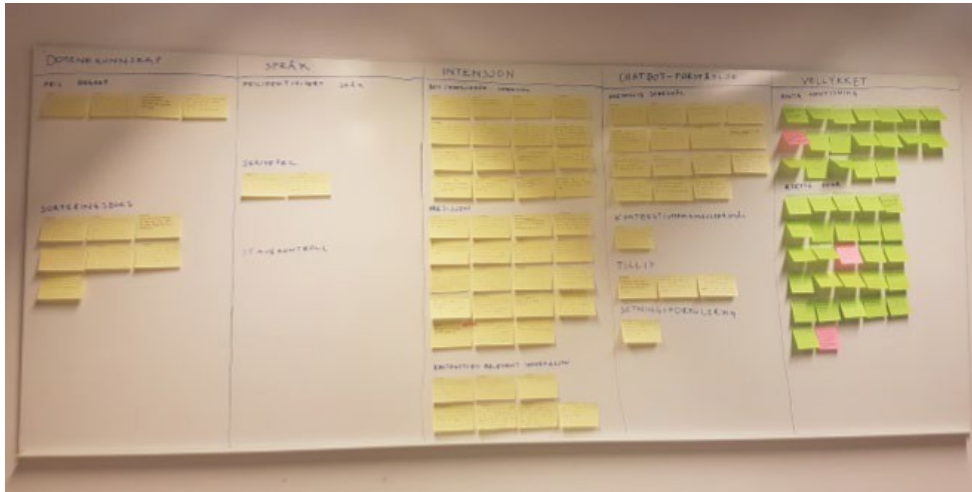
## KAPITTEL 4

det var også litt andre utfordringer som oppstod i disse telefonsamtalene enn hva vi har sett i chatloggene vi har analysert. Den kategorien som var mest relevant for oss herifra var 'Puttekassen'.(Verne, 2015)



Figur 17: Koding av første iterasjon

Vi så deretter gjennom et utvalg av logger fra én dag i mai, den 14.05, og sorterte disse ut ifra hvilke utfordringer hadde sett i chattene. Ved å først ha en induktiv tilnærming, ble det klart for oss at det var særlig fire kategorier av utfordringer som gikk igjen; Språk, Chatbotforståelse, Domenekunnskap og Intensjoner. Hver av disse kategoriene delte vi opp i underkategorier. I tillegg laget vi en kategori for 'Andre utfordringer' hvor vi samlet de chattene som ikke passet inne i noen av kategoriene vi hadde laget. Etter hvert kom vi også frem til noen underkategorier her.



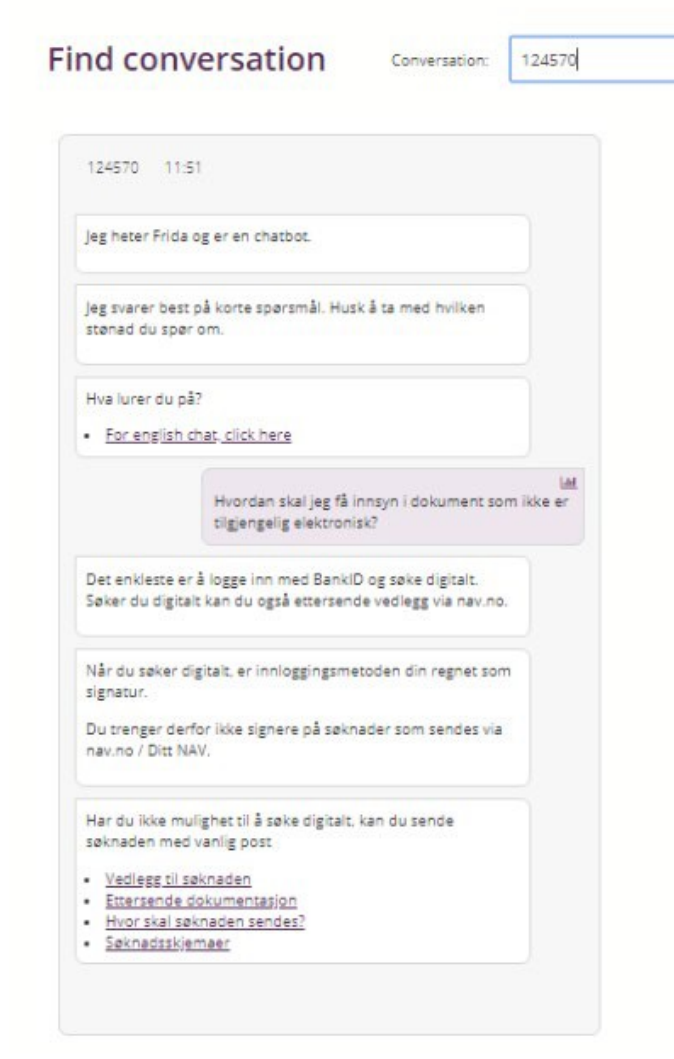
Figur 18: Koding av andre iterasjon

Vi hadde så en deduktiv tilnærming og prøvde å plassere et utvalg av chatter i de kategoriene vi kom frem til, for å se om vi kunne plassere alle chattene i en av kategoriene. Dette var for å se om utfordringene fortsatt var gjeldene for et større utvalg av chattene, men fokuset vårt var på hvilke problemer som gikk igjen, ikke at hver chat nødvendigvis skulle høre til en bestemt kategori. Noen chatter inneholdt også sammensatte problemer som kunne høre hjemme i flere kategorier, vi valgte da å kategorisere ut ifra hva som så ut til å være hovedutfordringen i den gitte samtalen.

### 4.3.2 Gjengivelse av chatlogger

I denne oppgaven vil jeg gjengi chatlogger hvor brukere har kommunisert med Frida. Eksemplene har blitt hentet ut fra adminpanelet til Boost og lastet ned som pdf og excel-fil. Deretter har jeg fjernet alt av personlige opplysninger som brukerne har oppgitt, og gjengitt chattene med både hva som ble skrevet og intensjonene som hører til. Både veilederne og chatbot-trenerne omtalte Frida som om hun hadde menneskelige egenskaper og kunne «forstå» tekst. Jeg har valgt og også omtale chatboten på denne måten i oppgaven min når jeg forklarer hva som skjer i loggene. Når jeg skriver at Frida «forstår», mener jeg altså at hun finner nøkkelord i brukerutsagn og kobler dette til riktig intensjoner.

Da jeg skulle analysere loggene måtte jeg bruke en NAV-pc og være på et NAV kontor for å få logget inn. I systemet ble chattene presentert på dette formatet:



Figur 19: Skjerm bilde av chat ifra mai 2019

For å kunne se på loggene utenom når vi var hos NAV, og for også å kunne fortsette analysen i perioden hvor vi ikke hadde tilgang til systemet fra våre kontoer, så lastet vi chattene ned som PDF- og Excelfiler.

# METODOLOGI

124874 (May 14, 2019)

Type	Text	Intent
Action	Jeg heter Frida og er en chatbot.	
	Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.	
	Hva lurer du på?	
	<ul style="list-style-type: none"> <li>For english chat, click here</li> </ul>	
Message	Hei! Jeg lurte på om far har rett på permisjon dersom jeg tar engangsstønad og har permisjon fra studiene?	Engangsstønad / ? Far vilkår
Action	Ja, i noen tilfeller kan far få engangsstønad.	
	Du kan lese mer om rettighetene til far ved å klikke på lenken under.	
	Kom gjerne tilbake til meg hvis det er mer du lurer på.	
	<ul style="list-style-type: none"> <li>Fars rettigheter</li> </ul>	

Figur 20: Skjerm bilde av en chat i en PDF-fil

PDF-filen viser tydelig hvordan brukerutsagn kobles opp mot ulike intensjoner. Som PDF er chattene lettere å lese enn i Excel, men det forsvinner også mye informasjon som man bare kan se i Excel-filen.

Message ID	Date	Time	User message	Action text	Conversa	Type ID	Type description	Prediction typ	Prediction typ	Pred
1	1118129	15.10.2019	23:26	<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva l	176290	2	Action sent to user	\N	\N	\N
2	1118130	15.10.2019	23:28	Hvordan beregnes arbeidsavklaringspenger i forhold til inntekt?	176290	1	Customer question	Regular/Valid		4
3	1118131	15.10.2019	23:28	<p>AAP blir fastsatt på grunnlag av din arbeidsinntekt i året før, eller de tre siste årene før arbeidsevnen din ble redus	176290	2	Action sent to user	\N	\N	\N
4	1118127	15.10.2019	23:23	<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva l	176289	2	Action sent to user	\N	\N	\N
5	1118128	15.10.2019	23:23	<p>If you have general questions regarding benefits for families and children, or AAP, I can transfer you to one of my	176289	2	Action sent to user	\N	\N	\N
6	1118113	15.10.2019	23:07	<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva l	176288	2	Action sent to user	\N	\N	\N
7	1118114	15.10.2019	23:07	Hei Frida	176288	1	Customer question	Perfect Match		7
8	1118115	15.10.2019	23:07	<p>Hei</p>. <p>Hva kan jeg hjelpe deg med?</p>.	176288	2	Action sent to user	\N	\N	\N
9	1118116	15.10.2019	23:07	Hei	176288	1	Customer question	Unknown		2
10	1118117	15.10.2019	23:07	<p>Jeg tror jeg har forsøkt å svare på dette, men det kan virke som at det ikke ble så vellykket som jeg ønsket. Jeg læ	176288	2	Action sent to user	\N	\N	\N
11	1118118	15.10.2019	23:08	<p>Du kan selvsagt få snakke med en av mine kolleger.</p>. <p>Hvordan ønsker du å snakke med oss?</p>. Chat, T	176288	2	Action sent to user	\N	\N	\N
12	1118119	15.10.2019	23:08	<p>Åpningstiden for chat er 9-14.30, mandag til fredag.</p>. <p>Skal jeg sette deg over?</p>. Ja, Nei takk	176288	2	Action sent to user	\N	\N	\N
13	1118120	15.10.2019	23:08	<p>Beklager, vi er stengt. Våre åpningstider er mandag til fredag, 09.00 - 14.30.</p>. <p>Sorry, we are closed. Our o	176288	2	Action sent to user	\N	\N	\N
14	1118121	15.10.2019	23:08	Hvorfor	176288	1	Customer question	Regular/Valid		4
15	1118122	15.10.2019	23:08	<p>Beklager, men jeg forstod ikke helt hva du spurte om.</p>. <p>Kan du prøve å spørre meg igjen på en litt annen n	176288	2	Action sent to user	\N	\N	\N
16	1118123	15.10.2019	23:09	Det er feil i oppsummering av foreldrepenge tid, hva skyldes det?	176288	1	Customer question	Regular/Valid		4
17	1118124	15.10.2019	23:09	<p>For å få foreldrepenge må du sende inn en digital søknad om foreldrepenge.</p>. <p>Mor og far eller medmor,	176288	2	Action sent to user	\N	\N	\N
18	1118125	15.10.2019	23:09	<p>Arbeidsituasjonen din og hva du søker om avgjør hvilke vedlegg du trenger til søknaden om foreldrepenge.</p>.	176288	2	Action sent to user	\N	\N	\N
19	1118126	15.10.2019	23:10	<p>Du kan ettersom vedlegg til søknad om foreldrepenge på Ditt NAV - Dine foreldrepenge, eller ved å følge linke	176288	2	Action sent to user	\N	\N	\N
20	1118103	15.10.2019	22:46	<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva l	176285	2	Action sent to user	\N	\N	\N
21	1118107	15.10.2019	22:55	arbeidsavklarings penger	176285	1	Customer question	Unknown		2
22	1118108	15.10.2019	22:55	<p>Beklager, men jeg forstod ikke helt hva du spurte om.</p>. <p>Kan du prøve å spørre meg igjen på en litt annen n	176285	2	Action sent to user	\N	\N	\N
23	1118109	15.10.2019	22:55	arbeidsavklaringspenger	176285	1	Customer question	Perfect Match		7
24	1118110	15.10.2019	22:55	<p>Arbeidsavklaringspenger (AAP) skal sikre deg inntekt i perioder du på grunn av sykdom eller skade har behov for hj	176285	2	Action sent to user	\N	\N	\N
25	1118102	15.10.2019	22:42	<p>Hei! Jeg er chatboten Frida. Jeg svarer best på korte spørsmål. Husk å si hvilken stønad du har spørsmål om. Hva l	176284	2	Action sent to user	\N	\N	\N
26	1118104	15.10.2019	22:48	Nav må kontakte meg forrest mulig på tlf.. <number>	176284	1	Customer question	Regular/Valid		4

Figur 16: Skjerm bilde av chatter i Excel-fil

Excel-filen viser flere detaljer om chattene, som hvilket klokkeslett hvert spørsmålene ble skrevet, hvilket språk chatboten har tolket brukerens utsagn som, og hvordan setningen blir prosessert av Frida.

## KAPITTEL 4

	Tid	Brukerutsagn	Frida sine responser	Samtale-ID														
1	Time	User message	Action tool	Conversation ID	Type ID	Type description	Prediction to	Predictor	Product	Predicted intent	Root intent	Root intent title	Translated sentence	Processed	Home Language			
2	23:28	«gmed? jeg er chatboten Frida. jeg ønsker best på korte spørsmål. Husk å si hvilken stemme du har sparsmbli om.		176290	2	Action sent to user	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	23:28	Hvordan beregnes arbeidsbidragsgjengen i forhold til lønnet?		176290	3	Customer question	Regular/Valid	4	2280	Beregning AAP	NA	2257 AAP	How calculat arbeidsbidragsgjengen	NA	NA	NA	NA	norwegian
4	23:28	«AAP blir fjernet på grunnlag av din arbeidsbidrag i året før, eller de tre siste årene før arbeidsbidraget din blir		176290	2	Action sent to user	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Figur 22: Excel-fil bestående av 18 kolonner med informasjon koblet til hvert utsagn

Ved å bruke av kombinasjon av Excel og PDF har vi kunnet hente ut mer informasjon av chattene og se hva som fører til hver av Frida sine responser.

### Forklaring av gjengivelse av chatlogger

Hver chat vil være gjengitt i denne formen:

ID: Identifikasjonsnummer for samtalen		
Hvem	Tekst	Intensjon
<p><i>Angir om det er Frida eller bruker som skriver</i></p>	<p><i>Gjengir det som har blitt skrevet, men navn og datoer er fjernet.</i></p> <p><i>Tekst</i></p> <ul style="list-style-type: none"> <li><i>Knapp</i></li> </ul>	<p><i>Dette feltet angir intensjonen Frida kobler bruker utsagnet opp mot, og hvilket intensjonstre dette hører til, på denne formen:</i></p> <p><i>Intensjonstre/Intensjonsgren</i></p>

Alle knapper i Frida sine svar vil derfor være gjengitt punktvis. Når brukeren trykker på en knapp vil det stå navnet på knappen og handling i parentes slik som dette:

Frida	<p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>Barnebidrag</li> <li>Foreldrepenger</li> </ul>	
-------	---	--

## METODOLOGI

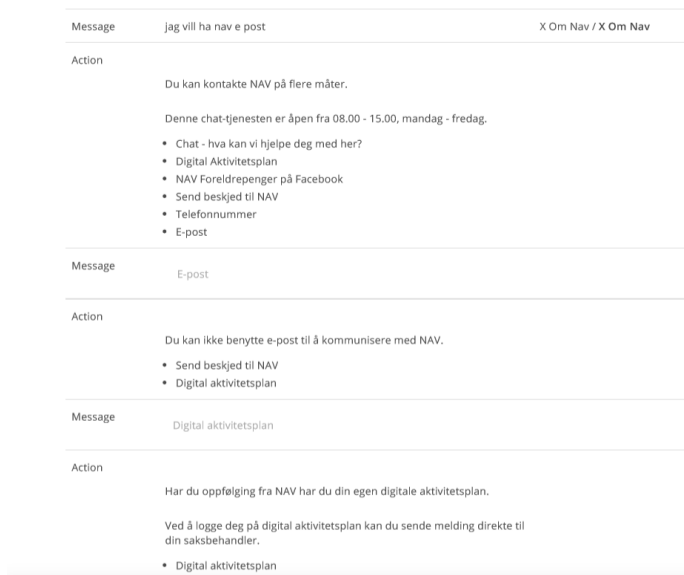
Bruker	Foreldrepenger (handling)	
--------	---------------------------	--

Knappen 'Foreldrepenger' har dermed blitt trykket på.

På feltene som indikerer hva bruker sier så har jeg også lagt til en mørkere farge for å gjøre det lettere å skille hva bruker skriver og hva Frida skriver. I de fleste tilfeller vil hele chatten gjengis. Dersom dette ikke er tilfelle, vil det spesifiseres.

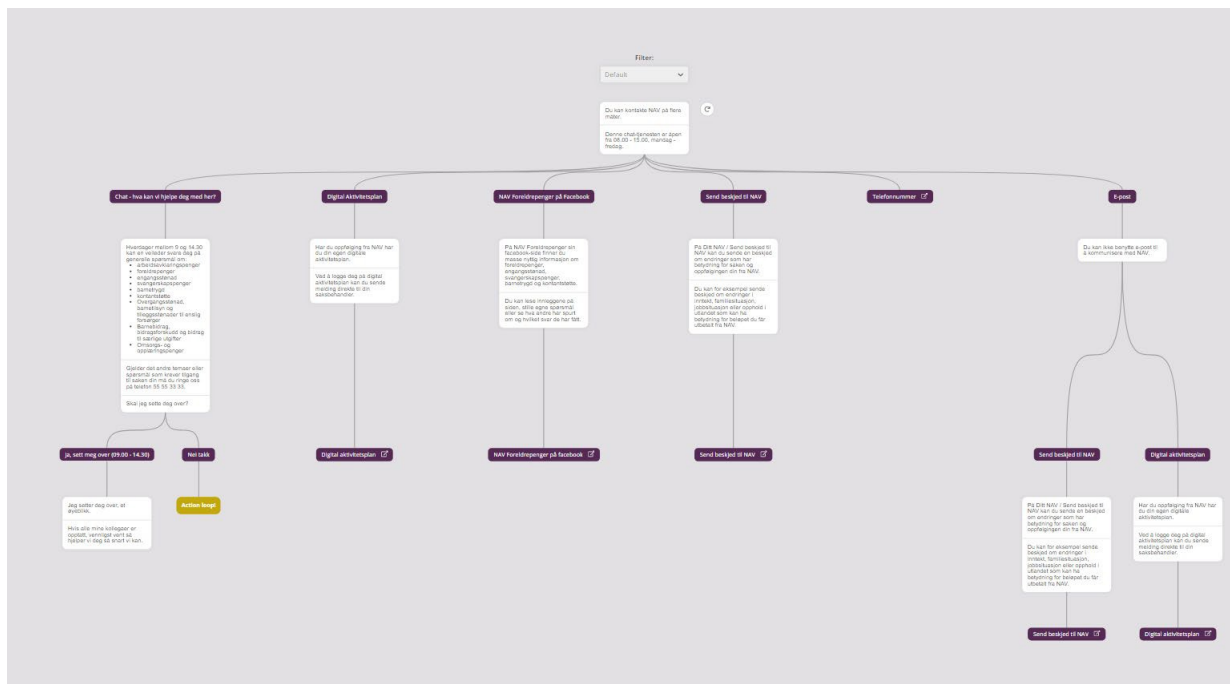
### Intensjoner

Når Frida skal analysere en setning, ser hun etter nøkkelord. Disse nøkkelordene kobles opp mot intensjoner. En intensjon vil være emne Frida kan gi informasjon om, slik som 'Foreldrepenger', 'Engangsstønad' eller 'Generelle spørsmål'. For hver av disse er det intensjonstrær med forskjellige underkategorier en bruker kan spørre om. Informasjon Frida kommer vil derfor være avhengig av hvilken intensjon hun tolker brukerutsagn opp mot. Brukerens oppfølgingsspørsmål vil ofte blir besvart ut ifra intensjonstreet Frida tolket forrige utsagn opp mot. Dette kan gjøre at brukeren får mer tilpasset informasjon, men kan også føre til følgefeil, om Frida tolker det første spørsmålet til brukeren opp mot feil intensjon.



Figur 23: Skjermbilde av utdrag fra chat hvor intensjonen 'Om NAV' forekommer (ID: 125235)

# KAPITTEL 4



Figur 24: Skjerm bilde av intensjonstre 'Om NAV'

Eksempel på hvordan en samtale kan se ut:

ID: 124874		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>For english chat, click here</li> </ul>	
Bruker	Hei! Jeg lurte på om far har rett på permisjon dersom jeg tar engangsstønad og har permisjon fra studiene?	Engangsstønad/ ?Far vilkår
Frida	Ja, i noen tilfeller kan far få engangsstønad.	



## METODOLOGI

	<p>Du kan lese mer om rettighetene til far ved å klikke på linken under.</p> <p>Kom gjerne tilbake til meg hvis det er noe mer du lurer på.</p> <ul style="list-style-type: none"><li>• Fars rettigheter(link)</li></ul>	
--	--	--

I samtalene over kan man se at bruker spør om hva far har rett på, og nevner også 'engangsstønad'. Frida tolker derfor dette til intensjonen 'Engangsstønad'. I intensjonstreet 'Engangsstønad' er det en gren som heter 'Far vilkår'. Frida svarer dermed basert på nøkkelordene 'engangsstønad' og 'far'. Brukeren sitt spørsmål inneholder også andre elementer enn disse to ordene, men Frida kommer ikke med noe informasjon om hvordan det fungerer hvis mor har permisjon ifra studiene. Dette er fordi 'permisjon' ville tilhørt et annet intensjonstre, og Frida analyserer sannsynligheten for at brukerens spørsmål passer med valgt intensjon, basert på tidligere spørsmål fra brukere. I dette tilfellet var dermed sannsynligheten for at bruker stilte et spørsmål om engangsstønad størst og det er derfor dette Frida oppgir informasjon om. Det er derfor større sannsynlighet for at man får svar på det man lurer på om man stiller korte spørsmål og bruker noen få konkrete nøkkelord uten å oppgi så mange detaljer om egen situasjon.

### 4.3.3 Analyse med Suchman sitt ramme verk

Lucy Suchman sitt rammeverk tar for seg hvordan det oppstår utfordringer når bruker utfører handlinger maskinen ikke oppfatter og maskinen utfører handlinger uten at dette er synlig for brukeren. Suchman forsket på samspillet mellom menneske og maskin. Hun hadde to mål med studiene sine; vise at vi alltid handler ut ifra situasjonen vi er i, og tydeliggjøre forholdet mellom hvordan vi tolker en handling og forutsetningene for handlingen. Hennes mest kjente eksempel er når to forskere skal bruke en kopimaskin for å kopiere tosidig. (Bratteteig & Herstad, 2020) Hun satte opp en tabell med en sekvens av handlinger for å kartlegge hvor problemet oppstod.

## KAPITTEL 4

### Sekvens av handlinger mellom bruker og Frida

*Sekvensdiagram der tidsaksen går nedover*

Bruker		Maskin	
Handling ikke tilgjengelig for maskin	Handling tilgjengelig for maskin	Effekt synlig for bruker	Designrasjonale

Figur 25: Illustrasjon av Suchman sitt analytiske rammeverk basert på (Suchman, 2007)s.123

Når vi har gått gjennom chatloggene, har vi kun loggene å gå ut ifra. Vi vet ikke hva brukeren gjorde utenom dette, slik som å søke på nettet, sjekke linkene Frida kommer med, eller spørre noen andre om hjelp. Det eneste aspektet vi kan se av hva brukeren gjør som ikke Frida ser, er tidsaspektet. Dette vil Frida kun ha en respons til hvis det går mer enn 30 min.

Men det vi kan se tydelig når vi går gjennom chatloggene er hva som skjer i maskineriet som ikke er synlig for brukeren. Og mange av utfordringene som oppstår i chattene, skjer fordi brukeren ikke ser hva som skjer. Brukeren kan ikke se hvilken intensjon Frida kobler spørsmålene til. Selv om det blir opplyst om at hun bruker nøkkelord, så vet ikke brukeren noe om hvordan Frida analyserer setninger og at de kobles til intensjonstrær. Dette kan føre til både forvirring og irritasjon.

I et vellykket eksempel vil det ofte se slik ut:

ID: 125287 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurer du på?  For english chat, click here	
Bruker	hvordan kan jeg søke om fredrekvote	Foreldrepenger / Hvordan søker jeg fedrekvote
Frida	Det enkleste er å bruke den digitale søknaden om foreldrepenger du finner på nav.no	

## METODOLOGI

	<p>Hvis du er arbeidstaker må i tillegg arbeidsgiveren din sende en digital inntektsmelding til NAV.</p> <ul style="list-style-type: none"> <li>• Søknadsfrist</li> <li>• Fleksibelt uttak</li> <li>• Søk her</li> </ul>	
--	--	--

Brukeren har stavet fedrekvote feil, allikevel kobler Frida det opp mot riktig intensjon, og brukeren får svar på det vedkommende lurer på.

Men om Frida kobler det opp mot feil intensjon og brukeren ikke forstår hva som foregår, kan det føre til problemer:

ID: 125247 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	[Frida informerer om foreldrepenger]	
	•	
Bruker	Hva er det her?	X Generelle spørsmål/ Er du der?
Frida	Ja, jeg er her! Hva lurer du på?	
Bruker	du sender jo bare masse tull	Foreldrepenger/ Hvor lang er fellesperioden
Frida	[Frida informerer om fellesperioden]	

*Utdrag fra chat*

Her kan man tydelig se at intensjonene Frida kobler brukerens utsagn til, ikke stemmer med det brukeren skriver. Brukeren blir oppgitt over Frida sitt svar og kan heller ikke se hvorfor denne feilen skjer.

Hentet fra veileder kapittel:

ID: 125007 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.	

## KAPITTEL 4

	<p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, <a href="#">click here</a></li> </ul>	
Bruker	Det er dessverre litt vanskelig å formulere kortere, kan jeg komme i kontakt med et menneske?	X Om Nav / Snakke med menneske
Frida	<p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat (knapp)</li> <li>• Telefonnummer (knapp)</li> </ul>	
Bruker	Chat(handling)	
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenge</li> <li>• engangsstønad</li> <li>• svangerskapspenge</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringspenge</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Sett meg over (08.00 - 15.00)</li> <li>• Nei takk</li> </ul>	

## METODOLOGI

Bruker	gjerne	X Generelle spørsmål / Hei
Frida	Hei! Hva kan jeg hjelpe deg med?	
Bruker	Hei!	Unknown
Frida	Jeg svarte kanskje ikke helt på det du lurte på? Kanskje du kan spørre litt annerledes?	
Bruker	Min samboer er selvstendig næringsdrivende, og har ikke kunnet tatt ut lønn så langt i år	Unknown
Frida	Beklager, men jeg forstår ikke hva du spør om.  Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.  Kan du stille spørsmålet på nytt? <ul style="list-style-type: none"> <li>• Trenger mer hjelp(knapp)</li> <li>• Nei(knapp)</li> </ul>	
Bruker	Beklager, tror ikke du kan hjelpe meg	X Generelle spørsmål / Hjelp
Frida	Hva kan jeg hjelpe deg med?	

I Suchman sitt rammeverk:

Handling ikke tilgjengelig	Handling tilgjengelig for maskin	Effekt synlig for bruker	Designrasjonale
		(Frida spør om bruker vil bli satt over)	
	gjerne		Intensjon: <i>X Generelle spørsmål / Hei</i>  Utsagnet 'gjerne' fører til at Frida kobler det til et nytt intensjonstre og

## KAPITTEL 4

			bruker blir ikke satt over
		Hei! Hva kan jeg hjelpe deg med?	
	Hei!		<i>Unknown</i>  Frida klarer ikke å koble det opp mot en intensjon
		Jeg svarte kanskje ikke helt på det du lurte på? Kanskje du kan spørre litt annerledes?	
	Min samboer er selvstendig næringsdrivende, og har ikke kunnet tatt ut lønn så langt i år		<i>Unknown</i>
		Beklager, men jeg forstår ikke hva du spør om.  Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.  Kan du stille spørsmålet på nytt?  <ul style="list-style-type: none"> <li>• Trenger mer hjelp(knapp)</li> <li>• Nei(knapp)</li> </ul>	
	Beklager, tror ikke du kan hjelpe meg		<i>X Generelle spørsmål / Hjelp</i>

## METODOLOGI

			Frida kobler det opp mot grenen 'hjelp' under intensjonstreet 'Generelle spørsmål'
		Hva kan jeg hjelpe deg med?	

Tabell 2: Chat i Suchman sitt rammeverk

Her ser ikke bruker at Frida ikke setter over på grunn av ordet 'jerne'. Hadde bruker i stedet skrevet 'ja' eller klikket på 'sett meg over', så hadde brukeren kommet til en veileder. Her forsøker brukeren å stille spørsmålet sitt i stedet fordi den ikke ser at den ikke er blitt satt over. Brukeren gjør heller ikke noe nytt forsøk på å bli satt over når den skjønner at den fortsatt snakker med chatboten, antagelig fordi den ikke skjønner hva som gikk galt i utgangspunktet. Hadde det her vært tydeligere for bruker hva som egentlig skjedde, hadde det vært lettere å skjønne for bruker hva som er nødvendig å gjøre for å bli satt over.

Suchman sitt rammeverk er derfor nyttig for å se på hva som foregår i maskineriet som brukeren ikke får med seg. Jeg vil gå nærmere inn på hvordan dette fungerer når jeg senere går gjennom de forskjellige kategoriene.

## 4.4 Formelle tillatelse for studien

### Taushetserklæring

NAV har et krav om at man skriver under en taushetserklæring før man får gjennomføre intervjuer og observasjoner hos NAV. Ved å skrive under på denne avtalen bekreftet jeg at jeg forstod min taushetsplikt, og at jeg ikke ville avsløre opplysninger om andre. Jeg bekreftet også at jeg ville hindre uvedkommende fra å få tilgang eller kjennskap til taushetsbelagte opplysninger.

### Databehandleravtale

En databehandleravtale skal sikre at personopplysninger blir behandlet i samsvar med regelverket og avklare hvordan personopplysningene skal behandles av databehandleren.

## KAPITTEL 4

(Datatilsynet, 2018) Jeg skrev under en databehandleravtale før jeg fikk tilgang til loggene. Denne inneholdt beskrivelse av hvordan behandlingen av data skulle foregå, med tanke på formål, oppbevaring og varighet. Ved å skrive under forpliktet jeg meg til å ikke behandle personopplysninger på en annen måte enn det som stod skrevet i avtalen.

### **Behandling av personopplysninger**

NSD(Norsk Senter for Forskningsdata) har en avtale med Universitetet i Oslo om at studier som inneholder behandling av personopplysninger må meldes hit. (NSD, 2020a) En personopplysning vil si en opplysning som kan identifisere en enkeltperson. Jeg ønsket til denne studien å ta lydopptak og innhente signert samtykke, og måtte derfor melde studien til NSD. Samtykkeskjemaene er utformet etter NSD sin mal(NSD, 2020b). Både samtykkeskjemaene og NSD sin vurdering av denne studiens behandling av personopplysninger, er lagt ved som vedlegg.

### **Etiske hensyn**

Jeg har anonymisert dataene jeg har samlet inn. Jeg har brukt pseudonymer i transkripsjonene og fjernet alt av personlig informasjon som navn, og sensitiv informasjon i chattene jeg gjengir.



## 5 Kategori: Vellykkede samtaler

NAV anser en samtale som vellykket hvis brukeren enten har fått svar på det vedkommende lurte på eller blitt riktig henvist til enten nettside, skjema, organisasjon eller til chat med menneskelig veileder.

Vi har derfor valgt å dele inn vellykkede chatter inn i tre ulike kategorier, *riktig svar*, *riktig henvisning til skjema eller nettside* og *riktig henvisning til veileder*. I riktig svar-kategorien har vi plassert de chat-samtalene der bruker har stilt et spørsmål og, som etter vår mening, har fått et tilstrekkelig godt svar. I riktig henvisning-kategorien har vi valgt å plassere de samtalene der chatboten, slik vi ser det, enten har satt over brukeren til en veileder på riktig sted, eller henvist brukeren til riktig nettside, søknadskjema eller organisasjon.



## 5.1 Riktig svar

Dette er chatter hvor Frida besvarer spørsmålet bruker spør om og bruker gir inntrykk av å være fornøyd med svaret.

ID: 125312 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	<p>Hei. Jeg har søkt pappapermisjon fra &lt;dato&gt;e mai, men nå er jeg sykmeld med rygge mitt</p>	Foreldrepenger/ Utsettelse egen sykdom fedrekvote
Frida	<p>Du kan utsette foreldrepengene hvis du er syk og er helt avhengig av hjelp til å ta deg av barnet. Du må søke om utsettelse, og legge ved en funksjonsbeskrivelse fra lege. Vanlig sykemelding er ikke tilstrekkelig.</p> <p>Dette må legen din ta med i funksjonsbeskrivelsen</p> <ul style="list-style-type: none"> <li>• Hvordan sykdommen påvirker evnen din til å ta deg av barnet</li> <li>• I hvilken grad du er avhengig av hjelp</li> <li>• Hvilken periode du er avhengig av hjelp</li> </ul> <p>Den andre forelderen kan som et alternativ til utsettelse, søke om å overta permisjonen din. Kravet er det samme som for utsettelse, du må være syk og helt avhengig av hjelp til å ta deg av barnet.</p> <ul style="list-style-type: none"> <li>• Søk her</li> </ul>	
Bruker	<p>Oke, vi skal gjøre dette, tusen takk for deres forte svar</p>	X Generelle spørsmål/ Takk
Frida	<p>Bare hyggelig</p>	

Her får bruker tilsynelatende svar på det vedkommende lurer på. Brukeren i eksempelet har søkt pappapermisjon, men har blitt sykmeldt. Chatboten svarer at det er mulig å utsette

## VELLYKKEDE SAMTALER

foreldrepengene eller la den andre forelderen overta permisjonen ved sykdom, dersom den som er sykmeldt er avhengig av hjelp til å ta seg av barnet. Chatboten opplyser bruker om at en søknad om dette må inneholde en funksjonsbeskrivelse fra en lege. Bruker takker så chatboten for svar, og skriver at vedkommende skal gjøre dette. Her ser man at chatboten oppfatter spørsmålet til tross for skrivefeil og klarer å gi riktig informasjon, og brukeren virker fornøyd med svaret. Men måten brukeren takker Frida på kan tyde på at brukeren ikke oppfattet at det var en chatbot brukeren snakket med.

## 5.2 Riktig henvisning til skjema/nettside

Frida henviser bruker til riktig skjema eller sted på nettsiden.

ID: 125261(14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurer du på? <ul style="list-style-type: none"><li>• For english chat, click here</li></ul>	
Bruker	Jeg har krysset av feil på søknad	Generelle spørsmål / Endre søknad
Frida	Hvis du vil endre på en innsendt søknad, må du som hovedregel sende en ny søknad. Det samme gjelder om du har gjort en feil.  Hva du skal søke om, avhenger av hvilken endring du vil gjøre.  Det samme gjelder om det er endringer i en sak du allerede har fått vedtak i. <ul style="list-style-type: none"><li>• Endring av foreldrepenger</li><li>• Trekke en søknad</li><li>• Søk her</li></ul>	

## KAPITTEL 5

Bruker	Trekke en søknad(handling)	Generelle spørsmål / Endre søknad
Frida	<p>Hvis du vil trekke en søknad du har sendt til NAV, kan du gjøre det ved å bruke tjenesten "Send beskjed til NAV".</p> <p>Husk at du må logge inn med BankID for å bruke den tjenesten.</p> <ul style="list-style-type: none"> <li>• Send beskjed til NAV</li> </ul>	

Her har en bruker krysset feil i en søknad. Frida opplyser om at hvis du vil gjøre endringer må du sende søknad på nytt. Bruker klikker så på 'Trekke en søknad' og Frida forklarer at man trekker en søknad ved å bruke tjenesten 'Send beskjed til NAV'. I denne chatten får brukeren forklaring på hva man gjør hvis man har gjort en feil i en søknad og henviser brukeren til riktig sted for å trekke søknad. Det er derfor en vellykket chat.

### 5.3 Riktig henvisning til veileder

I noen tilfeller skal Frida foreslå å sette over til veileder, enten fordi spørsmålet krever personopplysninger eller befinner seg utenfor de temaene Frida kan svare på. Disse samtalene vil derfor være vellykket om Frida forstår at dette noe veileder skal svare på.

ID: 616539/ 161084 (???.?.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hei hvordan er prosessen ved dødfødsel	X Generelle spørsmål / Død
Frida	<p>Beklager, da er det bedre om du snakker med mine kolleger.</p> <p>Mellom 08.00 og 15.00 kan du chatte med en av mine kollegaer.</p>	

## VELLYKKEDE SAMTALER

	<p>Du kan også snakke med en av oss på telefon. Da ringer du 55 55 33 33, telefonen er åpen mellom kl 8.00 - 15.30.</p> <ul style="list-style-type: none"> <li>• Chat - hva kan vi svare på her? (knapp)</li> </ul>	
Bruker	Chat - hva kan vi svare på her? (handling)	
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapsenger</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringspenger</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag - Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Ja, sett meg over (08.00 - 15.00) (knapp)</li> <li>• Nei takk (knapp)</li> </ul>	
Bruker	Ja, takk	Sett meg over
Frida	<p>Jeg setter deg over, et øyeblikk.</p> <p>Hvis alle mine kollegaer er opptatt, vennligst vent så hjelper vi deg så snart vi kan.</p>	

I dette eksempelet blir brukeren satt over til chat med veileder. Her hadde brukeren et spørsmål om dødfødsel. Ordet 'dødfødsel' er lagt inn som et nøkkelord hvor Frida direkte skal foreslå å sette over til veileder. Her henviser chatboten derfor riktig ved å overføre brukeren til menneskelig chat.

## 5.4 Oppsummering

I dette kapitlet har jeg chatter som jeg klassifiserer som vellykket. En chat havner i denne kategorien både når Frida henviser riktig og svarer riktig på det brukeren spør om

## 6 Når samtale settes over til veileder

Veilederne i NAV har ansvaret for den daglige kontakten med brukerne. De er tilgjengelige på chat og på telefon. En bruker som kontakter NAV via telefon, vil komme direkte til en veileder. En bruker som tar kontakt via chat, derimot, vil først settes i kontakt med Frida. Denne samtalen kan deretter bli satt over til en veileder, av ulike grunner.

Før kunne brukere velge å bli satt direkte over til veileder, men denne muligheten ble fjernet siden så mange valgte dette. Men det finnes fortsatt måter å lure systemet på. De såkalte ‘superbrukerne’ trykker på knappen ‘Engelsk’ for å bli direkte satt over.

En veileder kan ha opptil tre samtaler samtidig, men brukeren får allikevel svar relativt fort etter å ha blitt satt over. Veilederen vil først lese gjennom samtalen brukeren har hatt med Frida og så fortsette samtalen ved enten å kopiere inn ferdige fraser fra et word-dokument eller å skrive inn svar. Veilederen legger ofte vekt på å bruke muntlig språk for å gjøre det lettere for brukeren å forstå regelverket til NAV.

### Når samtale settes over

En samtale hvor Frida ikke forstår spørsmålet eller ikke kan svare, vil bli satt over til en veileder. Veilederen vil da kunne lese gjennom samtalen brukeren har hatt med Frida, før veilederen forsøker å besvare spørsmålet. Spørsmål som krever personopplysninger skal alltid rettes direkte til veileder, og komplekse spørsmål må også stilles til et menneske. Her er det meningen at Frida tidlig skal spørre om brukeren vil bli satt over til et menneske. Enkelte nøkkelord, sånn som ‘død’, vil også gjøre at samtalen blir satt over. Det kan være både brukeren som uttrykker ønske om å bli satt over og Frida som foreslår det.

### 6.1 Frida foreslår å sette over

Det er forskjellige grunner til at Frida foreslår å sette over. De vanligste er spørsmål som krever personopplysninger, spørsmål som kobles til samme intensjon som forrige spørsmål og spørsmål som ikke kan kobles opp mot noen intensjon(‘unknown’).

## KAPITTEL 6

### Spørsmål som krever personopplysninger

Spørsmål som krever personopplysninger skal alltid settes over til veileder ettersom at Frida ikke er en innlogget tjeneste. Ved spørsmål om status i sak og andre forhold som krever innsyn i personopplysninger, skal Frida derfor til foreslå å sette over til en veileder.

ID: 124953(14.09.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hei Frida, jeg har sendt inn en søknad om foreldrepenger men fortsatt ikke fått noe svar	Foreldrepenger / Saksbehandlingstid foreldrepenger
Frida	<p>Saksbehandlingstiden kan variere noe fra fylke til fylke.</p> <p>Du kan selv sjekke status i din sak, klikk "Status i sak" for mer informasjon.</p> <ul style="list-style-type: none"> <li>• Status i sak</li> <li>• Sjekk behandlingstid</li> </ul>	
Bruker	Status i sak(handling)	
Frida	<p>Jeg kan dessverre ikke svare deg på spørsmål som krever personopplysninger som for eksempel status i sak, innsendte dokumenter eller spørsmål om utbetaling.</p> <p>Det skyldes at denne tjenesten ikke er en innlogget tjeneste.</p> <p>Du kan enkelt følge saken din på Ditt NAV. Her finner du søknaden du har sendt oss og oversikt over alle dokumenter som er mottatt i saken.</p> <p>På Ditt NAV finner du også oversikt over vedtak og utbetalinger fra NAV.</p> <p>Du kan også ringe oss for hjelp på telefon 55 55 33 33 tastevalg 3 familie.</p> <ul style="list-style-type: none"> <li>• Saksbehandlingstider</li> </ul>	



## VELLYKKEDE SAMTALER

	<ul style="list-style-type: none"> <li>• Ditt NAV</li> <li>• Utbetaling</li> </ul>	
Bruker	vil snakket med et menneske	X Om Nav / Snakke med menneske
Frida	<p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	

I dette eksempelet lurer brukeren på hvorfor vedkommende ikke har fått svar på søknaden om foreldrepenger. Brukeren trykker så på alternativet 'status i sak'. Frida forklarer at hun ikke kan svare på dette ettersom at dette ikke er en innlogget tjeneste. Etter å ha fått tilbudet om å bli satt over til chat så avslutter brukeren samtalen. Her kan man anta at brukeren ringte telefonnummeret som Frida oppgir i stedet, for å komme i kontakt med et menneske. I slike tilfeller hvor brukeren lurer på noe som Frida ikke kan svare på, vil det alltid være riktig av Frida å informere om hvorfor hun ikke kan svare på det og foreslå å sette over. Dette er derfor, ifølge chatbot-trenerne, et eksempel på en vellykket samtale.

### Spørsmål som kobles til sammen intensjon

Når bruker stiller et spørsmål som vil resultere i et likt svar fra Frida(samme intensjon) som ved forrige spørsmål vil Frida foreslå å sette over. Brukeren kan ha omformulert spørsmålet sitt, men det inneholder gjerne de samme nøkkelordene og vil derfor kobles om mot den samme intensjonen. Frida vil da som regel respondere 'Jeg tror jeg har forsøkt å svare på dette spørsmålet, vil du settes over til veileder?'. Men om spørsmål med de samme nøkkelordene stilles med noen andre spørsmål imellom, så vil hun ofte bare gjenta svarene sine. Om brukeren varierer mellom to-tre like spørsmål, kan derfor samtalen gå i loop.

ID: 204183 (02.12.2019)		
Hvem	Tekst	Intensjon

## KAPITTEL 6

	<i>Bruker lurer på hva som skjer med AAP hvis man reiser til utlandet.</i>	
Frida	Når du har AAP, skal du ha jevnlig oppfølging. Som en del av oppfølgingen, skal det vurderes om vilkårene for AAP fortsatt er oppfylt.	
Bruker	Hva hvis jeg ikke får noe oppfølging?	AAP/ Oppfølging AAP
Frida	Som mottaker av AAP skal du ha jevnlig oppfølging fra veilederen din.  Hva som inngår i oppfølgingen vil være individuelt, og kan variere fra person til person.	
Bruker	Jeg får ingen oppfølging	AAP/ Oppfølging AAP
Frida	Jeg tror jeg har forsøkt å svare på dette, men det kan virke som at det ikke ble så vellykket som jeg ønsket. Jeg lærer hver dag og håper jeg kan svare bedre neste gang du chatter med meg  Hvis du ønsker kan du snakke med en av kollegaene mine <ul style="list-style-type: none"> <li>• Ønsker å snakke med et menneske</li> </ul>	
Bruker	Ønsker å snakke med et menneske(handling)	
Frida	Du kan selvsagt få snakke med en av mine kolleger.  Hvordan ønsker du å snakke med oss? <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefon</li> </ul>	

I dette eksempelet lurer prøver brukeren å formidle at vedkommende ikke får noe oppfølging. Dette oppfatter ikke Frida. Hun tolker begge utsagnene som Oppfølging AAP, og derfor sier hun at hun allerede har forsøkt å svare på dette spørsmålet. Her foreslår derfor Frida å sette over til veileder. I loggene ifra mai, vil slike samtaler som regel gå i loop, men på grunn av endringer lagt til før denne chatten i desember, så foreslår Frida å sette over når samme intensjon blir gjentatt.

### **Spørsmål som ikke kobles opp mot noen intensjon**

## VELLYKKEDE SAMTALER

Når brukerens utsagn ikke kan kobles opp mot noen intensjon ender Frida opp med intensjonen 'unknown'. Da vil Frida svare 'Jeg forstår ikke hva du mener, kan du stille spørsmålet på nytt?' Hvis man så svarer at man trenger mer hjelp så foreslår hun å sette over.

ID: 124913 (14.09.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hei. Eg lurte på om det var mulig å få støtte til omskolering	Unknown
Frida	<p>Beklager, men jeg forstår ikke hva du spør om.</p> <p>Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.</p> <p>Kan du stille spørsmålet på nytt?</p> <ul style="list-style-type: none"> <li>• Trenger mer hjelp</li> <li>• Nei</li> </ul>	
Bruker	Trenger mer hjelp(handling)	
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapsenger</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringsenger</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p>	

## KAPITTEL 6

	<p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Sett meg over (08.00 - 15.00)</li> <li>• Nei takk</li> </ul>	
--	---	--

I dette eksempelet er det en bruker som lurer på om det er mulig å få støtte til omskolering. Dette klarer ikke Frida å koble opp mot noen intensjon ettersom at det bruker skriver ikke inneholder noen nøkkelord som hun kan koble opp mot intensjoner. Frida ender derfor opp med ‘unknown’ og forklarer at hun ikke forstår spørsmålet. Når bruker klikker på ‘Trenger mer hjelp’ så forklarer hun hva hun kan svare på, og foreslår å sette over.

## 6.2 Bruker spør om å bli satt over

### Bruker spør om å bli satt over etter ikke å ha blitt forstått

I disse samtalene oppfatter Frida som regel bare deler av spørsmålet og kommer med svar som er irrelevante for det brukeren lurer på. Bruker uttrykker irritasjon og Frida spør om brukeren kan omformulere spørsmålet, men bruker spør da i stedet om å få snakke med et menneske.

ID: 124990 (14.09.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hvem har omsorgsdagene når mor og far ikke bor sammen?	Foreldrepenger / Dersom bare mor har rett
Frida	<p>I noen tilfeller kan mor få hele perioden med foreldrepenger.</p> <p>Hun kan enten være alene om omsorgen, eller bo sammen med barnets andre forelder som ikke har tjent opp rett til foreldrepenger.</p>	

## VELLYKKEDE SAMTALER

	<ul style="list-style-type: none"> <li>• Alene om omsorgen</li> <li>• Bor med far/medmor</li> <li>• Jeg mener barnetrygd</li> <li>• Jeg mener overgangsstønad</li> </ul>	
Bruker	Nei	X Generelle spørsmål / Nei
Frida	OK. Spør meg gjerne på nytt hvis du lurer på noe mer.	
Bruker	Ønsker å snakke med en person	X Om Nav / Snakke med menneske
Frida	<p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	

I denne chatten lurer bruker på hvordan omsorgsdager fordeles når man ikke bor sammen.

Frida tolker dette til intensjonen 'Dersom bare mor har rett' og forklarer hvordan mor kan få hele perioden med foreldrepenger, noe som ikke besvarer brukerens spørsmål. I stedet for å prøve å omformulere spørsmålet skriver brukeren ganske enkelt 'Nei', og spør deretter om å bli satt over.

### **Bruker spør om å bli satt over etter å ha fått riktig/tilstrekkelig svar**

Noen brukere stiller et spørsmål og får et synelatende tilstrekkelig svar, men spør deretter om å bli satt over. Det kan være at det er noe annet de vil spørre veilederen om, men dette kan også tyde på at de ikke stoler på at det svaret de får ifra Frida stemmer.

ID: 125210 (14.09.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p>	

## KAPITTEL 6

	<p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	kan jeg si opp barnehagen jeg har fått tildelt å heller søke kontantstøtte?	Kontantstøtte/ Kontantstøtte
Frida	<p>Kontantstøtte er en skattefri ytelse som kan gis for barn mellom 1 og 2 år (13 - 23 måneder) som ikke har plass i barnehage som mottar offentlig driftstilskudd.</p> <p>Redusert plass:</p> <ul style="list-style-type: none"> <li>• tom juli 2018 kan det gis redusert kontantstøtte for barn som har redusert plass (19 timer eller mindre timer per uke)</li> <li>• fom august 2018 kan det gis redusert kontantstøtte for barn som har redusert plass (32 timer eller mindre per uke).</li> </ul> <p>Den som søker må ha minst 5 års trygdetid i Norge og/eller EØS-land for å ha rett til kontantstøtten. Hvis barnet bor sammen med begge foreldrene, må begge foreldrene fylle vilkåret om 5 års trygdetid.</p> <ul style="list-style-type: none"> <li>• Om søknaden og hvordan søke</li> <li>• Stanse kontantstøtte</li> <li>• Hvor lenge kan jeg mottatt kontantstøtte?</li> <li>• Forholdet til foreldrepenger</li> <li>• Tildelt plass</li> <li>• Mer om kontantstøtte</li> </ul>	
Bruker	Tildelt plass(handling)	
Frida	<p>Det er tildelte timer per uke og ikke hvor mye barnehageplassen blir brukt som avgjør retten til kontantstøtte.</p> <p>Hvis du har tildelt plass du ikke bruker, må det dokumenteres at plassen brukes av et annet barn for at du skal kunne ha rett til kontantstøtte.</p> <p>Har du spørsmål om barnehageplassen og antall timer du har tildelt, kontakter du kommunen der du bor eller barnehagen.</p>	
Bruker	kan jeg få prate med en veileder?	X Om Nav / Snakke med menneske

## VELLYKKEDE SAMTALER

Frida	<p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	
Bruker	Chat (handling)	

I dette tilfellet lurer brukeren på om man kan si opp barnehageplassen og heller søke kontantstøtte. Frida sier at når et annet barn har tatt over plassen, så får man rett til kontantstøtte. Dette skulle i utgangspunktet besvare brukerens spørsmål, allikevel spør brukeren om å bli satt over til veileder. Det kan være at brukeren hadde oppfølgingsspørsmål som vedkommende følte ikke kom frem i svaret til Frida og ikke stolte på at Frida kunne svare på, og derfor vil snakke med veileder. Men det kan også være at brukeren tviler på svarene fra Frida, og ville forhøre seg med veileder om det stemte.

### **Bruker spør umiddelbart etter menneske, uten å prøve å spørre Frida**

Brukere som har snakket med Frida tidligere vet hvordan de skal få snakket med Frida og de skriver enten 'veileder' eller 'menneske'. Noen av dem trykker også på 'engelsk' med en gang for å komme i kontakt med veileder, og skriver så på norsk når de kommer til en veileder.

ID: 124985 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	

## KAPITTEL 6

Bruker	Snakke med menneske	X Om Nav / Snakke med menneske
Frida	<p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	
Bruker	Chat(handling)	
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenge</li> <li>• engangsstønad</li> <li>• svangerskapspenge</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringspenge</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Sett meg over (08.00 - 15.00)</li> <li>• Nei takk</li> </ul>	
Bruker	Ja	
Bruker	Sett meg over (08.00 - 15.00) (handling)	



## VELLYKKEDE SAMTALER

Frida	Jeg setter deg over, et øyeblikk.  Hvis alle mine kollegaer er opptatt, vennligst vent så hjelper vi deg så snart vi kan.	
-------	---	--

Denne brukeren spør om å bli satt over med en gang, uten å forsøke å spørre Frida først. Dette kan skyldes mangel på tillit til chatboten, men det kan også være at brukeren har et spørsmål den vet chatboten ikke kan hjelpe med.

### 6.3 Brukere som har problemer med å bli satt over

I enkelte chatter ser man eksempler på brukere som forsøker å bli satt over, men som ikke blir satt over av Frida. Dette er problematisk fordi brukere ikke kommer til en veileder uten å gå via Frida først. Brukere som ikke vet om snarveien ved at man kan trykke 'English' har ikke noe annet alternativ enn å prøve å få Frida til å sette dem over dersom de vil chatte med en veileder.

ID: 125007 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurer du på?  <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Det er dessverre litt vanskelig å formulere kortere, kan jeg komme i kontakt med et menneske?	X Om Nav / Snakke med menneske
Frida	Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.  Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.  Klikk "Chat" for informasjon om chat-tjenesten vår.  <ul style="list-style-type: none"> <li>• Chat (knapp)</li> <li>• Telefonnummer (knapp)</li> </ul>	
Bruker	Chat(handling)	

## KAPITTEL 6

Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapsenger</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringsenger</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Sett meg over (08.00 - 15.00)</li> <li>• Nei takk</li> </ul>	
Bruker	gjerne	X Generelle spørsmål / Hei
Frida	<p>Hei!</p> <p>Hva kan jeg hjelpe deg med?</p>	
Bruker	Hei!	Unknown
Frida	<p>Jeg svarte kanskje ikke helt på det du lurte på? Kanskje du kan spørre litt annerledes?</p>	
Bruker	Min samboer er selvstendig næringsdrivende, og har ikke kunnet tatt ut lønn så langt i år	Unknown
Frida	<p>Beklager, men jeg forstår ikke hva du spør om.</p> <p>Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.</p> <p>Kan du stille spørsmålet på nytt?</p> <ul style="list-style-type: none"> <li>• Trenger mer hjelp(knapp)</li> <li>• Nei(knapp)</li> </ul>	

## VELLYKKEDE SAMTALER

Bruker	Beklager, tror ikke du kan hjelpe meg	X Generelle spørsmål / Hjelp
Frida	Hva kan jeg hjelpe deg med?	

Her er det en person som gjerne vil bli satt direkte over til menneske fordi vedkommende har et spørsmål som det er vanskelig å formulere kort. Frida tilbyr brukeren å snakke med et menneske, men når brukeren svarer 'gjerne' blir dette tolket som 'Hei' og brukeren blir ikke satt over. Dette er også lett å misforstå for brukeren ettersom at brukeren venter på å bli satt over. I denne chatten ser det til at brukeren tror den har blitt satt over siden brukeren forsøker å stille spørsmålet sitt. Når brukeren skjønner at det er en chatbot så skriver brukeren 'Beklager, tror ikke du kan hjelpe meg'. I dette tilfelle handler det ikke bare om mangel på tillit fra brukerens side, ettersom at chatboten ikke klarer å svare på spørsmålet når brukeren spør heller. Chatten slutter her og brukeren blir heller ikke satt over. Men hadde brukeren bare trykke på 'Sett meg over' i stedet for å skrive 'gjerne', så hadde brukeren blitt satt over til veileder.

ID: 169479 (???.09.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Hei! Jeg er chatboten Frida og jeg svarer deg på generelle spørsmål om arbeidsavklaringspenger, foreldrepenger og andre ytelser angående barn og familie. Hva lurer du på?</p> <p>Får du ikke svar på det du spør om, kan det skyldes at jeg er under opplæring.</p> <p>Husk at du ikke må gi meg personlig informasjon som navn og personnummer.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• <a href="#">English? Click here</a></li> </ul>	
Bruker	Får jeg opptjent rett til sykepenger når jeg har foreldrepermisjon	Foreldrepenger / Opptjening likestilt med arbeid
Frida	Enkelte ytelser fra NAV kan gi opptjening av rett til foreldrepenger på lik linje med arbeid. Klikk	

## KAPITTEL 6

	<p>på "ytelse fra NAV" under for å lese mer om dette.</p> <p>Perioder du har vært i militær- eller obligatorisk sivilforsvarstjeneste telles som perioder med arbeid.</p> <ul style="list-style-type: none"> <li>• Opptjening på nav.no</li> <li>• Jeg har utbetaling fra NAV</li> </ul>	
Bruker	Jeg har utbetaling fra NAV(handling)	
Frida	<p>Likestilt med yrkesaktivitet er tidsrom du har mottatt</p> <ul style="list-style-type: none"> <li>• dagpenger</li> <li>• sykepenger</li> <li>• stønad ved barns sykdom m. m. etter kapittel 9 i folketrygdeloven</li> <li>• arbeidsavklaringspenger</li> <li>• foreldrepenger eller svangerskapspenger</li> </ul> <p>For å oppfylle inntektskrav må beregningsgrunnlaget ditt utgjøre minst 49 929 kroner (1/2 G)</p> <p>Startet foreldrepengeperioden din 30.4.19 eller tidligere måtte inntekten din opptjeningsperioden omregnet til årsinntekt være minst 48 441 kroner (1/2 G)</p>	
Bruker	Dette var ikke svar på spørsmål	X Generelle spørsmål / Hei
Frida	Hei	
Bruker	Kan du sette meg i kontakt med veileder?	??
Frida	Jeg kan svare på spørsmål om meg selv	
Bruker	ha det	X Generelle spørsmål / Ha det bra
Frida	<p>Ha det bra</p> <p>Takk for praten</p>	

I denne chatten lurer bruker på om man får opptjent rett til sykepenger når man har foreldrepermisjon, men Frida tolker det som intensjonen 'Foreldrepenger / Opptjening likestilt med arbeid'. Bruker forsøker å trykke på en av knappene i Frida sitt svar og får videre

## VELLYKKEDE SAMTALER

informasjon om opptjening av foreldrepenger. Brukeren skriver at dette ikke var svar på spørsmålet, men Frida tolker dette som intensjonen 'Hei'. Når bruker spør om å bli satt over til veileder, så svarer Frida at hun kan svare på spørsmål om deg selv og brukeren avslutter deretter samtalen. Slike tilfeller som dette forekommer ikke ofte, ettersom at nøkkelordet 'veileder' som regel fører til at brukeren blir direkte satt over. Allikevel er det enkelte tilfeller hvor andre nøkkelord blir prioritert av Frida. Flere av intensjonene i denne chatten ser ut til å ha blitt endret på i etterkant, så nå ville denne brukeren antagelig ha blitt satt over.

ID: 125181 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Veileder nav	X Om Nav / Timeavtale
Frida	<p>Du får ikke timeavtale til generell veiledning og søknadshjelp.</p> <p>Unntak er ved fastsetting av farskap. Da må du kontakte oss på 55 55 33 33, tastevalg familie for hjelp.</p> <p>Har du arbeidsrettet oppfølging fra NAV kan du kontakte din saksbehandler via den digitale aktivetsplanen.</p> <ul style="list-style-type: none"> <li>• Din situasjon</li> <li>• Digital aktivitetsplan</li> <li>• Hvor kan jeg få hjelp?</li> </ul>	

I dette eksempelet vil brukeren antagelig bli satt direkte over til veileder og skriver derfor 'veileder nav'. Men Frida kobler dette opp mot intensjonen 'Timeavtale' og svarer at man ikke kan få timeavtale, uten at brukeren blir satt over.

KAPITTEL 6

ID: 678937 (???.09.2019)		
Hvem	Tekst	Intensjon
Bruker	jeg tror ikke det jeg kan skrive kortere	
Frida	<p>Beklager, men jeg kan dessverre ikke svare på det du lurer på.</p> <p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapspengen</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringspengen</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Ja, sett meg over</li> <li>• Nei, ikke sett meg over</li> </ul>	
Bruker	Ja, takk, gjerne	X Generelle spørsmål / Takk
Frida	Bare hyggelig!	
Bruker	Er det lang ventetid?	Unknown
Frida	<p>Beklager, men jeg kan dessverre ikke svare på det du lurer på.</p> <p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapspengen</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> </ul>	

## VELLYKKEDE SAMTALER

	<ul style="list-style-type: none"><li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li><li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li><li>• Omsorgs- og opplæringspenger</li></ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"><li>• Ja, sett meg over</li><li>• Nei, ikke sett meg over</li></ul>	
Bruker	ja	
Frida	Jeg setter deg over, et øyeblikk.	

I dette eksempelet spør Frida om bruker vil bli satt over og bruker svarer 'Ja, takk, gjerne'.

Fordi brukeren skriver flere enn to ord blir dette prosessert som et spørsmål og blir koblet opp mot 'Takk' under generelle spørsmål. Når Frida svarer 'bare hyggelig' får bruker inntrykket av å bli satt over. Bruker spør så om det er lang ventetid og Frida tilbyr igjen å sette over til veileder. Denne gangen skriver bruker bare 'ja' og blir dermed satt over.

Slike ting, som at bruk av for mange ord gjør at det blir prosessert som spørsmål og at man ved å skrive et svar i stedet for å trykke på knappene kan risikere å ikke bli satt over, er det mange brukere som tilsynelatende ikke er klar over. I mange tilfeller vil bruken av knapper være et mer sikkert alternativ for at forespørselen skal bli prosessert riktig av Frida. Det kan tenkes at slike tilfeller kunne vært løst om bruker ikke hadde mulighet til å skrive et svar når Frida tilbyr knapper, men dette vil igjen vært problematisk ettersom at det hender at Frida sine knapper i passer med brukernes forespørsel. Allikevel, akkurat i dette tilfellet, hvor knappene Frida tilbyr er 'Ja, sett meg over' og 'Nei, ikke sett meg over', kunne en slik løsning med bare knapper muligens vært et alternativ.

## 6.4 Oppsummering

I dette kapitlet har jeg forklart hva som skjer når samtaler blir satt over til veileder og i hvilke tilfeller brukere har problemer med å bli satt over. Slike utfordringer vil jeg komme

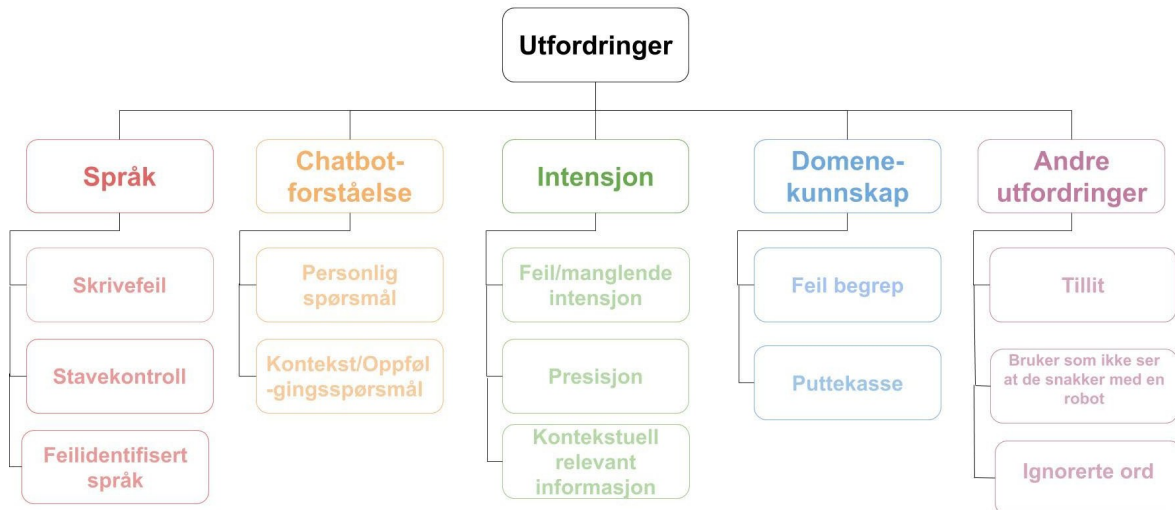
## KAPITTEL 6

nærmere inn på i de følgende kapitlene hvor jeg tar for meg de forskjellige kategoriene av utfordringer som har vært fremtredende i chattene jeg har analysert.



# 7 utfordringer

Chatter som ikke er vellykket har vi valgt å dele inn i fem underkategorier, ut ifra hvor utfordringene i chat-dialogen ligger; Intensjon, Språk, Chatbotforståelse, Domenekunnskap og andre utfordringer.



- Språk
  - Skrivefeil
  - Feilidentifisert språk
  - Stavekontroll
- Chatbot-forståelse
  - Personlige spørsmål
  - Kontekst / oppfølgingsspørsmål
- Intensjon
  - Feil / manglende intensjon
  - Presisjon
  - Kontekstuell relevant informasjon
- Domenekunnskap
  - Feil begrep
- Puttekassen
- Andre utfordringer
  - Tillit
    - Brukere som ikke ser at de snakker med en chatbot
    - Ignorerte ord

De forskjellige kategoriene av utfordringer (Språk, Chatbotforståelse, Domenekunnskap, Intensjoner og Andre utfordringer) vil utdypes i hvert sitt kapittel.



## 8 Språk

Flere av utfordringene vi har sett har vært knyttet til språkproblemer. Noen ganger oppstår problemene fordi brukeren har så mye skrivefeil at Frida ikke klarer å koble det opp mot noe nøkkelord, men det kan også være feilaktige rettinger av ord brukeren har skrevet riktig.

### 8.1 Utfordringer knyttet til språk

Kategorien språk omfatter alt som har å gjøre med *skrivefeil*, *stavekontroll* og *feilidentifisert språk*. Dette kan være at brukeren skriver feil slik at Frida misforstår hva det blir spurt om eller at Frida sin stavekontroll automatisk retter ord til noe annet enn hva brukeren mente. I noen tilfeller ser vi også at Frida identifiserer feil språk ut ifra det brukeren skriver, som igjen fører til følgefeil hvor Frida ikke forstår hva som skrives fordi hun analyserer det på bakgrunn av feil språk.

#### 8.1.1 Skrivefeil

Skrivefeil vil i noen tilfeller bli rettet opp til riktig uttrykk av stavekontroll slik at Frida kan koble det opp mot riktig intensjon, men i andre tilfeller vil dette føre til problemer. Her følger et eksempel på dette:

ID: 125269 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurer du på?  <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	hei	X Generelle spørsmål/ Hei
Frida	Hei, hva kan jeg hjelpe deg med?	
Bruker	Jeg er uført 100 prsent men kan jobb 40 prosent	Foreldrepenger/ Foreldrepenger
Frida	Foreldrepengene skal sikre foreldene inntekt i forbindelse med fødsel og adopsjon.	

## KAPITTEL 8

	<p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• Har jeg rett (knapp)</li> <li>• Periodens lengde (knapp)</li> <li>• Fleksibelt uttak(knapp)</li> <li>• Alt om søknaden(knapp)</li> <li>• Les mer (knapp)</li> <li>• Status i sak (knapp)</li> </ul>	
Bruker	Har jeg rett (handling)	
	<p>Opptjeningsreglene er de samme, uavhengig av hvilken yrkesstatus (arbeidstaker, selvstendig næringsdrivende, frilanser) du har.</p> <p>Ytelser som overgangsstønad, kontantstøtte og uføretrygd gir ikke opptjening til foreldrepenger.</p> <p>Er du eller skal du bli:</p> <ul style="list-style-type: none"> <li>• Mor(knapp)</li> <li>• Far(knapp)</li> <li>• Medmor(knapp)</li> <li>• Fosterforeldre(knapp)</li> </ul>	
Bruker	Er jeg	X Generelle spørsmål/ Jeg forstår
Frida	Er det noe annet jeg kan hjelpe deg med?	
Bruker	Ni takk	X Generelle spørsmål/Takk
Frida	Bare hyggelig!	

Denne chatten viser hvordan skrivefeil fører til problemer. Her skriver brukeren at vedkommende er ufør, men på grunn av feilstaving klarer ikke Frida å oppfatte dette nøkkelordet. Istedenfor kobler Frida det til intensjonen Foreldrepenger. Dette er koblet til sannsynlighetsberegningen Frida gjør. De spørsmålene Frida får oftest er om foreldrepenger, og 'Foreldrepenger' er derfor den vanligste intensjonen. Dette gjør at Frida ofte også tolker spørsmål som handler om noe helt annet til intensjonen 'Foreldrepenger' siden dette er det mest sannsynlige spørsmålet, basert på statistikk.

## SPRÅK

TOP 10 ROOT INTENTS			
#	Intent name	Count	Percentage
1	Foreldrepenger	2 457	33.75%
2	X Generelle spørsmål	1 558	21.40%
3	X Om Nav	652	8.95%
4	Unknown	586	8.05%
5	Generelle spørsmål	428	5.87%
6	Barnetrygd	408	5.60%
7	Bidrag	362	4.97%
8	Kontantstøtte	247	3.39%
9	Enslig forsørger	144	1.97%
10	Engangsstønad	114	1.56%

Denne oversikten viser de vanligste intensjonene ut ifra en gitt periode. Prosenten kan forandre seg etter hva Frida får flest spørsmål om, men intensjonen 'Foreldrepenger' vil som regel være på topp.

'Er jeg' tolkes som intensjonen 'Jeg forstår'. Dette kommer antagelig av at nøkkelordet 'jeg' kobles til frasen 'Jeg forstår' som er en gren på intensjonstreet for Generelle spørsmål.

Her kan det tenkes at brukeren hadde tenkt å stille et nytt spørsmål, men gir opp når Frida skriver 'Er det noe annet jeg kan hjelpe deg med'. Brukeren svarer da 'ni takk', som antagelig er en feilstaving av 'nei takk'. Frida oppfatter da bare ordet 'tak', og kobler det til intensjonen 'Takk', noe som resulterer i at hun svarer 'bare hyggelig'. I denne chatten burde nok Frida ha satt over til menneskelig veileder i stedet.

### 8.1.2 Feilidentifisert språk

Noen ganger tolker Frida brukerens utsagn som et annet språk. Dette skjer gjerne om brukeren skriver noe feil, slik at det ligner et dansk eller svensk ord. Frida vil da prøve å tolke det neste brukeren skriver ut ifra dette språket, noe som ofte leder til 'unknown'.

ID: N/A (???.???.2019)

## KAPITTEL 8

Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Far betaler ikke barnebidrag i tide	?Barnebidrag
Frida	<p>Jeg kan svare deg på generelle spørsmål om barnebidrag.</p> <p>Du kan skrive et spørsmål selv eller velge ut fra listen nedenfor.</p> <ul style="list-style-type: none"> <li>• Privat avtale (knapp)</li> <li>• Fastsetting av bidrag (knapp)</li> <li>• Over 18 år (knapp)</li> <li>• Søke bidrag (knapp)</li> <li>• Utbetaling (knapp)</li> </ul>	
Bruker	forsen betaling	Generelle spørsmål / Utbetaling  Språk: dansk
Frida	<p>På Dine utbetalinger finner du en oversikt over alle utbetalinger fra NAV de siste 3 årene.</p> <p>Logg inn på Dine utbetalinger, eller les mer på lenkene</p> <p>Hvis du ikke finner det du ser etter på Dine utbetalinger, må du ringe 55 55 33 33.</p> <ul style="list-style-type: none"> <li>• Mangler utbetaling? (knapp)</li> <li>• Stemmer ikke beløpet du har fått? (knapp)</li> <li>• Er du arbeidsgiver? (knapp)</li> <li>• Utbetalingsdato (knapp)</li> <li>• Dine utbetalinger (knapp)</li> </ul>	
Bruker	barnebidrag	Unknown

## SPRÅK

Frida	Beklager, men jeg forstår ikke hva du spør om. Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder. Kan du prøve å stille spørsmålet på en annen måte?  <ul style="list-style-type: none"> <li>• Trenger mer hjelp (knapp)</li> <li>• Nei (knapp)</li> </ul>	
Bruker	far betaler ikke	Utbetalingsdato
Frida	Hvilken dag utbetalingen kommer, avhenger av hvilken ytelse det gjelder. Frida: Du kan se utbetalingene fra NAV på «Dine utbetalinger» på nav.no, men de vises ikke før de er kommet inn på konto.	

I denne chatten er det utsagnet ‘forsen betaling’ som Frida tolker som dansk. Ordet barnebidrag blir dermed tolket som et dansk utsagn og siden Frida ikke klarer å tolke det opp mot et dansk ord så blir dette i stedet ‘unknown’. Her er det altså Frida sin språk-tolkning som skaper problemer. Selv om brukeren skriver neste ord ‘barnebidrag’ riktig, noe som vanligvis ville fulgt direkte til intensjonen Barnebidrag, så klarer ikke Frida å koble det opp mot et nøkkelord. Brukeren blir heller ikke gjort oppmerksom på at Frida nå er i et dansk ‘spor’ og at det er derfor hun ikke kan forstå hva brukeren skriver. Dette gjør det vanskelig for bruker å forstå hvorfor det oppstår feil og hva man skal skrive for at Frida skal forstå.

### 8.1.3 Stavekontroll

Stavekontrollen kan i noen tilfeller endre brukerens utsagn til noe annet enn det brukeren prøvde å formidle. I enkelte tilfeller kan også ord som er stavet riktig bli rettet av stavekontroll til andre ord.

ID: N/A (???.???.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurere du på?  <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	

## KAPITTEL 8

Bruker	Må far velge 80% utbetaling i sin pappaperm, dersom mor velger dette?	??Foreldrepenger/Overta fedrekvoten
Frida	<p>Mor kan overta fedrekvoten hvis ett eller flere av punktene under er oppfylt:</p> <ul style="list-style-type: none"> <li>• Hun er alene om omsorgen for barnet</li> <li>• Far har ikke rett til foreldrepenger</li> <li>• Far er innlagt i helseinstitusjon</li> <li>• Far har en sykdom eller skade som gjør at han ikke er i stand til å ta seg av barnet alene</li> <li>• Far er syk (knapp)</li> </ul>	
Bruker	Må begge foresatte ha samme prosentandel utbetalt ved permisjon?	Unknown
Frida	Beklager, men jeg forstod ikke helt hva du spurte om. Kan du prøve å spørre meg igjen på en litt annen måte? De er best om du skriver korte og enkle spørsmål.	

Her spør brukeren først om far må velge 80 % utbetaling i sin pappaperm, dersom mor velger dette. Frida tolker det som om spørsmålet dreier seg om overtakelse av fedrekvoten, og svarer bruker med at mor kan overta fedrekvoten dersom visse krav er oppfylt. Bruker omformulerer seg, og spør så om begge foresatte må ha samme prosentandel utbetalt ved permisjon. Frida svarer at hun ikke forstår spørsmålet. Grunnen til dette er at 'foresatte' i spørsmålet 'Må begge foresatt ha samme prosentandel utbetalt ved permisjon?' rettet til 'fortsatt'. Frida klarer dermed ikke å koble brukerens utsagn opp mot noen intensjon.

Det bruker skriver	Stavekontroll	Det Frida leser	Intensjon
Må begge <b>foresatte</b> ha samme prosentandel utbetalt ved permisjon?	Foresatt=fortsatt	Må begge <b>fortsatt</b> ha samme prosentandel utbetalt ved permisjon?	Frida finner ikke nøkkelordet foresatt, og ender opp med 'Unknown'

Illustrasjon av hvordan stavekontrollen hindrer Frida i å koble utsagnet opp mot riktig nøkkelord.



## 8.2 Språk som utfordring

I disse eksemplene har jeg vist hvordan språk fører til utfordringer grunnet Frida sitt forsøk på å rekke brukerfeil. I første eksempel fører brukerens skrivefeil til at Frida ikke forstår hva brukeren prøver å spørre om, men som vi kan se i de neste to eksemplene er det også feil som oppstår fordi Frida forsøker å rette det opp. Misforståelser knyttet til språk trenger derfor ikke å være fordi bruker skriver feil, men rett og slett fordi Frida ikke kjenner til ordet bruker skriver og derfor endrer til et ord hun kjenner til, selv om det ikke var dette bruker mente. Om bruker hadde skrevet feil i utgangspunktet er dette en nyttig funksjon ettersom at det gjør at Frida kan få koblet til riktige intensjoner selv om bruker har skrivefeil. Problemet oppstår når hun endrer på ting som egentlig er riktig, fordi hun ikke finner nøkkelord. Men dette er noe hun vil bli bedre på etter hvert som det foretas flere treningsrunder og legges til flere synonymer.

### Mangel på innsyn i maskineri

Når Frida forsøker å korrigere ord eller språk, får ikke brukerne beskjed om dette. De vet ikke hvor misforståelsen har oppstått og hva de må gjøre for at Frida skal forstå spørsmålet riktige. Om Frida i stedet hadde svart med «Dette forstår ordet forstår jeg ikke, kan du spørre på en annen måte?» kunne bruker fått en forklaring på hvorfor det ikke fungerer og deretter spurt på en annen måte. Når hun kobler et nøkkelord til en intensjon kunne også gjerne ha startet med «Jeg forstår det sånn/oppfatter at du spør om dette», slik at bruker skjønner hva som foregår og hva Frida svarer på.

Denne mangelen på informasjon om hva som skjer og hvorfor de ikke får svar på det de lurer på, fører ofte til svar fra brukerne som indikerer at de er irriterte, noe som kan gjøre at de gir opp å få besvart spørsmålene sine på denne måten. NKS har som mål å bidra med informasjon og bedre brukerkontakt, men misforståelse slik som dette, kan være med på å forme inntrykket brukerne har av NAV.

## 8.3 Oppsummering

I dette kapitlet har jeg gjort rede for utfordringer tilknyttet språk. De fleste feilene tilknyttet språk oppstår altså ikke kun på bakgrunn av at bruker skriver feil, men fordi Frida forsøker å rette det opp. Om Frida retter noe feil vil dette kunne oppleves forvirrende for bruker og Frida burde derfor tydeliggjøre hva som skjer.

## 9 Chatbotforståelse

Mange brukere har manglende forståelse for hvordan chatboter fungerer, dette kan føre til flere problemer. Samhandling med chatbot krever forståelse for chatbot-mekanikken, ettersom at spørsmål til en chatbot ofte må formuleres litt annerledes enn om du hadde snakket med menneske. Frida oppgir at brukeren må bruke nøkkelord og stille generelle spørsmål, men det er allikevel mange brukere som stiller kompliserte personlige spørsmål som er vanskelig for Frida å besvare.

### 9.1 utfordringer knyttet til chatbotforståelse

Chatbot-forståelse handler om brukerens forståelse av hvordan chatboter fungerer, inndelt i tre underkategorier: *personlig spørsmål og kontekst/oppfølgingsspørsmål*. Her plasserer vi samtaler hvor brukeren viser mangel på forståelse for hvordan chatboten fungerer ved enten å stille personlig spørsmål, selv om det presiseres at chatboten bare kan svare på generelle spørsmål, eller å formulere spørsmålene på en måte som gjør at chatboten ikke oppfatter hva det spørres om ved å bruke for mange nøkkelord. I noen tilfeller prøver også brukeren å stille oppfølgingsspørsmål, selv om Frida ikke forstår konteksten til hva brukeren spør om. Alle disse tingene er relatert til at brukerne har for lite kunnskap om hvordan chatboter fungerer og hvordan man skal kommunisere med den.

#### 9.1.1 Personlig spørsmål

Personlige spørsmål er spørsmål som gjelder brukerens spesifikke situasjon og er vanskelig å svare generelt på. Frida sine generelle svar passer derfor gjerne dårlig til å besvare brukernes spørsmål.

ID: 125330 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	

## KAPITTEL 9

Bruker	Hva betyr uke syv etter fødsel	Foreldrepenger / Fødsel utenfor termin
Frida	<p>Ved fødsel er 3 uker av foreldrepengeperioden er forbeholdt tiden før termin/planlagt keisersnitt. Disse ukene inngår i foreldrepengeperioden, men er ikke en del av hverken mødrekvoten eller fellesperioden.</p> <p>De 3 ukene kan ikke brukes etter fødsel.</p> <p>NAV får melding om fødsel fra Folkeregisteret som får fødselsmelding fra sykehuset.</p> <p>Du trenger derfor ikke gi beskjed om fødsel før eller etter termin dato.</p> <ul style="list-style-type: none"> <li>• Endret termin dato?</li> <li>• Fødsel før termin</li> <li>• Fødsel etter termin</li> </ul>	
Bruker	Om man har fått barn <dato>, hvilken dato er syv uker fra fødsel å regne?	Unknown
Frida	Jeg svarte kanskje ikke helt på det du lurte på? Kanskje du kan spørre litt annerledes?	
Bruker	Hva er syv uker etter <dato>	Unknown
Frida	<p>Beklager, men jeg forstår ikke hva du spør om.</p> <p>Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.</p> <p>Kan du stille spørsmålet på nytt?</p> <ul style="list-style-type: none"> <li>• Trenger mer hjelp</li> <li>• Nei</li> </ul>	
Bruker	Hva betyr syv uker etter fødsel	Foreldrepenger / Fødsel utenfor termin
Frida	<p>Ved fødsel er 3 uker av foreldrepengeperioden er forbeholdt tiden før termin/planlagt keisersnitt. Disse ukene inngår i foreldrepengeperioden, men er ikke en del av hverken mødrekvoten eller fellesperioden.</p> <p>De 3 ukene kan ikke brukes etter fødsel.</p> <p>NAV får melding om fødsel fra Folkeregisteret som får fødselsmelding fra sykehuset.</p>	

## CHATBOTFORSTÅELSE

	Du trenger derfor ikke gi beskjed om fødsel før eller etter termin dato. <ul style="list-style-type: none"><li>• Endret termin dato?</li><li>• Fødsel før termin</li><li>• Fødsel etter termin</li></ul>	
--	--	--

Dette eksempelet viser en mangel på forståelse for hvordan en chatbot fungerer fra brukerens side. Frida opplyser om at hun svarer på korte spørsmål, og at hun ikke forstår hva brukeren spør om. Brukeren forsetter allikevel å stille samme spørsmålet om igjen, på en litt annen måte. Frida forstår det fortsatt ikke og kommer kun med informasjon om fødsel utenfor termin. Brukeren får ikke svar på det vedkommende lurer på, og blir heller ikke satt over. Dette er et to-delt problem. Brukeren forstår ikke at dette er et spørsmål Frida ikke kan svare på, og Frida foreslår heller ikke å sette over. Dette kunne vært løst enten ved å kommunisere tydeligere til brukeren hva slags spørsmål Frida ikke kan svare på, eller å gjøre at Frida lettere setter over i slike tilfeller. I etterkant har dette blitt endret slik at to 'unknown'-intensjoner etter hverandre gjør at Frida foreslår å sette over.

### 9.1.2 Kontekst/oppfølgingsspørsmål

En bruker vil ofte stille oppfølgingsspørsmål innenfor et tema. Hvis Frida da er på riktig intensjonstre, så kan bruker få riktig svar på dette, men hvis Frida ikke lenger er i samme intensjonstre eller er i et annet intensjonstre så vil brukeren få feil svar på dette.

ID: 125329 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.  Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.  Hva lurer du på? <ul style="list-style-type: none"><li>• For english chat, click here</li></ul>	
Bruker	Hvordan overføre foreldrepermisjon fra mor til far når mor er sykemeldt	Foreldrepenger/ Utsette egen sykdom - mødrekvote
Frida	(...)  Vanlig sykemelding er ikke tilstrekkelig.	

## KAPITTEL 9

	<p>Dette må legen din ta med i funksjonsbeskrivelsen</p> <ul style="list-style-type: none"> <li>• hvordan sykdommen påvirker evnen din til å ta deg av barnet</li> <li>• i hvilken grad du er helt avhengig av hjelp</li> <li>• hvilken periode du er avhengig av hjelp</li> </ul> <p>Den andre foreldren kan som et alternativ til utsettelse, søke om å overta permisjonen din. Kravet er det samme som for utsettelse; du må være syk og helt avhengig av hjelp til å ta deg av barnet.</p> <ul style="list-style-type: none"> <li>• Søk her(Knapp)</li> </ul>	
Bruker	Den andre foreldren kan som et alternativ til utsettelse, søke om å overta permisjonen din.	Foreldrepenger/ Utsettelse søke
Frida	<p>Jeg anbefaler at du søker om utsettelse samtidig som du søker om foreldrepenger.</p> <p>Det er også mulig å bare søke om utsettelse.</p> <p>Hvis du skal utsette foreldrepengeperioden fordi du skal jobbe heltid eller ta ut lovbestemt ferie, må NAV ha mottatt søknaden din før du begynner å jobbe eller før ferien din starter.</p> <p>Dersom du skal utsette fedrekvoten eller kvoten til medmor, må søknaden din være mottatt hos NAV senest den siste dagen i fellesperioden. Da må du søke om fedrekvote/medmorkvote og utsettelse samtidig.</p> <ul style="list-style-type: none"> <li>• Nytt arbeidsforhold (knapp)</li> <li>• Rett til utsettelse? (knapp)</li> <li>• Endringer (knapp)</li> <li>• Om søknaden (knapp)</li> </ul>	
Bruker	hvordan gjør jeg dette	X Generelle spørsmål/ Hva gjør jeg
Frida	Hva trenger du hjelp til?	
Bruker	Endringer(handling)	
Frida	<p>Hvis du vil</p> <ul style="list-style-type: none"> <li>• endre en innsendt søknad</li> <li>• endre foreldrepengeperioden</li> </ul>	

## CHATBOTFORSTÅELSE

	<ul style="list-style-type: none"> <li>• utsette eller gradere</li> <li>• søke om gjenstående foreldrepenger må du sende ny søknad.</li> </ul> <p>Hva du skal søke om, avhenger av hvilken endring du vil gjøre.</p> <p>Er barnet født og du allerede har fått innvilget foreldrepenger, kan du bruke den digital søknaden og velge "Søk endring"</p> <ul style="list-style-type: none"> <li>• Hvordan fylle ut søknad</li> <li>• Status i sak</li> <li>• Trekke søknad</li> <li>• Gå til søknad</li> <li>• Mer om fleksibelt uttak</li> </ul>	
--	--	--

Frida tolker første utsagn riktig med intensjonen 'Foreldrepenger'. Begge de to første utsagnene til brukeren blir koblet til grener under intensjonstreet 'Foreldrepenger', henholdsvis 'Foreldrepenger/ Utsette egen sykdom – mødrekvote' og 'Foreldrepenger/ Utsettelse søke'. Men når brukeren spør 'hvordan gjør jeg dette?', blir dette i stedet koblet til Generelle spørsmål. Her er problemet at det er større sannsynlighet for at spørsmålet er 'Hva gjør jeg?', noe som er en gren på intensjonstreet for generelle spørsmål. Frida svarer derfor 'Hva trenger du hjelp til?'. Dette resulterer i at brukeren går tilbake i samtalen og forsøker å trykke på en knapp i stedet, som var tilknyttet foreldrepenger.

Problemet her er sammensatt. Det er en kombinasjon av at Frida ikke svarer ut ifra samme intensjonstre og at brukeren stiller et oppfølgingsspørsmål på samme måte som man ville gjort til et menneske uten å tenke på at en chatbot trenger kontekst for spørsmålet, ikke et spørsmål revet ut av sammenheng. Om du lurer på hvordan du skal søke om noe må du derfor skrive ordet søke og hva du skal søke om i samme setning.

Det hender at bruker spør urelaterte spørsmål og dette vil derfor i andre sammenhenger bli riktig når Frida svarer med et svar fra et annet intensjonstre. Dette er derfor vanskelig å få rettet opp, og løsningen ligger nok derfor heller i at bruker trenger mer kunnskap om hvordan chatbot fungerer for å stille spørsmålene på en annen måte.

## 9.2 Chatbotforståelse som utfordring

Mange brukere vet ikke hvordan de skal interagere med en chatbot. Det er derfor viktig at det gjøres klart hvordan de skal formulere spørsmål og hva de kan forvente at chatboten skal kunne svare på. Personlige spørsmål fører gjerne til 'unknown' eller feil intensjon, og oppfølgingsspørsmål blir gjerne knyttet til feil intensjon om ikke Frida er inne på riktig intensjonstre. Det oppgis at Frida bare kan svare på generelle spørsmål, men det er allikevel mange brukere som forsøker å stille personlige spørsmål. Det kan også være vanskelig å få formidlet til brukere hvordan Frida bruker intensjonsgrener til å svare på oppfølgingsspørsmål. I noen tilfeller kan oppfølgingsspørsmål bli riktige, i andre ikke. Dette kan lett gjøre at brukere gir opp, fordi de ikke skjønner hva som skjer. Å prøve å få formidlet hva Frida gjør, samtidig som man tar hensyn til dem med manglende chatbotforståelse er derfor viktig.

Større forståelse for hvordan chatboter fungerer, vil føre til at brukerne i mindre grad stiller personlige og lange spørsmål. Bruker med erfaring med bruk av chatbot vil derfor ha større sjanse for å få besvart sitt spørsmål. Det er derfor viktig at det gjøres tydelig hvordan man skal kommunisere med chatboten.

## 9.3 Oppsummering

I dette kapitlet har jeg vist at manglende forståelse for hvordan chatboter fungerer kan gjøre det vanskelig å kommunisere med Frida. Brukere som ikke stiller generelle spørsmål, men heller lange og personlige spørsmål, vil ofte bli koblet til 'unknown', slik at Frida ikke kan svare på spørsmålene. Dette skjer enten fordi bruker skriver for lange spørsmål eller stiller et personlig spørsmål som det ikke finnes noe generelt svar til. Oppfølgingsspørsmål fører også til problemer i noen tilfeller, selv om det kan fungere i andre tilfeller. Det er derfor viktig at Frida forklarer tydelig hva hun er i stand å svare på og hvordan man kommuniserer med henne.



## CHATBOTFORSTÅELSE

# 10 Domenekunnskap

Brukere som ikke har vært i kontakt med NAV tidligere har ofte varierende kunnskap om NAV sine tjenester og ytelser. Mange av de som kontakter Frida/NKS gjør det fordi de synes noe er forvirrende med NAV sitt begrepsapparat. Men når Frida baserer seg på at brukeren vet hvilke uttrykk og nøkkelord de skal bruke, så fører dette til problemer når brukeren ikke vet hvilket nøkkelord de skal bruke fordi de ikke kjenner til begrepene NAV bruker. Når de snakker med Frida klarer de da ikke formulere spørsmål sitt og får heller ikke svar ifra Frida. Dette kan ha store konsekvenser dersom brukerne tror de har fått riktig svar fordi de ikke skjønner at Frida har svart angående en annen ordning enn den de lurte på, noe som kan resultere i at de tror de har krav på penger de ikke har rett til, eller ikke blir opplyst om at de må søke for å få det de har krav på.

## 10.1 utfordringer knyttet til domenekunnskap

Kategorien domenekunnskap omfatter samtaler der bruker uttrykker manglende domenekunnskap. Dette innebærer at brukeren enten benytter seg av feil begrep, eller har problemer med å se hvordan eget liv passer inn i NAV. Vi har valgt å kalle disse to underkategoriene for *feil begrep* og *puttekassen*. En samtale vil havne i feil begrep-kategorien dersom en bruker har benyttet seg av feil begrep for å stille et spørsmål. Dette kan for eksempel være at ordet barnetrygd har blitt brukt om foreldrepenger etc. En samtale vil havne i kategorien puttekassen dersom en bruker har problemer med å forstå hvordan egen livssituasjon kan tilpasses NAV sine begrep, stønadsordninger og liknende.

### 10.1.1 Feil begrep

Mange brukere er ikke kjent med NAV sine begreper. De bruker da gjerne NAV sine begreper feil eller bruker andre begreper som ikke er en del av begrepsapparatet til Frida. Om de bruker andre uttrykk som ikke brukes i NAV vil Frida ofte ikke forstå spørsmålet og ende opp med intensjonen 'unknown'. Om brukeren bruker begrepet feil vil Frida svare ut ifra det gale begrepet. Brukeren blir da enten irritert fordi Frida ikke forstår spørsmålet, eller godtar svaret uten å se at det er feil.

DOMENEKUNNSKAP

ID: 125314 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	trygd innbetaling barn	Bidrag / Utbetaling bidrag
Frida	<p>Dere kan ordne betaling av bidraget privat, eller søke om at NAV Innkreving håndterer innkreving og utbetaling.</p> <p>Dette gjelder både når dere har en privat avtale og når bidraget er fastsatt av NAV.</p> <p>Selve innkrevingen er gebyrfri.</p> <ul style="list-style-type: none"> <li>• Søk om innkreving</li> <li>• Betalingsfrist</li> <li>• Fra innbetaling til utbetaling</li> <li>• Spørsmål om betaling og utbetaling</li> <li>• Meld fra om endringer</li> </ul>	
Bruker	stønad til barn under 18 år skal utbetales til barnets kontonummer	Bidrag / Over 18 bidrag
Frida	<p>Plikten til å betale barnebidrag varer normalt ut den måneden barnet fyller 18 år.</p> <p>Barn som går på videregående skole eller tilsvarende, kan ha rett til bidrag etter fylte 18 år.</p> <ul style="list-style-type: none"> <li>• Hvem søker? (knapp)</li> <li>• Privat avtale (knapp)</li> <li>• Beregne bidrag (knapp)</li> <li>• Fastsetting (knapp)</li> <li>• Vilkår (knapp)</li> <li>• Utbetaling (knapp)</li> </ul>	
Bruker	Utbetaling (handling)	
Frida	<p>Det er barnet som er bidragsmottaker og som mottar bidraget.</p> <ul style="list-style-type: none"> <li>• Utbetaling (knapp)</li> </ul>	
Bruker	Utbetaling/Betaling (handling)	

## KAPITTEL 10

Frida	<p>Dere kan ordne betaling av bidraget privat, eller søke om at NAV Innkreving håndterer innkreving og utbetaling.</p> <p>Dette gjelder både når dere har en privat avtale og når bidraget er fastsatt av NAV.</p> <p>Selve innkrevingen er gebyrfri.</p> <ul style="list-style-type: none"> <li>• Søk om innkreving (knapp)</li> <li>• Betalingsfrist (knapp)</li> <li>• Fra innbetaling til utbetaling (knapp)</li> <li>• Spørsmål om betaling og utbetaling (knapp)</li> <li>• Meld fra om endringer (knapp)</li> </ul>	
Bruker	Meld fra om endringer	
Frida	<p>Når NAV krever inn og utbetaler barnebidraget må NAV få beskjed ved endringer hvis:</p> <ul style="list-style-type: none"> <li>• dere blir enige om et nytt beløp i den private avtalen</li> <li>• dere har bidrag fastsatt av NAV og går over til en privat avtale</li> <li>• bidraget skal opphøre (stanses)</li> <li>• det er endringer som kan ha betydning for retten til bidragsforskudd</li> <li>• Send beskjed til NAV (knapp)</li> </ul>	

I dette eksempelet kjenner ikke bruker til de begrepene som brukes hos NAV og klarer ikke å gjøre seg forstått i samtale med Frida. Bruker spør om 'trygd innbetaling barn', men får ikke svar på det vedkommende lurer på. Frida svarer i stedet med informasjon om utbetaling av bidrag. Brukeren prøver å formulere spørsmålet på nytt, men bruker fortsatt feil begreper og samtalen går i en loop hvor Frida kommer med de samme svarene flere ganger før bruker til slutt avslutter samtalen. Her gjør altså brukeren manglende domenekunnskap at Frida svarer på noe helt annet, som gjør at brukeren ikke får svar på det vedkommende lurer på.

### 10.1.2 Puttekassen

Brukere som ikke klarer å plassere sin egen situasjon ut ifra NAV sine tjenester og ytelser og ikke vet hvordan de skal forklare egen situasjon ut ifra NAV sine begreper. Dette kan være

## DOMENEKUNNSKAP

brukere som ikke vet hva de har krav på, hvordan de skal forklare sitt tilfelle eller hvordan NAV sine begreper passer med deres situasjon.

ID: N/A (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Skal ha barn, hva må jeg søke om?	Unknown
Frida	<p>Beklager, men jeg forstår ikke hva du spør om.</p> <p>Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.</p> <p>Kan du stille spørsmålet på nytt?</p>	
Bruker	Hvordan får jeg penger når jeg venter barn?	Utbetaling/ Utbetalingsdatoer
Frida	<p>Hvilken dag utbetaling kommer, avhenger av hvilken ytelse det gjelder.</p> <p>Du kan se utbetalingene fra NAV på Dine utbetalinger på nav.no, men de vises ikke der før de er kommet på konto.</p> <ul style="list-style-type: none"> <li>• Oversikt utbetalingsdatoer (knapp)</li> <li>• Utbetalingsoversikt (knapp)</li> <li>• Ikke fått penger (knapp)</li> </ul>	

I dette eksempelet er det en bruker som ikke har kjennskap til NAV sine begreper og lurere på hva man må søke om når man skal ha barn. Brukeren viser mangel på domenekunnskap, og ettersom at brukeren ikke bruker de riktige begrepene, blir dette heller ikke koblet opp mot de riktige intensjonene. Frida spør først om brukeren kan omformulere spørsmålet før hun kobler det opp mot intensjonen 'utbetalingsdatoer', noe som gjør at brukeren ikke får svar på det vedkommende lurere på.

## KAPITTEL 10

ID: 125315 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	kan en ung ufør som er gift med utenlandsk enslig mor få barnetrygd og barnetillegg	Barnetrygd / Utvidet barnetrygd
Frida	<p>Er du enslig mor eller far, kan du ha rett til utvidet barnetrygd hvis du bor alene med barn.</p> <p>Den utvidede barnetrygden er barnetrygd for ett barn mer enn du faktisk bor sammen med.</p> <p>Du må søke om utvidet barnetrygd, den blir ikke automatisk innvilget.</p> <ul style="list-style-type: none"> <li>• Rett til utvidet barnetrygd?</li> <li>• Delt barnetrygd / bosted</li> <li>• Endringer i familiesituasjonen</li> <li>• Søknadsveileder</li> <li>• Særfradrag</li> <li>• Utbetaling</li> </ul>	
Bruker	Rett til utvidet barnetrygd? (handling)	
Frida	<p>Vilkår for utvidet barnetrygd - du bor alene med barn og:</p> <p>er separert, skilt eller gjenlevende ektefelle</p> <ul style="list-style-type: none"> <li>• er ugift og ikke lever sammen med den andre forelderen til barnet.</li> <li>• Samlivsbrudd mellom samboere som har felles barn under 16 år, må bekreftes med meklingsattest</li> <li>• har vært faktisk separert i minst 6 måneder uten separasjonsbevilling eller dom</li> <li>• har en ektefelle/samboer som er fengslet med ubetinget straff på minst 6 måneder, eller som har sittet i varetekt i minst 6 måneder</li> <li>• har en ektefelle/samboer som har vært forsvunnet i minst 6 måneder</li> </ul>	

DOMENEKUNNSKAP

	<p>Hvis du ikke har felles barn med den du har bodd sammen med, trenger vi bare dokumentasjon på at dere ikke bor samme. Det kan for eksempel være:</p> <p>Erklæring om samlivsbrudd</p> <ul style="list-style-type: none"> <li>• Seperasjonsbevilling</li> <li>• Når opphører retten til utvidet barnetrygd?</li> <li>• Søk utvidet barnetrygd</li> <li>• Delt barnetrygd</li> </ul>	
Bruker	Når opphører retten til utvidet barnetrygd?	Barnetrygd / Utvidet barnetrygd
Frida	<p>Utvidet barnetrygd faller bort hvis:</p> <ul style="list-style-type: none"> <li>• Du inngår ekteskap</li> <li>• Får felles barn med ny samboer</li> <li>• Har vært samboer med en du ikke har felles barn med, i 12 måneder</li> <li>• Flytter sammen med barnefar</li> </ul> <p>Er du enslig mor eller far, kan du ha rett til utvidet barnetrygd hvis du bor alene med barn.</p> <p>Den utvidede barnetrygden er barnetrygd for ett barn mer enn du faktisk bor sammen med.</p> <p>Du må søke om utvidet barnetrygd, den blir ikke automatisk innvilget.</p> <ul style="list-style-type: none"> <li>• Rett til utvidet barnetrygd?</li> <li>• Delt barnetrygd / bosted</li> <li>• Endringer i familiesituasjonen</li> <li>• Søknadsveileder</li> <li>• Særfradrag</li> <li>• Utbetaling</li> </ul>	
Bruker	Utbetaling (handling)	

## KAPITTEL 10

	<p>Barnetrygden utbetales siste ukedag hver måned. Barnetrygden er skattefri.</p> <p>Husk at du kan se utbetalinger fra NAV ved å logge på Ditt NAV.</p> <p>På nav.no finner du oversikt over utbetalingsdatoer for barnetrygd.</p> <ul style="list-style-type: none"> <li>• Hva kan jeg få?</li> <li>• Mangler utbetaling?</li> <li>• Må jeg søke?</li> <li>• Endringer i familiesituasjon</li> <li>• Når et barn fyller 18 år</li> <li>• Jeg mener barnebidrag</li> </ul>	
Bruker	kan en utenlandsk enslig mor få barnetillegg	Foreldrepenger / Dersom bare mor har rett
Frida	<p>I noen tilfeller kan mor få hele perioden med foreldrepenger.</p> <p>Hun kan enten være alene om omsorgen, eller bo sammen med barnets andre forelder som ikke har tjent opp rett til foreldrepenger.</p> <ul style="list-style-type: none"> <li>• Alene om omsorgen</li> <li>• Bor med far/medmor</li> <li>• Jeg mener barnetrygd</li> <li>• Jeg mener overgangsstonad</li> </ul>	
Bruker	Jeg mener barnetrygd (handling)	



Frida	<p>Er du enslig mor eller far, kan du ha rett til utvidet barnetrygd hvis du bor alene med barn.</p> <p>Den utvidede barnetrygden er barnetrygd for ett barn mer enn du faktisk bor sammen med.</p> <p>Du må søke om utvidet barnetrygd, den blir ikke automatisk innvilget.</p> <p>Rett til utvidet barnetrygd?</p> <ul style="list-style-type: none"> <li>• Delt barnetrygd / bosted</li> <li>• Endringer i familiesituasjonen</li> <li>• Søknadsveileder</li> <li>• Særfradrag</li> <li>• Utbetaling</li> </ul>	
-------	--	--

Bruker stiller i dette eksempelet spørsmålet “kan en ung ufør som er gift med utenlandsk enslig mor få barnetrygd og barnetillegg”. Bruker får svar fra Frida som omhandler støtte til enslig mor eller far. Bruker stiller så et spørsmål om en enslig utenlandsk mor kan få barnetillegg. Frida gir da informasjon om mors rettigheter som enslig, noe som ikke stemmer med brukers situasjon ettersom at de er gift så mor er ikke enslig. Her er det tydelig at det er en person som har problemer med å forklare situasjonen sin, slik at Frida kan gi riktige svar. Bruker får ikke svar på sitt spørsmål, og blir heller ikke henvist.

## 10.2 Domenekunnskap som utfordring

Mange brukere har manglende kunnskap om NAV sine ytelser og støtteordninger. Veiledere tilpasser seg gjerne med et mer muntlig språk, dette gjør ikke Frida. Der hvor en veileder har fordel av at de kan tilpasse seg brukeren de snakker med, så vil Frida ha forhåndsbestemte fraser som skal passe alle. Om en bruker skriver feil begrep vil Frida ofte ikke forstå hva brukeren spør om, og om Frida bruker begreper bruker ikke kjenner til kan det gjøre at bruker blir usikker på om de har fått svar på det de lurer på. Det er derfor viktig at språket Frida bruker stemmer overens med det språket brukeren benytter. Frida burde derfor tilpasses de uttrykkene som går igjen hos brukerne. Hun gjør dette allerede til en viss grad ved at hun skjønner at bruker som spør om støtteordninger når de skal ha barn mener foreldrepenger,

men det er fortsatt utrykk brukerne benytter som fører til misforståelser. Guri Verne sin studie av Skatteetaten(2015) viser at veilederne kom med tilleggsinformasjon om de skjønnte at dette var brukere med manglende domenekunnskap. Men dette kan være vanskelig for en chatbot å plukke opp. De mest alvorlige tilfellene er der hvor bruker ikke skjønner at svarene fra Frida ikke er relevante for dem og ikke får klarhet i dette i løpet av samtalen. Dette kan føre til at brukerne tror de har rett på noe de ikke har rett på, eller tror at de ikke trenger å søke for å få det de har krav på.

### **10.3 Oppsummering**

I dette kapitlet har jeg vist at brukere med manglende domenekunnskap har problemer med å få formulert spørsmålene sine slik at Frida forstår dem. De bruker begreper som ikke benyttes av NAV og klarer ikke å plassere situasjonen sin ut ifra NAV sine begreper. I slike situasjoner får ofte brukerne svar ifra Frida som ikke er relevante for dem og det er ikke alltid brukeren oppfatter at svaret ikke gjelder for deres situasjon.

## DOMENEKUNNSKAP

# 11 Intensjoner

Fridas intensjoner passer ikke alltid med brukerens spørsmål og noen ganger vil Frida svare med intensjoner som er irrelevante for det brukeren spør om. Dette kapittelet vil ta for seg utfordringene knyttet til dette.

## 11.1 Utfordringer knyttet til intensjoner

Kategorien Intensjon omfatter samtaler der Frida kobler spørsmålet til brukeren opp mot en annen intensjon enn det brukeren spør om. Dette kan da være knyttet til *feil/manglende intensjon*, *presisjon* eller *kontekstuell relevant informasjon*. Feil eller manglende intensjon vil si at Frida enten kobler brukerens spørsmål opp mot en annen intensjon enn hva brukeren spør om eller at en intensjon for det brukeren spør om ikke finnes, selv om spørsmålet skal være innenfor Fridas omfang. Presisjon handler om at Frida gir et veldig generelt svar som ikke nødvendigvis er relevant på et presist spørsmål. Mangel på kontekstuell relevant informasjon kan være at brukeren spør om å få snakke med veileder og blir sendt videre, selv om det er utenfor åpningstiden eller får beskjed om at man kan bruke e-post i tilfeller hvor det ikke mulig.

### 11.1.1 Feil/manglende intensjon

Noen spørsmål er det enten ikke meningen at Frida skal kunne svare på, eventuelt ikke lagt til intensjon for enda. Disse spørsmålene vil det da ofte mangle en passende intensjon for. Helst skulle Frida da ha enten henvist til riktig sted eller opplyst om at hun ikke kan svare på spørsmålet. I stedet hender det at hun kobler det opp mot en irrelevant intensjon. I andre tilfeller kan riktig intensjon eksistere, men sannsynlighetsberegningen til Frida knyttet utsagnet til brukeren opp mot feil intensjon.

ID: 125324 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	Jeg heter Frida og er en chatbot.	

## INTENSJONER

	<p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, <a href="#">click here</a></li> </ul>	
Bruker	Kan man studere mens man er i mammapermisjon	Foreldrepenger/ Fleksibelt uttak
Frida	<p>Du kan dele opp perioden med foreldrepenger, eller kombinere foreldrepenger og arbeid.</p> <p>Skal du jobbe på fulltid, ta ut lovbestemt ferie, er syk eller innlagt på sykehus?</p> <p>Da kan du utsette foreldrepengene.</p> <p>Jobber du mindre enn 100%, kan du kombinere jobb og foreldrepenger.</p> <ul style="list-style-type: none"> <li>• Jobbe deltid</li> <li>• Utsettelse</li> <li>• Må jeg søke?</li> <li>• Når må foreldrepengene være brukt?</li> <li>• Foreldrepengeplanlegger</li> <li>• Samtidig uttak</li> </ul>	
Bruker	Studere foreldrepermisjon	Foreldrepenger/ foreldrepengeperioden
Frida	<p>Foreldrepengeperioden ved fødsel er 49 uker med 100 prosent foreldrepenger og 59 uker med 80 prosent foreldrepenger.</p> <p>For adopsjon er den henholdsvis 46 (100%) og 56 uker (80%).</p> <p>Fordelingen av foreldrepengeperioden er sånn:</p> <ul style="list-style-type: none"> <li>• 3 uker forbeholdt mor og tiden før termin</li> <li>• Mødrekvote</li> <li>• Fedrekvote eller kvote til medmor</li> <li>• Fellesperiode</li> </ul> <p>Bruk foreldrepengeplanleggeren, da får du god oversikt over alle periodene.</p> <ul style="list-style-type: none"> <li>• Mødrekvote</li> <li>• Fedrekvote</li> <li>• Kvote til medmor</li> </ul>	

## KAPITTEL 11

	<ul style="list-style-type: none"> <li>• Fellesperiode</li> <li>• Foreldrepenge-planlegger</li> </ul>	
--	---	--

I dette eksempelet passer ikke intensjonen med det bruker spør om. Bruker lurer på om man kan studere mens man er i mammapermisjon. Frida svarer med informasjon om hvordan du kan kombinere foreldrepenger og arbeid, og om lengden på foreldrepengeperioden, i stedet for å svare på hvordan det fungerer når man er student. Brukeren prøver også å formulere seg med bare nøkkelord, men fortsatt får ikke brukeren svar på det vedkommende lurer på av Frida. Her kan det være at en intensjon om hvordan man kombinerer studier og mammapermisjon mangler eller at Frida svarer med feil intensjon.

ID: 124900 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hei	X Generelle spørsmål / Hei
Frida	<p>Hei!</p> <p>Hva kan jeg hjelpe deg med?</p>	
Bruker	Hei	Unknown
Frida	<p>Jeg svarte kanskje ikke helt på det du lurte på?</p> <p>Kanskje du kan spørre litt annerledes?</p>	
Bruker	Omsorg	X Generelle spørsmål / Ha det bra
Frida	<p>Ha det bra!</p> <p>Takk for praten!</p>	

I denne chatten prøver bruker å skrive ordet ‘omsorg’ for å få informasjon fra Frida. Brukeren forsøker tilsynelatende å benytte nøkkelord, noe som Frida vanligvis vil koble opp mot en intensjon. Men ordet ‘omsorg’ er antagelig ikke lagret som noe synonym for noen ytelser og Frida kobler det derfor til ‘ha det bra’ i stedet.

## 11.1.2 Presisjon

Denne kategorien omfatter spørsmål hvor bruker stiller et veldig spesifikt spørsmål og får et generelt svar. Ofte uttrykker bruker at de er misfornøyde med svaret og at Frida sin respons ikke besvarer spørsmålet deres, selv om svaret gjerne omhandler temaet brukerne lurer på.

ID: 125209 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	jeg lurer på om starter fødselspermisjonen før termin, blir de dagene/ukene redusert fra det totale 49 ukene?	Foreldrepenger/ Oppstart
Frida	<p>Foreldrepengene ved fødsel starter senest 3 uker før fastsatt termin.</p> <p>Du kan velge å starte med foreldrepenger inntil 12 uker før termin dato. Husk at det du bruker utover 3 uker tas fra fellesperioden. Perioden med foreldrepenger etter fødsel blir tilsvarende kortere</p> <p>Viktig! Starter du tidligere enn 3 uker før termin forskyver du også opptjeningsperioden tilsvarende. Opptjeningsperioden blir da de 10 siste månedene før du starter med foreldrepenger.</p> <p>Det kan få konsekvens for retten til foreldrepenger.</p> <ul style="list-style-type: none"> <li>• Far/medmor og oppstart</li> <li>• Fødsel utenfor termin</li> <li>• Opptjeningsperioden</li> <li>• Eksempler</li> </ul>	
Bruker	Far/medmor og oppstart(handling)	
Frida	<p>Som far/medmor er hovedregelen at du kan starte med foreldrepenger fra uke 7 etter fødsel.</p> <p>Er mor syk og helt avhengig av hjelp til å ta seg av barnet eller innlagt på helseinstitusjon, kan du</p>	

## KAPITTEL 11

	søke om å starte foreldrepengeperioden fra og med fødselsdatoen.	
Bruker	det er ikke det jeg lurte på	X Generelle spørsmål / X Generelle spørsmål

### *Utdrag fra chat*

Bruker lurer på om antall uker blir redusert hvis man tar ut fødselspermisjon før termin. Frida forklarer hvordan det fungerer med oppstart av foreldrepenger, men bruker oppfatter ikke dette som et svar på det spørsmålet. Her er svaret fra Frida generelt og bruker anser ikke dette som et svar på spørsmålet fordi det ikke er presist nok.

ID: 125192 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Hei, har spørsmål om søknad på barnetrygd	Barnetrygd/ Søknad barnetrygd
Frida	<p>Hvis barnet blir født i Norge, vil mor automatisk få innvilget barnetrygd omtrent 1-2 måneder etter fødsel.</p> <p>Hvis du har adoptert og/eller kan ha rett til utvidet barnetrygd må du søke om barnetrygden.</p> <p>Det er også andre situasjoner som krever at du søker, blant annet om barnet flytter og en annen skal motta barnetrygden, hvis du er verge eller fosterforelder. Det kan du lese mer om på nav.no</p> <ul style="list-style-type: none"> <li>• Utbetaling</li> <li>• Hvordan søke</li> <li>• Status i sak</li> <li>• Endringer i familie situasjon</li> <li>• Søk barnetrygd</li> </ul>	
Bruker	Hvordan søke(handling)	
Frida	Vi anbefaler at du søker digitalt (krever innlogging med BankID).	



## INTENSJONER

	<p>Du kan også bruke papirsøknaden som du sender i posten. Da trenger du en førsteside (strekkodeark) som du laster ned sammen med søknaden.</p> <p>På førstesiden finner du adressen du skal sende søknaden til.</p> <p>Du kan følge saken din på Ditt NAV som du finner på nav.no.</p> <ul style="list-style-type: none"> <li>• Søknad</li> <li>• Vedlegg</li> <li>• Søke tilbake i tid</li> </ul>	
Bruker	Må man søke både om både barnetrygd og utvida barnetrygd	Barnetrygd/ Utvidet barnetrygd
Frida	<p>Er du enslig mor eller far, kan du ha rett til utvidet barnetrygd hvis du bor alene med barn.</p> <p>Den utvidede barnetrygden er barnetrygd for ett barn mer enn du faktisk bor sammen med.</p> <p>Du må søke om utvidet barnetrygd, den blir ikke automatisk innvilget.</p> <ul style="list-style-type: none"> <li>• Rett til utvidet barnetrygd?</li> <li>• Delt barnetrygd / bosted</li> <li>• Endringer i familiesituasjonen</li> <li>• Søknadsveileder</li> <li>• Særfradrag</li> <li>• Utbetaling</li> </ul>	

I dette eksempelet får bruker tilsynelatende svar på sitt spørsmål, men det kan virke som om bruker ikke oppfatter dette. Bruker skriver først at vedkommende har et spørsmål om søknad om barnetrygd. Frida svarer som om bruker hadde et generelt spørsmål om barnetrygd-søknaden. I svaret fra Frida informeres det om at barnetrygd er noe som automatisk blir innvilget dersom barnet blir født i Norge, og at mor vil få første utbetaling omkring 2 måneder etter fødsel. Videre informeres det at dersom “du har adoptert og/eller kan ha rett til utvidet barnetrygd må du søke om barnetrygden”. I svaret fra Frida blir det også informert om at det finnes andre situasjoner der det er nødvendig å sende inn en søknad. I tillegg til dette presenterer chatboten fem ulike lenker, inkludert en lenke ved navn “Søk barnetrygd”. Bruker klikker på lenken ved navn “Hvordan søke”, og får informasjon om at det er mulig å søke barnetrygd både digitalt og via en papirsøknad som kan sendes i posten. Etter dette skriver

bruker inn et spørsmål. Bruker lurer på om man må søke om “både barnetrygd og utvida barnetrygd”. I svaret fra Frida informeres det om hvem som kan ha rett på utvidet barnetrygd, samt at dette er noe som det må søkes om. I svaret står det ingenting om at barnetrygd, som altså ikke er utvidet, blir innvilget automatisk. Det er uklart om bruker forstår den informasjonen som har blitt presentert.

### 11.1.3 Kontekstuell relevant informasjon

Noen ganger utelater Frida informasjon som ville vært relevant å oppgi, og kommer med alternativer som ikke er reelle i det gitte tidspunktet. Det kan for eksempel være å foreslå å sette over til veileder når det er utenfor tidsrommet man kan snakke med veileder, eller å foreslå e-post selv om man ikke kan bruke e-post til å kontakte NAV.

ID: 125318 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	kan summen av foreldrepengene endres?	Foreldrepenger/ Endre søknad foreldrepenger
Frida	<p>Hvis du vil</p> <ul style="list-style-type: none"> <li>• endre en innsendt søknad</li> <li>• endre foreldrepengeperioden</li> <li>• utsette eller gradere</li> <li>• søke om gjenstående foreldrepenger må du sende ny søknad.</li> </ul> <p>Hva du skal søke om, avhenger av hvilken endring du vil gjøre.</p> <p>Er barnet født og du allerede har fått innvilget foreldrepenger, kan du bruke den digital søknaden og velge "Søk endring"</p> <ul style="list-style-type: none"> <li>• Hvordan fylle ut søknad</li> <li>• Status i sak</li> <li>• Trekke søknad</li> <li>• Gå til søknad</li> </ul>	

## INTENSJONER

	<ul style="list-style-type: none"><li>• Mer om fleksibelt uttak</li></ul>	
Bruker	Hvordan fylle ut en søknad (handling)	
Frida	Den digitale søknaden er lett å bruke, prøv så skal du se!  Trenger du hjelp?  Jeg kan sette deg over til en av mine kollegaer sånn at du kan få hjelp.  Skal jeg det? <ul style="list-style-type: none"><li>• Ja</li><li>• Nei, ikke sett over</li></ul>	
Bruker	Ja (handling)	
Frida	Beklager, vi er stengt. Våre åpningstider er mandag til fredag, 08.00 15.00.  Sorry, we are closed. Our opening hours are Monday to Friday, 08.00 15.00.	

I dette eksempelet svarer bruker svarer ja til å snakke med veileder og får så beskjed om at det ikke er mulig å snakke med veileder ettersom at chatten er stengt. Her burde Frida ha opplyst om at chatten var stengt tidligere i stedet for å spørre om bruker vil bli satt over.

### 11.2 Intensjon som utfordring

Intensjonstrærne blir stadig videreutviklet av chatbot-trenerne, og særlig 'Unknown' som havner under 'feil/manglende intensjon', jobbes det mye med for å avdekke når Frida ikke klarer å svare på noe, og om det er fordi en intensjons for dette mangler eller fordi det ikke kobles til den riktige intensjonen.

'Kontekstuell relevant informasjon' er lettere å gjøre noe med, ettersom at Frida her kan stille avklaringspørsmål før hun kommer med informasjon. Hun kan la være å spørre om bruker vil bli satt over om det er utenfor tidsrommet, kun opplyse om når det er mulighet for det. Andre intensjonsgrener, slik som e-post, kan fjernes, fordi dette ikke lenger vil være aktuelt da dette er en kommunikasjonskanal NAV ikke lenger benytter. Denne kategorien

sammenfaller med retningslinjen fra Amershi m.fl.(2019) om at man alltid skal vise kontekstuell relevant informasjon.

Eksempelene som havner i 'Presisjon' vil være mer komplisert å tilpasse/forme svar til. Frida skal i utgangspunktet bare svare på generelle spørsmål og sette over til veileder hvis det er noe hun ikke kan svare på. Men ofte stiller brukere spørsmål som fint kan besvares med generelle regler, uten at de forstår at reglene gjelder for deres tilfelle. Veilederne kan skrive personlige svar som gjør at brukerne får svar som er tilpasset deres situasjon, det kan ikke Frida. Når veilederne ble spurt om spørsmålene brukerne kommer med, ofte kan finnes på nav.no, så uttalte en av dem dette:

*«Svaret er ja, det finnes ofte på NAV.no Men vi skjønner at de chatter med oss. For det er kanskje mye informasjon også. Også har de lest det, men ikke forstått det helt.» «Ofte vil de ha bekreftelse på ting, fordi det er et komplisert regelverk og mye nyanser»*

Det er derfor viktig at Frida kan få brukeren til å forstå at dette svaret gjelder deres situasjon. Noe av fordelen med en chatbot er at den kan sile ut informasjon for brukeren. På NAV sine nettsider hvor det er mye informasjon, så kan chatboten plukke ut akkurat den informasjonen(intensjonsgrenen) som er relevant. Det er allikevel ikke alltid brukeren skjønner at dette er relevant for dem. Å avklare brukerens situasjon med oppfølgingsspørsmål før Frida kommer med et svar er derfor viktig, slik Frida også gjør til en viss grad, sånn at brukeren skjønner at spørsmålet deres er forstått og at svaret er relevant for dem.

Ved å dele opp tekstblokker og stille oppfølgingsspørsmål vil svaret føles mer presist, og chatloggene viser at brukere ofte er mer fornøyde med svarene om svarene til Frida er kortere.

### 11.3 Oppsummering

I dette kapitlet har jeg vist at brukerutsagn noen gang kobles til feil intensjoner, ofte fordi bruker benytter nøkkelord som kobles opp mot ulike intensjoner. I noen tilfeller opplever også bruker svarene som ikke relevante, fordi de ikke er presise nok eller fordi de ikke passer med konteksten. Mer bruk av oppfølgingsspørsmål kunne muligens løst noe av problemet, men intensjonene er noe som stadig forbedres og utvides av chatbot-trenerne.

## INTENSJONER

# 12 Andre utfordringer

Det finnes også andre utfordringer som ikke passer inn i noen av kategoriene som er beskrevet i de foregående kapitlene. Disse utfordringene vil bli gjennomgått i dette kapitlet.

## 12.1 Tillit

Mange brukere stoler ikke på svarene chatboten gir. Noen stiller først spørsmålet de lurer på, før de avslutter samtalen når de skjønner at det ikke er et menneske som svarer, noe som tyder på mangel på tiltro til chatboten. Andre spør først om de kan få snakke med et menneske, før de stiller spørsmålet til chatboten når de skjønner at det er utenfor åpningstiden. Disse får ofte tilstrekkelig svar av chatboten, selv om de tilsynelatende ikke hadde tro på at en chatbot kunne svare på det i utgangspunktet. I disse tilfellene kan vi ikke vite sikkert om brukeren egentlig hadde tenkt å spørre om noe annet når de spurte etter veileder. Det som fremkommer fra chatten, tyder allikevel på at enkelte brukere er skeptiske til å stole på chatboten sine svar.

ID: 125213 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	<p>hei jeg skal søke foreldrepenger men husker ikke hvor mange uker jeg hadde rett på ?</p>	Foreldrepenger / Søknad foreldrepenger
Frida	<p>Jeg hjelper deg gjerne med foreldrepengeøknaden.</p> <p>Husk at det er den som skal ha foreldrepenger som søker selv.</p> <ul style="list-style-type: none"> <li>• Hva lurer du på?</li> <li>• Hvordan søker jeg?</li> <li>• Når skal jeg søke?</li> <li>• Bosted</li> <li>• Endringer</li> <li>• Status søknad/sak</li> </ul>	

## ANDRE UTFORDRINGER

	<ul style="list-style-type: none"> <li>• Vedlegg</li> </ul>	
Bruker	er dette en pc som svarer?	X Generelle spørsmål / Er du en robot
Frida	Ja, jeg er en robot.	
Bruker	oki hade	X Generelle spørsmål / Ha det bra
Frida	Ha det bra!  Takk for praten!	
Bruker	jeg skal spise is no	X Generelle spørsmål / X Generelle spørsmål
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> <li>• engangsstønad</li> <li>• svangerskapsenger</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringsenger</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Ja, sett meg over (knapp)</li> <li>• Nei, ikke sett meg over (knapp)</li> </ul>	

Bruker lurer på hvor mange uker med foreldrepenger vedkommende har rett på, og Frida svarer med informasjon om foreldrepenger og forskjellige lenker tilknyttet dette som bruker kan trykke på. I stedet for å klikke på noen av lenkene spør brukeren om det er en pc som svarer og avslutter samtalen når Frida svarer ja. Her ser det ut som at brukeren ikke har tiltro

## KAPITTEL 12

til at Frida kan gi riktige svar, ettersom at brukeren velger å avslutte samtalen når Frida sier at hun er en robot.

ID: 125246 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurere du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	personal	Bidrag / Privat bidrag
Frida	<p>Forsørging av barn er i utgangspunktet et privat forhold mellom foreldrene.</p> <p>Hvis dere sammen kan komme fram til en avtale om barnebidrag som begge er tilfreds med, trenger dere som hovedregel ikke å informere eller involvere NAV.</p> <p>Under "hvor mye" kan du lese mer om hvordan bidrag fastsettes. Trenger dere hjelp til å komme frem til et beløp, bruk gjerne bidragskalkulatoren.</p> <ul style="list-style-type: none"> <li>• Inngå privat avtale</li> <li>• Hvor mye?</li> <li>• Bidragskalkulator</li> </ul> <p>Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.</p> <p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	
Bruker	Chat(handling)	
Frida	<p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"> <li>• foreldrepenger</li> </ul>	



## ANDRE UTFORDRINGER

	<ul style="list-style-type: none"> <li>• engangsstønad</li> <li>• svangerskapspenger</li> <li>• barnetrygd</li> <li>• kontantstøtte</li> <li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li> <li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li> <li>• Omsorgs- og opplæringspenger</li> </ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Ja, sett meg over (knapp)</li> <li>• Nei, ikke sett meg over (knapp)</li> </ul>	
Bruker	Ja, sett meg over	
Frida	<p>Beklager, vi er stengt. Våre åpningstider er mandag til fredag, 08.00 15.00.</p> <p>Sorry, we are closed. Our opening hours are Monday to Friday, 08.00 15.00.</p>	
Bruker	foreldrepenger tvillinger	Foreldrepenger / Flere barn samtidig
Frida	<p>Hvis dere får flere barn ved fødsel eller adopterer flere barn samtidig, vil dere få flere uker med foreldrepenger.</p> <p>Antall uker bestemmes av om dere har valgt 80% eller 100% foreldrepenger, antall barn og når barna er født eller adoptert.</p> <p>Disse ekstraukene kan tas fra og med tidspunkt for fødsel eller omsorgsovertakelse. Foreldrene kan selv velge når disse ekstra ukene skal brukes og hvem som skal bruke de.</p> <p>Ukene kan også brukes til samtidig uttak sånn at mor og far/medmor har foreldrepenger samtidig.</p> <p>80 % foreldrepenger 100 % foreldrepenger Regler for bruk</p>	

I denne chatten tolker jeg det som at bruker først forsøker å skrive det norske ordet ‘personal’. Dette tolker Frida som ‘personlig’ og kobler det derfor til intensjonen ‘Bidrag / Privat Bidrag’. Frida informerer så om privat bidrag, men spør også om bruker vil bli satt over. Bruker trykker da på ‘chat’, men dette er utenfor åpningstidene og bruker blir derfor ikke satt over. Bruker forsøker da å spørre Frida ved å skrive ‘foreldrepenger tvillinger’ og Frida klarer å koble dette til riktig intensjon og besvarer spørsmålet. Jeg tolker dette som at bruker ikke trodde at Frida kunne svare på spørsmål og derfor ville bli satt over med en gang. Det kan også hende at det var noe annet brukeren tenkte å spørre veileder om, men jeg tolker det som at denne brukeren har manglende tillit til chatboter.

## 12.2 Brukere som ikke ser at de snakker med en chatbot

I mai var det mange brukere som ikke oppfattet at de snakket med en chatbot før langt uti samtalen. I etterkant har det blitt gjort endringer på blant annet Frida sin åpningsfrase som gjør at flere får med seg dette i starten av samtalen. Men fortsatt er det flere eksempler på brukere som forsøker å stille flere spørsmål før de til slutt spør «Er du en robot?». Når de ikke får med seg dette, forsøker de ikke å stille spørsmålene tilpasset chatboten og de blir skuffet over at de ikke får svar.

ID: 141736 (???.???.2019)		
Hvem	Tekst	Intensjon
Frida	Beklager, men jeg forstår ikke hva du spør om.  Jeg skjønner det best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.  Kan du stille spørsmålet på nytt? <ul style="list-style-type: none"> <li>• Trenger mer hjelp</li> <li>• Nei</li> </ul>	
Bruker	Er du en robot	X Generelle spørsmål / Er du en robot
Frida	Ja, jeg er en robot	
Bruker	Du fungerer dårlig	

## ANDRE UTFORDRINGER

Frida	<p>Beklager, men jeg kan dessverre ikke svare på det du lurer på.</p> <p>Her på chat kan vi svare deg på generelle spørsmål om:</p> <ul style="list-style-type: none"><li>• foreldrepenger</li><li>• engangsstønad</li><li>• svangerskapsenger</li><li>• barnetrygd</li><li>• kontantstøtte</li><li>• Overgangsstønad, barnetilsyn og tilleggsstønader til enslig forsørger</li><li>• Barnebidrag, bidragsforskudd og bidrag til særlige utgifter</li><li>• Omsorgs- og opplæringsenger</li></ul> <p>Har du spørsmål om andre ytelser eller spørsmål om søknad, sak, utbetaling eller andre spørsmål som krever personopplysninger for å besvares må du ringe oss for hjelp, 55 55 33 33.</p>	
-------	---	--

*Utdrag fra chat*

I denne chatten er det tydelig at brukeren ikke leste hele åpningsfrasen til Frida og ikke oppfattet at det var en chatbot brukeren snakket med.

### 12.3 Ignorerte ord/Stoppord

Et annet problem er ordene som blir ignorert. Frida ser bare på nøkkelordene, og selv om mange ord ofte er irrelevante for temaet/intensjonen, så er det også enkelte ord som blir ignorert som er viktige for at spørsmålet skal bli forstått.

Et problem som gikk igjen, var at brukere ble irritert over at Frida ikke oppfattet ordet 'ikke'.

ID: N/A (???.???.2019)		
Hvem	Tekst	Intensjon
Bruker	Jeg informerte om at det ikke skal søkes om foreldrepenger	Foreldrepenger/ Søke
Frida	Jeg hjelper deg gjerne med foreldrepengesøknaden.	

## KAPITTEL 12

	<p>Husk at det er den som skal ha foreldrepenger som søker selv.</p> <ul style="list-style-type: none"> <li>• Hva lurer du på?</li> <li>• Hvordan søker jeg?</li> <li>• Når skal jeg søke?</li> <li>• Bosted</li> <li>• Endringer</li> </ul>	
Bruker	Drit å dra	Fedrekvoten
	<p>Fedrekvoten er forbeholdt fra, men han trenger ikke ut hele. Han mister de dagene han ikke tar ut.</p> <p>Fedrekvoten kan også utsettes eller deles opp over en lengre periode.</p>	

### *Utdrag fra chat*

Her poengterer brukeren at det ikke skal søkes om foreldrepenger, men Frida ignorerer ordet 'ikke' og kommer med informasjon om hvordan man søker om foreldrepenger, noe som tilsynelatende gjør at brukeren blir irritert.

ID: 172477 (08.10.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <p>Får du ikke svar på det du spør om, kan det skyldes at jeg er under opplæring.</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Har sendt dokumentasjon, men tror at jeg samtidig har søkt om utvidet barnetrygd. Det skulle jeg ikke...	Foreldrepenger/ Søke
Frida	<p>Du kan ha rett til utvidet barnetrygd hvis du bor alene med barn.</p> <p>Utvidet barnetrygd er barnetrygd for et barn mer enn du faktisk bor sammen</p>	

## ANDRE UTFORDRINGER

	Du må søke om utvidet barnetrygd, den blir ikke automatisk innvilget.	
Bruker	Jeg skal IKKE søke utvidet barnetrygd	
Frida	<p>Jeg tror jeg har forsøkt å svare på dette, men det kan virke som at det ikke ble så vellykket som ønsket. Jeg lærer hver da og håper jeg kan svare bedre neste gang du chatter med meg.</p> <p>Har du generelle spørsmål kan jeg sette deg over til en kollega (Mandag Fredag 08.00 - 15.00).</p> <p>Skal jeg sette deg over?</p> <ul style="list-style-type: none"> <li>• Sett meg over (08.00 - 15.00)</li> <li>• Nei takk</li> </ul>	
Bruker	Ja	
Frida	<p>Jeg setter deg over, et øyeblikk.</p> <p>Hvis alle mine kollegaer er opptatt, vennligst vent så hjelper vi deg så snart vi kan.</p>	

Her er det en bruker som mistenker at h\*n har søkt barnetrygd ved en feil, og spør derfor Frida om hjelp. Frida kommer da med informasjon om barnetrygd i stedet. Brukeren prøver da å presisere at vedkommende ikke skal ha barnetrygd. Frida svarer da at hun allerede ha prøvd å besvare spørsmålet, uten å merke ordet 'ikke'. Brukeren blir derfor satt over til veileder i stedet.

Når Frida misforstår på denne måten, som vist i disse to eksemplene fører det ofte til at brukeren enten benytter store bokstaver eller banneord, noe jeg tolker som at bruker blir irritert. Studier tyder på at brukere har lettere for å ty til banneord i samtale med en chatbot, sammenlignet med et menneske (Hill m.fl., 2015). Ofte vil en bruker bruke ordet 'ikke' nettopp fordi det har vært en misforståelse tidligere. Når en bruker da prøve å presisere ved å skrive at det ikke var det brukeren trengte det hjelp til, kommer bare Frida med mer informasjon om dette.

Her følger et eksempel på hvordan analysen til Frida fungerer når brukerutsagnet inneholder 'ikke'.

ID: ??? (???.???.2019)
------------------------

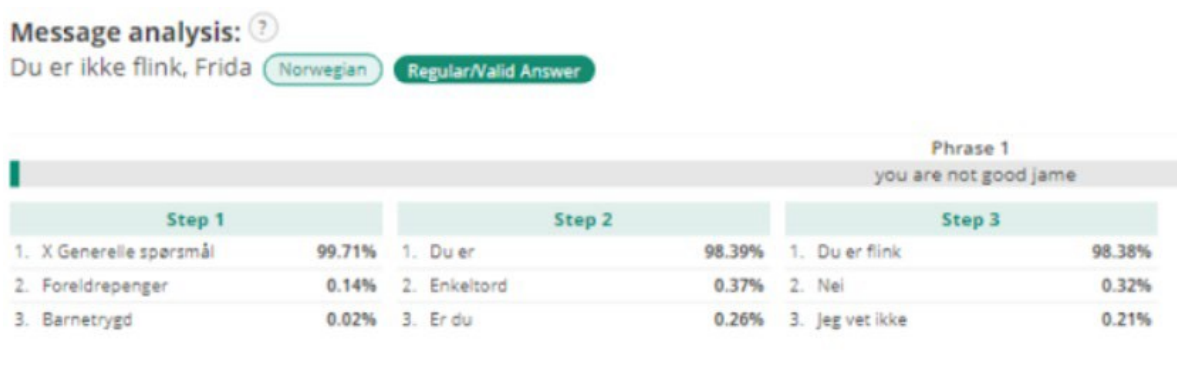
## KAPITTEL 12

Hvem	Tekst	Intensjon
Frida	(...)	
Bruker	Faen	Unknown
Frida	Beklager, men jeg forstår ikke hva du spør om  Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.  Kan du stille spørsmålet på nytt?  <ul style="list-style-type: none"> <li>• Trenger mer hjelp</li> <li>• Nei</li> </ul>	
Bruker	Du er ikke flink, Frida	X Generelle spørsmål / Du er flink
Frida	Takk! Det var hyggelig å høre.	
Bruker	Hade din sopp	X Generelle spørsmål / Ha det bra
Frida	Ha det bra!	

*Utdrag fra chat*

Her er brukeren tilsynelatende ikke fornøyd med Frida sine svar og skriver ‘Du er ikke flink, Frida’, noe Frida tolker som et kompliment ettersom hun ignorerer ordet ‘ikke’.

Slik ser analysen ut:



Første steg tilsier at det er 99.71% sjanse for at det er et generelt spørsmål og hører til intensjonstreet ‘Generelle spørsmål’. I neste steg finner Frida en 98.39% match med ‘Du er’, noe som i steg 3 leder til en 98.38% match med ‘Du er flink’. Siden det er over 45% sjanse

for denne intensjonen og differanse på minst 15%, vil reglene for setningsanalyse gjøre at Frida svarer basert på intensjonen 'Du er flink'.

### 12.4 Oppsummering

I dette kapitlet har jeg vist at det er flere utfordringer som kan oppstå som ikke hører hjemme i noen av kategoriene jeg har presentert tidligere. Brukerne kan vise tegn på manglende tillit; ved å spørre om å bli satt over heller enn å forsøke å stille Frida spørsmål selv, ikke legge merke til at Frida er en robot; ved å stille spørsmål om dette langt ute i samtalen, eller så kan ord Frida ignorerer føre til problemer; ved at brukeren ikke skjønner hvorfor Frida svarer som hun gjør.

# 13 Diskusjon

I denne delen vil jeg diskutere utfordringene vi har funnet og hvordan disse sammenfaller med studiene jeg presenterte i kapittel 2.

## 13.1 Skjulte funksjoner

Jeg har gjennom oppgaven vist en rekke eksempler i fra chatlogger hvor man kan se at intensjonene til Frida, samt funksjoner som stavekontroll og identifisering av språk kan føre til frustrasjon hos brukeren fordi brukeren ikke ser hva som skjer. I slike situasjoner vet ofte ikke brukerne hva de skal gjøre annerledes heller, fordi de ikke forstår hva som skjer. Dette støttes opp av studien til Luger and Sellen(2016) av Siri og Google Now som konkluderte med at brukere som ikke vet hva kunstig intelligens er i stand til, eller hvordan den fungerer, ofte vil bli frustrerte og gi opp. De som derimot forstod hvordan Siri og Google Now fungerte ville heller prøve å omformulere spørsmålene sine og bruke nøkkelord. Dette ser man også ved bruk av Frida. Noen brukere forstår hvordan de skal omformulere spørsmålene og stiller korte spørsmål eller bare skriver nøkkelord, mens andre brukere viser mangel på forståelse og vil gjerne skrive det samme spørsmålet på en like kompleks måte og deretter gi opp. Disse brukerne burde i større grad bli gjort klar over hva Frida er i stand til å svare på, og hvordan de skal interagere med henne. Jeg har i denne studien ikke hatt anledning til å intervju brukerne, men det kan tenkes at også brukere som i større grad forstår hvordan chatboten fungerer, også kan ha problemer med å forstå hva som skjer om Frida plutselig er inne på feil intensjon eller feil språk. I disse tilfellene vil det som regel heller ikke hjelpe å formulere spørsmålet på annen måte eller bruke nøkkelord fordi man allerede er inne på feil spor. Da må man i så fall prøve å nullstille Frida, ved å hente få to like intensjoner på rad eller en 'unknown'-intensjon, men dette vil være vanskelig å gjøre for bruker når de ikke ser hva som skjer.

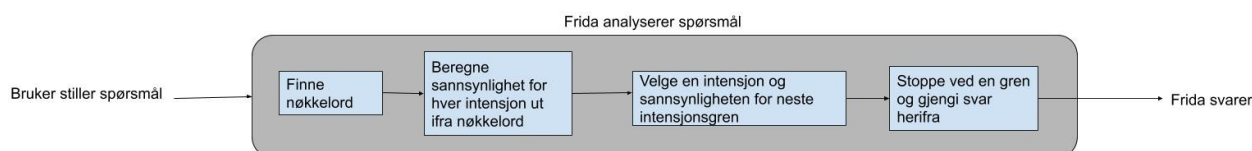
Ord som blir ignorert, slik som 'ikke', fører også til irritasjon hos brukerne fordi de ikke forstår hva som skjer. Dette fører ofte til økt irritasjon fordi problemer oppstår når brukerne prøver å fortelle Frida at hun har misforstått og at det ikke var det de lurte på. Og det er vanskelig å rette på for brukerne når man ser Frida er på feil spor.



## DISKUSJON

Når man starter en ny chat med Frida, får man kun beskjed om at man skal skrive korte setninger og at Frida ser på nøkkelord. Man får ikke beskjed om at Frida kobler nøkkelord opp mot en intensjon og deretter bruker sannsynlighetsberegning for å finne riktig intensjonstre eller intensjonsgren og at nye spørsmål kan bli koblet opp mot en intensjon basert på forrige spørsmål. De får heller ikke vite det, om lav sannsynlighet gjør at Frida går bort i fra forrige tema/intensjon og at brukers oppfølgingsspørsmål derfor ikke vil analysert ut ifra det temaet brukers foregående spørsmål har dreid seg om. Men det kan tenkes at om Frida hadde synliggjort hvordan hun tolket setninger kunne dette ha virket avskrekkende for brukere med lav teknisk kompetanse. Det er allikevel mulig at en mellomting ville ha gjort bruken av Frida enklere.

Dette blir særlig tydelig om man bruker Suchman sitt rammeverk. Flere steg blir gjort mellom hvert brukerutsagn og hvert svar fra Frida, uten at bruker ser det. Første steg er å velge ut hvilke ord som er relevante. Noen ord vil bli fokusert på, såkalte nøkkelord, mens andre ord vil bli ignorert, såkalte stoppord. Allerede her kan det hende at en viktig del av brukers spørsmål forsvinner, uten at brukeren er klar over det. Deretter vil Frida bruke sannsynlighetsberegning til å velge intensjon. Om Frida her velger feil intensjon, så blir ikke brukeren gjort klar over dette. Man kan derfor risikere at brukeren får feil svar uten å være klar over det og deretter stiller oppfølgingsspørsmål som Frida besvarer ut ifra den gale intensjonen hun er inne på.



Å gjøre noe av dette synligere kunne derfor muligens ha endret hvordan brukerne hadde interagert med Frida.

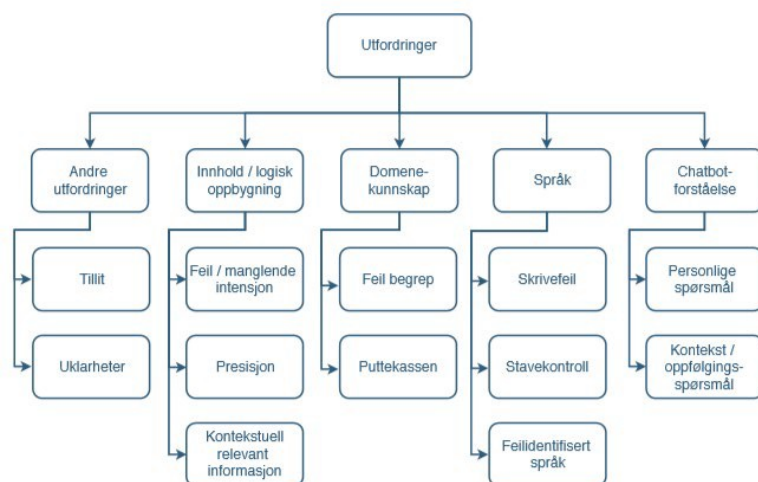
Når Frida får spørsmål om et skjema er det derfor viktig at hun skriver tydelig hvilken ytelse det skjema hun henviser brukeren til, gjelder for, slik at brukeren forstår det om skjema ikke stemmer med det brukeren spør om. Dette sammenfaller også med retningslinjen om å gjøre det klart hvorfor systemet gjorde som det gjorde (Amershi m.fl., 2019). Ved å også innlede svarene med formuleringer av type: «Jeg forstår det sånn at du lurte på dette» så viser hun tydelig hva hun har forstått slik at bruker kan se at svaret stemmer overens med spørsmålet.

Dette kan øke tilliten til svarene og minske sjansen for å få feil svar uten å oppdage det, og sjansen for at man stiller oppfølgingsspørsmål innenfor feil intensjon.

Frida gjør dette til en viss grad ved å komme med fraser som «Jeg hjelper deg gjerne med foreldrepengesøknaden» før hun forteller videre om hvordan. Men i de tilfellene hvor hun er inne på en intensjonsgren i et intensjonstre, kan hun også svare mer konkret om søknader og utbetalinger uten at det nødvendigvis er riktig skjema, eller utbetaling av riktig ytelse.

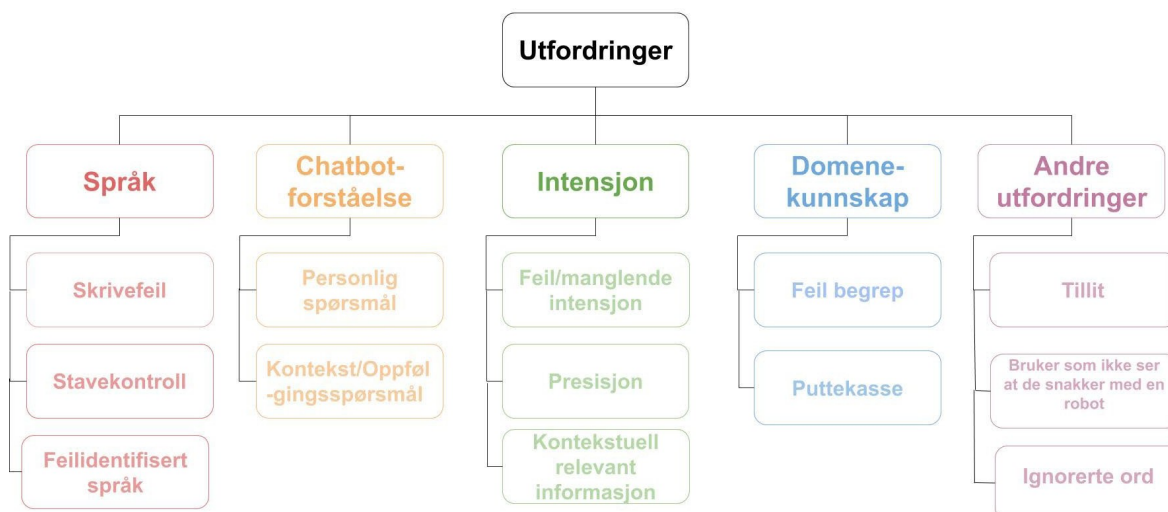
## 13.2 Kategoriseringen av utfordringene

Kategoriseringen av utfordringer har vært en kontinuerlig prosess hvor underkategoriene har blitt tilpasset etter hvert som det karakteristiske ved hvert problem har blitt mer tydelig. Inndelingen jeg har kommet frem til i denne masteroppgaven varierer derfor litt fra Linett Simonsen sin masteroppgave, som igjen varierer litt fra artikkelen vi skrev. De utfordringene Linett presenterte i sin masteroppgave er de utfordringene vi hadde kommet frem til i fellesskap på dette tidspunktet.



Figur 17: Oversikt over utfordringene presentert i Linett sin oppgave (Simonsen, 2019)

## DISKUSJON



Figur 18: Min oversikt over utfordringer

Linett har en underkategori som heter 'Uklarheter' under 'Andre utfordringer' som jeg har gått bort ifra. Jeg syntes denne underkategorien var lite presis så jeg jobbet videre med den og valgt til slutt å fjerne den og heller legge til to nye underkategorier her: 'Brukere som ikke ser at de snakker med en chatbot' og 'Ignorerte ord'. Jeg har også endret navn på kategorien 'Innhold/logisk oppbygning' til 'Intensjon'. Dette er fordi jeg mener at underkategoriene 'Presisjon' og 'Kontekstuell relevant informasjon', også dreier seg om intensjoner. 'Presisjon' handler om at bruker stiller et veldig spesifikt spørsmål og får et generelt svar. Her er det intensjoner med lange svar, uten avklaringer rundt brukers situasjon på forhånd. 'Kontekstuell relevant informasjon' handler om at nøkkelordet blir koblet opp mot en intensjon som ikke er relevant for konteksten. Begge disse underkategoriene dreier seg dermed om å tilpasse intensjonene, slik at de føles mer treffende for brukeren, selv om det er et generelt svar.

De type problemene som har oppstått, har også endret seg underveis. Vi har begge sett på sett på logger ifra mai om startpunkt for analysen, men jeg har også sett på logger helt frem til desember. Noen av problemene vi fant i mai vil derfor være rettet opp i chattene fra

desember. Jeg har fortsatt valgt å ta med slike problemer i denne oppgaven, men jeg har da forklart hvordan dette har blitt endret i etterkant.

Jeg hadde også et møte med NAV i slutten av august hvor jeg diskuterte noen av de første utfordringene jeg la merke til, blant annet at brukerne tilsynelatende ikke fikk med seg all teksten Frida hadde skrevet fordi svarene var for lange, og at mange spurte langt ut i chattene om Frida var en chatbot. NAV sin informant kunne da fortelle at dette var ting de jobbet med å forbedre.

Ettersom at perioden av logger jeg så på var ifra mai-desember 2019, så kan mange endringer ha blitt gjort allerede, i etterkant av min datainnsamling. Det kan derfor hende at noen av kategoriene ikke er like relevante nå, etter hvert som det gjøres flere endringer. Noen utfordringer tenker jeg allikevel vil være relevante selv om chatbotens design og intensjoner har blitt endret på.

Språklige utfordringer som stavekontroll og identifisering av språk, vil nok bli mindre relevante etter hvert som Frida blir bedre til å gjenkjenne språk og flere synonymer blir lagt til av chatbot-trenerne. Utfordringer knyttet til Intensjoner er også noe det stadig jobbes med å forbedre. I de tilfellene hvor det har manglet intensjoner vil dette bli lagt til, og etter hvert som Frida blir bedre vil det være større sjanse for at brukernes utsagn vil kobles til riktig intensjon. Men selv om problemene vil bli mindre gjeldende over tid, er det vanskelig å fjerne slike problemer helt.

Chatbotforståelse og domenekunnskap, derimot, tenker jeg er mer universelle utfordringer som kan være vanskeligere å gjøre noe mer. Chatbotforståelse handler om at brukerne ikke skjønner hvordan de skal interagere med chatboten og hvordan de skal formulere spørsmålene sine. Mangel på domenekunnskap gjør at brukerne ikke klarer å formidle spørsmålene sine og bruke riktige begreper. Disse to tingene kan det være utfordrende å finne en løsning på, og det er derfor også disse utfordringene jeg vil fokusere på i den videre diskusjonen.

### **13.3 Domenekunnskap**

I artikkelen “I’m disabled and married to a foreign single mother” Public service chatbot’s advice on citizens’ complex lives, som jeg skrev sammen med Linett Simonsen, Guri Verne og Tone Bratteteig, diskuterte vi utfordringen domenekunnskap. I denne artikkelen tok vi

## DISKUSJON

utgangspunkt i datainnsamlingen og analysen Linett Simonsen og jeg hadde gjort i forbindelse med masteroppgavene våre, og gikk i dybden på hvordan domenekunnskap kan gjøre bruk av chatbot vanskelig. Artikkelen tar for seg tre underkategorier til domenekunnskap 'feil begrep', 'puttekassen' og 'a logical/legal flaw', og designimplikasjoner.

### **Feil begrep**

Veilederne i Verne sin studie gjenkjente mangel på domenekunnskap ved bruk av feil eller inkonsekvente begreper fra innringerne. (Verne, 2014) En studie blant immigranter i Norge viste at selv om de hadde bodd i Norge i flere tiår og snakket norsk, var NAV sine begreper vanskelig å forstå fordi begrepene ikke hadde noen tilsvarende i deres morsmål. (Simonsen m.fl., 2020 : Nguyen, 2019) Dette gjelder også innbyggere født i Norge uten kunnskap om NAV sine tjenester. Veilederne fortalte oss at de vanligvis bruker «muntlig språk for da blir det mer forståelig for brukerne» i chattene, og de deler opp informasjonen for å gjøre det lettere å forstå. «Det er veldig sjelden at du bare kan bruke standardsvarene» sa en annen veileder, og svarene må tilpasses, «Jeg ser hver chat litt an, og konkret tilpasser informasjonen – ikke for mye, ikke for lite». I tillegg så legger de vekt på at brukerne spør på mange forskjellige måter: «Det er utrolig mange nyanser og utrolig mange ulike måter å løse ting på». (Simonsen m.fl., 2020)

Dette kan man også se i eksemplene jeg har presentert i kapittel 10.1.1. Brukere med mangel på domenekunnskap benytter ofte uttrykk som ikke samsvarer med Frida sine nøkkelord, noe som gjør det vanskelig å bli koblet til riktige intensjoner. I eksempelet presentert i 10.1.1 kobles 'trygd innbetaling barn' til intensjonen 'Utbetaling' og bruker får ikke svar på det vedkommende lurer på. Der hvor en veileder kunne ha tilpasset svarene når de så at brukeren uttrykte mangel på domenekunnskap, så har ikke Frida denne muligheten. Brukeren får et standard svar og i dette tilfellet passer heller ikke svaret, siden brukeren ikke klarer å formidle spørsmålet sitt til Frida.

### **Puttekaske**

Som vi poengterer i artikkelen, ser også den andre underkategorien 'puttekaske' ut til å være et språkproblem. Brukeren lurer gjerne på om vedkommende har rett på en spesifikk ytelse. Dette problemet blir bekreftet av Verne. Hennes studie konkluderte med at mange av innringerne lurte på hvordan deres spesifikke livssituasjon passet med reglene. Selv innringere med god kunnskap om reglene ville ha bekreftelse på om de hadde tolket reglene

riktig. Dette kom også frem i intervjuer med veiledere som sa: ««Ofte vil de ha bekreftelse på ting, fordi det er et komplisert regelverk og mye nyanser» og «Brukerne kontakter oss for å forsikre seg om at de har gjort ting riktig». Veilederne sa også at «Svaret er ja, det finnes ofte på nav.no Men vi skjønner at de chatter med oss. For det er kanskje mye informasjon også. Også har de lest det, men ikke forstått det helt.» De sa også at «Søknadsskjema har blitt mye lettere, men folk slutter ikke å kontakte NAV.» En veileder forklarte at «En utfordring hos NAV er at ting overlapper mye. Det er ingen klare grenser, foreldreting flyter over i [andre temaer] Det er mange terminologier ligner, men som ikke er helt like, men nesten like: barnetrygd, barnetillegg.»(Simonsen mfl., 2020)

Frida kan ikke svare på spørsmål om puttekassen fordi chatboten bare svarer på generelle spørsmål, og referer til generelle regler og regelverk. Frida kan hjelpe brukerne med å finne generell informasjon, men hun klarer ikke å koble det opp mot den faktiske situasjonen til innbyggerne. I eksempelet 679785 i 10.1.2, skriver en bruker ‘Skal ha barn, hva må jeg søke om?’, noe Frida ikke klarer å koble opp mot en intensjon. Bruker skriver deretter ‘Hvordan får jeg penger når jeg venter barn?’ som kobles opp mot utbetalingsdatoer. Eksempler som dette viser at mange brukere sliter med å forstå hvordan deres spesifikke livssituasjon passer med NAV sine tjenester og ytelser.

### **A logical/legal flaw**

I artikkelen har vi også skrevet om en tredje type domenekunnskap som vi har kalt ‘A logical/legal flaw’. Denne varianten av domenekunnskap er en misforståelse av reglene. Ettersom at chatboten ikke registrerer denne misforståelsen, er det en fare for at brukeren blir feilinformert, og får en ubehagelig overraskelse hvis vedkommende regnet med en ytelse som det viser seg at brukeren ikke har rett på. Veilederne fortalte oss at mange ikke forstår hva som står skrevet på nav.no. Kombinert med hva veilederne fortalte oss om tilpasninger av språket og svaret, til den individuelle brukeren, og de mange nyansene i både spørsmål og svar, ser vi at menneskelig chat tilbyr hva kontaktsenter vanligvis gjør «Jeg svarer på deres nivå».(Simonsen mfl., 2020)

Jeg har valgt å ikke inkludere denne varianten av domenekunnskap i masteroppgaven min. Misforståelse av reglene har også en sammenheng med forståelse av NAV sine tjenester og ytelser og hvordan egen livssituasjon passer med dette. I chatloggene jeg har gått gjennom har jeg sett mange eksempler på at det ofte er en kombinasjon av disse som skaper problemer. Brukeren har misforstått aspekter ved tjenestene og ytelsene, og har samtidig en manglende

## DISKUSJON

forståelse av hvordan dette passer med egen livssituasjon. En misforståelse om ytelser og hvem en ytelser gjelder for, gjør det også vanskelig å plassere egen situasjon. Jeg har derfor valgt å slå sammen disse og plassere begge problemene under puttekasse.

Jeg har da plassert chat 125315, som vi i denne artikkelen har plassert i A logical/legal flaw', under 'Puttekasse' i stedet. Men dette er et eksempel på en chat som det er vanskelig å knytte opp mot en spesifikk utfordring, ettersom at problemet her er sammensatt. Det er flere aspekter rundt brukers situasjon som trenger å bli avklart for at brukeren skal kunne få riktig svar. Dette var en av chattene vi diskutert med en veileder, som uttalte at det i dette tilfellet ville være nødvendig å stille flere spørsmål for å avklare hva brukeren egentlig lurte på. 'Barnetillegg' er tilknyttet Arbeidsavklaringspenger, mens 'barnetrygd' er tilknyttet knyttet til Familie, og brukeren hadde derfor antagelig blitt satt over til Arbeid og Helse om vedkommende hadde ringt.

### **Øke kunnskapsnivået til bruker**

En løsning på dette kan være at chatboten kan komme med mer informasjon, om brukeren for eksempel bruker begreper som ikke brukes av NAV. Om chatboten kan plukke opp kunnskapsnivå til bruker og har egne intensjoner for dette, vil også brukeren kunne føle seg sikrere på at de ordningene brukeren bli informert om stemmer for sin situasjon. Chatboten må dermed reagere på nøkkelord som kommer ifra dagligtale og ikke bare NAV-begreper, og også kunne tilpasse mengden informasjon som er nødvendig for brukeren.

Mange brukere vil ikke lese alt om Frida kommer med for mye informasjon på en gang, derfor er det viktig at dette kan tilpasses i en viss grad, ved at de som bruker riktig begreper og spør om riktige støtteordninger blir koblet om mot andre intensjoner med presise svar uten for lange blokker av informasjon. Skatteetaten sin rapport fra 2018(Arlov, m.fl) viser også at det er viktig å avklare hva brukeren egentlig er ute etter, og dermed unngå misforståelser og feilinformasjon, ved å spørre 'hva trenger du den til?', i stedet for 'hvilket skjema ser du etter?'.

Veilederne vi snakket med utrykte at det var flere brukere som var usikre på hvordan NAV sine støtteordninger fungerte. En av veilederne sa dette: «*Jeg opplever at det er veldig mange som spør om sånn 'Hei, jeg planlegger barn. Jeg har termin, jeg kommer nok til å ha termin sånn og sånn. Hva burde jeg passe på? Hvor søker jeg?'*» Dette viser at noe av den viktigste jobben til NAV sin førstelinjetjeneste er å lede brukeren til riktig ordning og forklare hvordan

## KAPITTEL 13

den fungerer. Mye av informasjonen finnes på NAV.no, men allikevel er det mange som synes dette er uklart og ikke skjønner hvordan begrepene passer med deres situasjon. Frida må derfor kunne veilede brukeren frem til riktig informasjon og avklare at det er den ordningen som passer for brukerens situasjon.

Det vil allikevel være en utfordring knyttet til bruk av chatbot ettersom at Frida ikke vil kunne oppdage en misforståelse om ikke brukeren selv ser at informasjonen Frida kommer med, ikke passer med deres situasjon. Det er derfor viktig at Frida avklarer ting tidlig i samtalen, ved å stille brukeren spørsmål og komme med tilstrekkelige forklaringer på hva hun svarer på.

Det er også flere tilfeller der brukere kontakter Frida kun for å spørre om hva et begrep betyr. Eksempler på dette under:

ID: 125146 (14.05.2019)		
Hvem	Tekst	Intensjon
Frida	<p>Jeg heter Frida og er en chatbot.</p> <p>Jeg svarer best på korte spørsmål. Husk å ta med hvilken stønad du spør om.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• For english chat, click here</li> </ul>	
Bruker	Terminbekreftelse, hvilken skal skrives her	Foreldrepenger/ Foreldrepenger
Frida	<p>Foreldrepengene skal sikre foreldrene inntekt i forbindelse med fødsel og adopsjon.</p> <p>Hva lurer du på?</p> <ul style="list-style-type: none"> <li>• Har jeg rett?</li> <li>• Periodens lengde</li> <li>• Fleksibelt uttak</li> <li>• Alt om søknaden</li> <li>• Les mer</li> <li>• Status i sak</li> </ul>	
Bruker	rådgiver	X Om NAV/ Snakke med et menneske
	Hvis du ønsker å snakke med oss, kan du ringe 55 55 33 33 alle hverdager mellom kl 08.00 og 15.30.	



## DISKUSJON

	<p>Gjelder det pensjon/uføretrygd eller hvis du er arbeidsgiver, samhandlingspartner eller behandler, se telefonnummeroversikten.</p> <p>Klikk "Chat" for informasjon om chat-tjenesten vår.</p> <ul style="list-style-type: none"> <li>• Chat</li> <li>• Telefonnummer</li> </ul>	
--	--	--

I denne chatten prøver bruker å spørre om hva som skal skrives under 'Terminbekreftelse'. Frida informerer så om foreldrepenger i stedet, og bruker spør om å bli satt over i stedet. Jeg tolker dette som at bruker vil snakke med et menneske for å få hjelp til å tyde begreper på grunn av manglende domenekunnskap.

ID: 125212 (14.05.2019)		
Hvem	Tekst	Intensjon
Bruker	Hva er tilsynstid?	Unknown
Frida	<p>Beklager, men jeg forstår ikke hva du spør om.</p> <p>Jeg skjønner best hvis du skriver korte spørsmål og tar med hva slags stønad det gjelder.</p> <p>Kan du stille spørsmålet på nytt?</p> <ul style="list-style-type: none"> <li>• Trenger mer hjelp</li> <li>• Nei</li> </ul>	
Bruker	Hva er tilsynstid, søknad om barnebidrag?	Bidrag / Søknad bidrag
Frida	<p>Hvis du og den andre forelderen ikke kommer fram til en avtale om barnebidraget privat, kan hver av dere søke om å få bidraget fastsatt av NAV.</p> <p>Husk at du må betale gebyr når NAV gjør vedtak om bidrag etter søknad fra enten deg eller den andre parten.</p> <ul style="list-style-type: none"> <li>• Når søke?</li> <li>• Hvordan søke?</li> <li>• Mer om gebyr</li> <li>• Barn over 18 år</li> <li>• Endre bidraget</li> </ul>	

*Utdrag fra chat*

I denne chatten spør bruker om hva tilsynstid er, men bruker klarer ikke å koble dette opp mot noen intensjon. Bruker forsøker å omformulere spørsmålet sitt ved å skrive ‘Hva er tilsynstid, søknad om barnebidrag?’ Frida kobler da dette til ‘Søknad bidrag’, men bruker får ikke svar på hva tilsynstid er.

I begge chattene over er det brukere som lurere på spesifikke begreper. Frida klarer ikke å koble dette opp mot riktige begreper og brukerne får heller ikke svar på det de lurere på.

### 13.4 Chatbotforståelse

Som vist i eksemplene i kapittel 9, er det mange som har problemer med å forstå hvordan de skal formulere seg for at chatboten skal forstå dem. Mange stiller lange personlige spørsmål som Frida ikke kan svare på. Dette ble også bekreftet av veilederne.

«Noen bare fortsetter på samme måten. [skriver lange spørsmål] Selv om hun på en måte sier: «Jeg jobber best med korte setninger. Jeg skjønner ikke hva du mener.» Men da er hun veldig rask med å på en måte sette over til en kollega. Så jeg har sett de samme spørsmålene som de på en måte har spurt Frida, hun setter over til, så bare copy-paster de og limer inn de samme spørsmålene til meg.»

Mange brukere har manglende forståelse for hvordan de skal formulere et spørsmål slik at chatboten forstår det. Det er viktig at disse brukerne får tilstrekkelig informasjon slik at de forstår hvordan de skal kommunisere med Frida. Dette er også grunnen til at hun oppgir at du må bruke korte setninger og gjerne oppgi hvilken ytelse det gjelder. Hvis man i tillegg har manglende domenekunnskap, så kan dette gjøre kommunikasjonen med Frida vanskelig.

Et problem er at Frida er en chatbot du kontakter i en spesifikk livssituasjon, i motsetning til en chatbot du kommuniserer med over tid, slik som Siri. Luger and Sellen(2016) sin studie viste at mange lærte seg hvordan de skulle bruke nøkkelord og korte setninger etter hvert, men denne type læring via utforskning, vil ikke være like relevant for en slik chatbot som Frida. Med en chatbot man interagerer med over tid, vil man gjerne tilpasse seg og forstå hvordan man skal formulere seg til neste gang man skal kommunisere med den. En NAV-bruker vil ikke nødvendigvis prøve å ta kontakt på samme måte igjen. Frida må derfor gjøre det tydelig hva hun kan, og hvordan du kan interagere med henne.

## DISKUSJON

Det kan allikevel tyde på at noen brukere klarer å tilpasse seg Frida i løpet av samtalen og forstår hvordan de må formulere seg for å få svar. En veileder uttalte at «Jeg har sett en del opp i gjennom som har prøvd og så skrive til Frida, for eksempel formulere et spørsmål, først så starter de kanskje med å formulere spørsmålet på én måte, men så greier ikke Frida å besvare det, for at det er noe med setningen som gjør at Frida ikke greier å prosessere det, og så prøver de kanskje å forenkle språket sitt noe veldig da, for å få Frida til og så greie å prosessere det, da» Dette tyder på at enkelte brukere skjønner hvordan de må omformulere seg for å bli forstått. En annen ting som varierte mellom samtalen med chatboter og med mennesker i denne studien, var at i samtalen med maskiner fremstod menneskene i større grad som uhøflige og irriterte. Dette kan også være knyttet til at brukerne tenkte over at det var en robot de snakket til og ikke følte behovet for den høfligheten de ville brukt om de snakket med et virkelig menneske, fordi de ikke tenkte på hvordan de ville bli oppfattet. (Hill m.fl, 2015) En veileder uttalte dette: «Men jeg har også sett de hvor de vet at det er robot så de svarer veldig kort. «Ja, det er foreldrepenger.» Også får du jo ulike alternativer av Frida, ikke sant? Og hvis de da ikke skjønner helt, så skriver de «jeg vil snakke med en veileder». Og da skriver de lengre spørsmål, lengre avhandlinger, fordi da skjønner de at de snakker med et menneske. Og kan på en måte skriver litt mer som de vil da.» En annen veileder uttalte seg også om adferden til brukerne ut ifra hvem de snakket med «(...) så litt mer kommanderende fordi at de vet at de snakker med en robot, og det er ikke så mye bruk av smiley-fjes og sånt til Frida, kanskje.» Dette viser at noen bruker stiller spørsmålene på en annen måte, når de snakker med en chatbot.

Det viktigste tiltaket for at brukerne i større grad skal forstå hvordan de skal kommunisere med en chatbot, er å gjøre det tydelig både hva Frida kan og hvorfor hun gjør det hun gjør. Ved at Frida viser hvilken intensjon hun er inne på, ved å gjøre det klart at hun oppfatter at brukeren lurte på foreldrepenger og utbetalinger knyttet til dette, om brukeren bare skriver utbetaling som oppfølgingsspørsmål, vil det også gjøre at brukeren skjønner hva som skjer. Da kan brukeren igjen spesifisere hvilken ytelse det er snakk om, slik at Frida kommer seg inn på riktig intensjonstre, uten at brukeren trenger å bli informert om hvordan intensjonstrærne fungerer.

### 13.5 Brukere som blir satt over

Som nevnt i kapittel 6 er det tre hovedgrunner til at brukere spør om å bli satt over til veileder: Bruker blir ikke forstått, Bruker får relevant svar, men stoler ikke på svaret og Bruker får svar, men skjønner ikke at de har fått svar fordi det er så generelt. I tillegg er det brukere som har spørsmål som skal settes over til veileder, hvis det for eksempel krever innsyn i sak. Men de tilfellene som er mest interessante å se på er der hvor Frida kan svare, men brukere allikevel vil bli satt over. Dette kan skyldes både mangel på forståelse av hvordan chatboten fungerer og mangel på tillit.

Veilederne fortalte at det var mange som ville rett til menneske-chat: «Så på noen, de vi får, så ser vi jo at de kanskje skrive 'veileder' med en gang. De er ikke interessert i å chatte med Frida. Det er gjerne de som er sånn superchatterne våre, som kanskje bruker det oftere. De vet at hvis de skriver 'menneske', 'veileder' så tar det under et minutt så har hun på en måte satt over til en ledig menneskelig veileder da.»

Andre skriver først spørsmålet sitt til Frida, men stiller så det samme til en veileder etterpå. På spørsmål om dette skjer ofte sa en veileder: «Ja, det er noen ganger. Det er ikke alle som på en måte. Det blir for mye informasjon for de da, kanskje. Når Frida gjør det, så kan det gå noen sekunder og så har hun skrevet mange linjer, ikke sant? Så blir man litt sånn forvirret av «oi, hva er det som skjer her, får jeg egentlig svar?» Det er ikke så mange som leser de svarene der. De tenker heller, «okei det her ble for mye. 'Snakk med et menneske'». Og da ser jeg at de egentlig stiller de samme spørsmålene. Det er bare det at det er kanskje lettere for meg å forklare personen da. Med enkle ord.. «Ja, det stemmer. Kontantstøtte kan du tidligst søke når barnet fyller ett år.»» Dette kan tyde på at det ikke bare er mangel på tillit, men også lengden på svarene som kan være et problem. En person som har fått riktig svar og så spør om å bli satt over kan dermed gjøre det på bakgrunn av mangel på tillit, for generelle svar og for lange svar. Noen legger heller ikke merke til at det er en chatbot før lenger ute i samtalen, og spør da om å bli satt over når de innser det. Dette skjer gjerne hvis samtalen går i loop eller Frida ikke svarer på det hun blir spurt om.

### **Bruk av kanaler**

Studiene til Skaarup(2012) og Ebbers(2008) viste at tilbudet av informasjonskanaler kan også ha stor innvirkning på forholdet til myndighetene. På samme måte vil det forholdet man allerede har påvirke hvilke former for kommunikasjon man føler seg komfortabel med. Dette ble også nevnt av veilederne vi intervjuet: “De har aldri vært borti NAV før, så kan jo faktisk

## DISKUSJON

NAV virke veldig skummelt. Folk har jo ulike bilder av NAV. Og kanskje ikke alltid like positivt. I forhold til media og andre ting. Så tror jeg det er derfor at det kanskje er litt ubehagelig å ta kontakt med oss på telefonen. (...) Så tror jeg den chatten, chat-kanalen gjør det litt lettere for de som virkelig gruer seg til å ta den samtalen. De tror kanskje vi er skumle. Og da kan det være greit at vi starter med en melding, eller chat og viser at vi, vi er egentlig ikke så skumle.» Mange føle seg nok mer komfortable med å snakke chat enn telefon og chattefunksjonen til NAV er derfor viktig. Noen kan også foretrekke chatbot fremfor chat med menneske om de ser på NAV som noe skummelt og ikke vil forholde seg til mennesker.

Studien til Hill m.fler(2015) som handlet om hvordan mennesker kommuniserte med en chatbot, sammenlignet med andre mennesker, konkluderer med at brukerne var mer komfortable med chatboten enn antatt, ettersom at de sendte mange meldinger. Men basert på hva jeg har sett i chatloggene jeg har analysert, vil samtaler med mange meldinger fra bruker som regel skyldes at brukeren ikke får svar på det vedkommende lurte på. Det å sende mange meldinger på rad trenger derfor ikke nødvendigvis å være knyttet til å være komfortabel, men heller det at de brukerne blir irriterte over mangel på tilstrekkelige svar. Studien til Følstad m.fl.(2017) tyder allikevel på at bekymringer rundt hvordan man blir oppfattet kan ha en viss innvirkning på hvordan kommunikasjonen med chatboter fungerer i forhold til menneskelige veiledere. Og det er nok også brukere som føler det er lettere å kommunisere med chatbot fremfor å forholde seg til et menneske

En annen veileder snakket også om at mange kunne oppfatte chat som et mer lavterskel tilbud: «Fordelen for oss og for bruker er kanskje det at bruker oppsøker hjelp i god tid, før det er for sent, da. Mens har man kun det telefontilbudet, så er det kanskje veldig mange som aldri tar kontakt fordi de er litt redd for den kanalen.» Å ha flere type kanaler for forskjellige type spørsmål og ulike brukergrupper, er derfor veldig viktig. Frida fører derfor ikke bare til større effektivitet, men også at det kan bli lettere å ta kontakt med NAV.

### **13.6 Riktige forventinger**

Skatteetaten sin rapport(Arlov m. fler) viste også et vanlig problem var at mennesker trodde de snakket med mennesker og derfor også venter menneskelig fleksibilitet. Det er derfor viktig å gjøre det klart hva chatboten er i stand til. For å tydeliggjøre dette er det viktig å

unngå personnavn, bruke avatar eller logo som ikke ligner et menneske og bruke en vennlig tone, men unngå menneskelige responser. Man må gjøre det klart at det er en robot og hva den er i stand til. NAV sin chatbot har både navn og menneskelignende avatar, noe som kan påvirke de forventningene brukeren har. Dette kan igjen ha en påvirkning på hvordan de formulerer spørsmål og hvordan de reagerer når Frida ikke klarer å svare.

Frida 19:28



Figur 19: Skjerm bilde av Frida sin avatar

Også Luger og Sellen sin studie bekreftet at forventningene brukeren har vil påvirke bruk og tillit. Dette betyr at selv om brukerne får svar på det de lurer på, kan de fortsatt spørre om å bli satt over fordi de ikke har tillit til chatboten, som tidligere vist i delkapittel 13.1. Dette påvirker derfor både kategoriene utfordring og tillit.

Forventninger vil ha en stor innvirkning på chatbotforståelse, ettersom at gale forventninger gjør det vanskelig å forstå hvordan man skal interagere med en chatbot. Det vil også påvirke hvordan man oppfatter Frida sine svar om hun svarer med feil intensjon, og hvor forståelsesfull brukeren er om Frida misforstår. Chatbotloggene vi studerte inneholdt også antropomorfiske responser slik som «Å, du får meg til å rødme» og «Jeg forstår godt hva du mener.» Analysen vår tyder på at ikke alle brukere forstår at de kommuniserer med en maskin, og vi foreslår derfor i artikkelen vår å endre designet til chatboten slik at den ikke fremstår som et menneske, men heller klart viser sine ikke-menneskelige egenskaper. (Simonsen mfl., 2020) Slike programmerte svar kan føre til urealistiske forventninger som påvirker brukeropplevelsen (Simonsen mfl., 2020; Maass, S., Rommes, E., 2007)

### 13.7 Sammensatte problemer

## DISKUSJON

Som tidligere nevnt, er det flere faktorer som bidrar til at problemer kan oppstå i samtaler med chatbot. Ofte vil det også være en kombinasjon av utfordringer som skaper problemer. Flere av de ulike utfordringene henger også sammen med manglende tillit. En bruker med manglende domenekunnskap kan være usikker på om svaret er riktig fordi de ikke vet hvordan ordningene fungerer. Mangel på chatbotforståelse vil ofte også medføre mangel på tiltro til Frida sine svar, ettersom at brukerne ikke vet sikkert hva en chatbot kan og hvor sikre svarene de får er. Dette sammenfaller med retningslinjene om å gjøre det klart hva systemet kan gjøre og gjøre det klart hvor bra systemet kan gjøre disse tingene.(Amershi m.fl., 2019) Disse utfordringene kan man prøve å imøtekomme med aspekter ved designet. Om brukere som benytter feil begreper får ekstra informasjon, kan de også føles seg sikrere på svaret. Og om brukere med manglende chatbotforståelse får vite noe om hvor sikker Frida er på svarene sine, så kan det muligens bli lettere for dem å stole på at de får riktige svar.

I mange av eksemplene jeg har presentert fører også ofte flere av utfordringene til feil intensjon. Språklige feil fører gjerne til feil intensjoner, og brukere som stiller personlige spørsmål ønsker gjerne mer presise svar. I tilfellene hvor feil kobling til intensjon er hovedproblemet, hadde muligens mer bruk av knapper gjort dette enklere. Det er mange tilfeller hvor Frida stiller et avklaringsspørsmål og bruker sitt svar blir koblet opp mot en annen intensjonsgren, enten fordi bruker skriver for mange ord eller ord som ligner andre nøkkelord. Ved bruk av knapper kunne slike situasjoner vært unngått. En løsning kan være å oppfordre brukeren til å trykke på knapper i stedet for å skrive et svar, eventuelt fjerne muligheten til å skrive et svar i noen tilfeller. Særlig i situasjoner hvor Frida spør et 'ja'- og 'nei'-spørsmål, slik som når Frida spør om bruker vil bli satt over til menneske, kunne dette gjort overgangen til veileder mer effektiv. På denne måten kunne man unngått problemer hvor bruker skriver 'ja, takk, gjerne'(chat: 678937) og Frida ikke skjønner at bruker vil bli satt over.

Både Intensjoner og Språk er knyttet til designet av Frida og handler om funksjonene til chatboten. Om en bruker stiller et spørsmål riktig, kan brukeren fortsatt få feil svar om Intensjonene er koblet feil eller om stavekontrollen endrer en frase til feil ord.

Chatbotforståelse og domenekunnskap, derimot, handler like mye om måten brukerne stiller spørsmål på. En bruker med mangel på enten chatbotforståelse eller domenekunnskap vil kunne ha problemer med å bruke Frida, samtidig som at en bruker som er vant med chatboter og kjent med NAV sine begreper vil kunne benytte chatboten uten disse utfordringene. Å

tilpasse Frida slik at designet også tar hensyn til disse brukergruppene, vil derfor også gjøre chatboten mer inkluderende.

## 13.8 Designimplikasjoner

Jeg har gjennom denne diskusjonen kommet med en del forslag til designgrep for å møte de ulike utfordringene jeg har presentert. Jeg mener at designet burde skape realistiske forventninger ved å gjøre det tydelig at Frida ikke er menneske, og i tillegg gjøre det tydelig hva hun kan svare på og hvor sikker hun er på svarene. Ved å ha en avatar som viser at hun ikke er menneske og gjøre det klart hva hun kan svare på, kan man løse noen av problemene knyttet til chatbotforståelse. Når brukerne skjønner hva Frida er i stand til å svare på, vil dette kunne føre til færre personlige og lange spørsmål. Det vil også kunne øke bruken av spørsmålsformuleringer som er mer tilpasset en chatbot, med korte konkrete spørsmål om generelle temaer. Ved å vise hvor sikker Frida er på hvert spørsmål, kan man også øke tilliten hos brukerne. Når man ser at Frida er usikker på et spørsmål, vet man at man burde dobbeltsjekke at dette passer med situasjonene man er i, og om Frida sier hun er sikker kan man i større grad stole på det, fordi man vet at hun ville ha gjort det klart om dette ikke var tilfellet.

Det burde også være mulig for brukerne å lære domenekunnskap underveis. Hvis Frida kan komme med ekstra informasjon til brukere som benytter feil begrep, vil brukerne lære seg de riktige begrepene og forstå hvilke begreper de skal bruke for at Frida skal kunne koble det til riktig intensjon og gi et riktig svar. Ekstrainformasjon kan også føre til at det blir lettere for brukere som har problemer med å plassere sin egen situasjon ut ifra NAV sine begreper, fordi de får en større innsikt i hva NAV sine begreper dreier seg om. Dette vil også kunne øke tillit, ved at brukerne ser at de får svar på det de faktisk lurte på, og det vil også minske sjansen for at brukerne tror de har fått et riktig svar når svaret ikke stemmer. Lange tekstsvare fra Frida kan ofte føre til at brukerne ikke leser gjennom alt, det er derfor viktig at denne informasjonen er delt opp i forskjellige tekstblokker gjerne med bruk av knapper underveis for å presisere brukerens situasjon og komme med svar som passer til dette. Mer bruk av knapper kan også føre til at man i større grad blir satt over til riktig intensjon, og unngå at brukeren får et svar som ikke er relevant.



## DISKUSJON

Et annet problem jeg har sett, er at mange brukere har problemer med å bli satt over. Å bruke knapper i disse tilfellene, i stedet for tekstsvar, vil derfor kunne gjøre det lettere å bli direkte satt over. Ved å gjøre det tydelig hva Frida kan svare på, vil også brukere forhåpentligvis i større grad forsøke å stille Frida generelle spørsmål, i stedet for å spørre om å bli satt over med en gang. Dette kan man se i enkelte eksempler jeg har presentert, slik som i chat 125246 hvor bruker først spør Frida etter å ha fått beskjed om at det ikke går an å bli satt over på dette tidspunktet, selv om dette er et generelt spørsmål som Frida fint kan svare på.

## KAPITTEL 14

## 14 Konklusjon

I denne oppgaven har jeg forsøkt å svare på følgende forskningsspørsmål:

- Hvilke utfordringer oppstår ved bruk av chatbot i NAV?
- Hva gjør at samtaler blir satt over til veileder?

Jeg har foretatt intervjuer og observasjoner med chatbot-trenere og veiledere, og jeg har analysert chatbotlogger for å finne svar på disse spørsmålene. Til å analysere loggene har jeg brukt Suchman sitt analytiske rammeverk og fokusert på de skjulte funksjonene som brukeren ikke blir gjort oppmerksom på. I analysen presenterer jeg flere kategorier av utfordringer som oppstår, og jeg viser eksempler på disse. Jeg setter disse i sammenheng ved å diskutere sammensatte problemer for brukerne. Til slutt kommer jeg med forslag til mulige designgrep for å møte disse utfordringene.

Gjennom loggene jeg har analysert, har jeg kommet fram til at det er særlig fire kategorier av utfordringer som går igjen: Språk, Chatbotforståelse, Domenekunnskap og Intensjoner. *Språk* handler om at Frida ikke klarer å koble brukeren sitt ord opp mot et nøkkelord og endrer det til et ord hun kjenner til. *Chatbotforståelse* handler om at brukeren ikke skjønner hvordan spørsmål skal stilles og hva slags spørsmål en chatbot kan svare på. *Domenekunnskap* handler om at bruker mangler kunnskap om NAV og NAVs tjenester, noe som gjør det vanskelig å kommunisere med Frida ettersom at systemet baserer seg på at bruker skal kjenne til NAV sine begreper. *Intensjon* handler om situasjoner hvor brukerens utsagn kobles opp mot feil intensjon, eller at en intensjon for det brukeren spør om, mangler. Jeg har også funnet noen utfordringer som ikke passer i noen av disse kategoriene, slik som tillit, ignorerte ord og brukere som ikke ser at de snakker med en chatbot. I tillegg har jeg sett på hvilke faktorer som gjør at en samtale settes over til veileder og i hvilke situasjoner brukere har problemer med å bli satt over: dette er noe som gjerne skjer når brukere skriver for mange ord eller bruker ord Frida ikke kjenner igjen. Jeg har også diskutert sammensatte problemer og reflekter over mulige designgrep for å møte noen av disse utfordringene.

Språk og Intensjoner er lettest å gjøre noe med, ved å legge til flere synonymer og nye intensjoner, og fortsette treningen av Frida. Domenekunnskap og chatbotforståelse er mer komplisert. Problemet med både domenekunnskap og chatbotforståelse er todelt; brukerne har

mangel på kunnskap og systemet legger opp til at brukerne skal ha kunnskap for å bruke det. Det er derfor viktig både å tenke over hva man forventer av brukerne og hvordan man kan tilrettelegge systemet for at det skal brukes av flest mulig, uavhengig av forkunnskaper. Jeg mener derfor at systemet burde tilrettelegges slik at det skal være lett å benytte seg av for brukere med mangel på domenekunnskap og mangel på chatbotforståelse, ved å gi ekstra informasjon om både begreper og regler, og vise tydelig hvordan Frida fungerer. I tillegg burde Frida tilpasse seg brukerne, ved å lære seg de uttrykkene brukerne pleier å benytte. Selv om det er flere designgrep man kan gjøre, vil ingen av utfordringene forsvinne helt. Det er derfor viktig å vise begrensninger og gjøre det klart hva som skjer, ved å vise hva som foregår i systemet når man snakker med Frida.

### **Mitt bidrag**

Jeg har jobbet videre med kategoriene presentert i Linett sin masteroppgave (Simonsen, 2019) og i artikkelen ‘ "I'm disabled and married to a foreign single mother" Public service chatbot's advice on citizens' complex lives' (Simonsen m.fl., 2020) . Ved å gå gjennom og analysere flere logger har jeg endt opp med å slå sammen noen av kategoriene presentert i disse tekstene, slik som ‘puttekasse’ og ‘logical/legal flaw’, og jeg har funnet noen nye kategorier, slik som ‘Ignorerte ord’. Jeg har også sett nærmere på overgangen fra chatbot til veileder, noe det finnes lite forskning på fra tidligere studier.

Tidligere forskning på chatboter har vist at flere meldinger på rad fra brukere kan tyde på at brukerne er mer komfortable (Hill m.fl., 2015). Dette har jeg vist at ikke nødvendigvis stemmer. I mange av loggene jeg har presentert, er grunnen til at brukerne sender mange meldinger, en konsekvens av at de ikke får svar på spørsmålene de stiller.

Jeg foreslår et design hvor det skapes realistiske forventninger til hva Frida er i stand til. Dette kan man gjøre ved å

- Gjøre det tydelig at Frida ikke er et menneske. Unngå bruk av menneskelignende avatar og antropomorfiske responser.
- Gjøre det tydelig hva slags type spørsmål Frida kan svare på.
- Gjøre usikkerhetsmomentene tydeligere. Vise hvor sikker Frida er på svarene slik at brukerne føler at de kan stole på svarene de får.

## KONKLUSJON

Designet burde også legges til rette for brukere med manglende domeneskunnskap og gjøre det mulig for brukere å lære domeneskunnskap underveis i samtalen. Ved at Frida kommer med ekstrainformasjon til brukere som benytter synonymer som ikke er en del av NAV sitt begrepsapparat, kan brukerne lære seg de riktige begrepene samtidig som at de kan føle seg trygge på at de får svar på det de faktisk spurte om. Utrykkene til brukerne kan også legges inn som synonymer slik at Frida lærer seg brukerne sine uttrykk. For å få svarene til å føles mer relevante for brukeren kan chatboten også stille flere spørsmål underveis i dialogen for å avklare situasjonen til brukerne og gi et mer presist svar. Men dette vil kreve en del flere intensjoner og intensjonstrær. I tilfeller hvor Frida forsøker å avklare situasjonen til bruker, kan også knapper benyttes i større grad slik at det blir koblet opp mot riktige intensjoner. Når Frida spør om bruker vil settes over, kan bruk av knapper, uten mulighet for tekstsvar, gjøre at brukere i større grad blir satt direkte over, uten misforståelser. Frida burde også vise tydelig hva hun svarer på i responsene sine, slik at brukeren skjønner hva som skjer og ser at de får svar på det de lurte på, eventuelt skjønner hvor misforståelsen oppsto slik at de vet at de må formulere spørsmålet sitt på nytt.

### **Fremtidig forskning**

I denne studien har jeg ikke involvert brukere. Å gjennomføre intervjuer med brukerne om hvordan de opplever interaksjoner med Frida kan derfor være et forslag til videre forskning. Hadde jeg hatt tid og mulighet til å samle inn mer data, ville jeg også ha sett nærmere på forskjellene på chat med veileder og chat med chatbot. Å se videre på hvilke utfordringer som oppstår i hver av disse type interaksjonene, samt se på hvordan en veileder ville ha besvart spørsmål hvor Frida har problemer, ville vært et interessant utgangspunkt for videre forskning. Det kunne også vært interessant å følge chatter som blir satt over fra chatbot til menneske, og se på hvordan veilederen besvarer disse spørsmålene. En annen mulighet for videre forskning er å gjøre en kvantitativ studie med utgangspunkt i de kategoriene av utfordringer som jeg har presentert.



# Litteraturliste

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennet, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., Horwitz, E. (2019) 'Guidelines for Human-AI Interaction' CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems doi: <https://doi.org/10.1145/3290605.3300233>

Arlov, L., Borgen, E., & Qetkaj, S. (2018). Sluttrapport fra Pilot-teamet. Skatteetaten.

Barland, M. (2016) 'Slik blir du overvåket' *Teknologirådet* postet 28.januar 2016 Hentet 01.06 fra <https://teknologiradet.no/slik-blir-du-overvaket-2/>

Boost AI (2019a) Rapport 'What are intents'

Boost AI(2019b) Rapport 'Core rules'

Boost AI(2019c) Rapport 'Extended rules'

Boost AI(2019d) Rapport 'What is training data'

Boost AI(2019e) Rapport 'What is test data'

Boost AI(2019f) Rapport 'Processed sentence'

Boost AI(2019g) Rapport 'Unknown word penalties'

Botwiki(2018) Botwiki Hentet illustrasjon 01.06 fra <https://botwiki.org/bot/eliza/>

Brandtzæg, P. B. & Følstad, A. (2017) 'Why people use chatbots' *4th International Conference on Internet Science 2017* Hentet fra: [https://www.researchgate.net/publication/318776998\\_Why\\_people\\_use\\_chatbots](https://www.researchgate.net/publication/318776998_Why_people_use_chatbots)

Bratteteig, T. & Herstad, J. (2020) «Notat om sekvens av handlinger mellom menneske og maskin» Hentet fra <https://www.uio.no/studier/emner/matnat/ifi/IN1030/v20/Pensumliste%20/notat-sekvensavhandling2020jh.pdf>

Datatilsynet (2018) 'Kunstig intelligens og personvern' Rapport. Datatilsynet. Hentet 01.06 fra <https://www.datatilsynet.no/globalassets/global/dokumenter-pdfer-skjema-ol/rettigheter-og-plikter/rapporter/rapport-om-ki-og-personvern.pdf>

Datatilsynet(2019) 'Hva er en personopplysning' Hentet 01.06.20 fra: <https://www.datatilsynet.no/rettigheter-og-plikter/personopplysninger/>

Ebbers, W.E., Pieterse, W.J., Noordman, H.N. N. (2008) "Electronic government: Rethinking channel management strategies," *Gov. Inf. Q.*, vol. 25, no. 2, pp. 181–201, Apr. 2008 doi: <https://doi.org/10.1016/j.giq.2006.11.003>

Følstad, A., Nordheim, C.B., Bjørkli, C.A 'Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study'(2018) *The Fifth International Conference on Internet Science* doi: 10.1007/978-3-030-01437-7\_16

Gibbs, S. (2015) 8. juli 2015 *The Guardian* 'Women less likely to be shown ads for high-paid jobs on Google, study shows' Hentet 26.06.20 fra <https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study>

Government: 6th International Conference (EGOV 2007), vol. 4656/2007. Springer Berlin/Heidelberg, pp. 216–227, 2007 doi: [https://doi.org/10.1007/978-3-540-74444-3\\_19](https://doi.org/10.1007/978-3-540-74444-3_19)

Grönlund, A., Hatakke, M., Ask, A. (2007) “Inclusion in the E-Service Society – Investigating Administrative Literacy Requirements for Using E-Services,” *Electronic*

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’ *Science Direct: Computers in human behaviour* doi: <https://doi.org/10.1016/j.chb.2015.02.026>

Hunt, E. (2016) 24. mars 2016 ‘Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter’ *The Guardian*. Hentet 26.03.20 fra <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>

Ireland, C.(2012) ‘Alan Turing at 100’ *The Harvard Gazette* Hentet 05.06.20 fra <https://news.harvard.edu/gazette/story/2012/09/alan-turing-at-100/>

Larsen, T. (2016). ‘Den digitale transformasjonen av NAV — drivere, utfordringer og erfaringer.’ Hentet 05.06.20, fra [https://www.uio.no/om/samarbeid/samfunn-og-naringsliv/partnerforum/arrangementer/nettverk/digitalisering/2016/larsen\\_nav\\_20160823.pdf](https://www.uio.no/om/samarbeid/samfunn-og-naringsliv/partnerforum/arrangementer/nettverk/digitalisering/2016/larsen_nav_20160823.pdf)

Luger, E. & Sellen, A. (2016). ‘Like Having a Really Bad PA: The Gulf Between User Expectation and Experience of Conversational Agents.’ I Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (s. 5286–5297). CHI ’16. doi:10.1145/2858036.2858288

Maass, S., Rommes, E.: *Uncovering the Invisible: Gender-Sensitive Analysis of Call Center Work and Software, Gender Designs IT: Construction and Deconstruction of Information*

Mugås, H. O. (2019) ‘Foreldrepengeprosjektet i NAV vant Digitaliseringsprisen 2019’ *MEMU* Hentet 10.12.19 fra <https://memu.no/artikler/foreldrepengeprosjektet-i-nav-vant-digitaliseringsprisen-2019/>

Myers, M. D. “Qualitative Research in Information Systems,” (1997) *MIS Quarterly* (21:2), June 1997, pp. 241-242. *MISQ Discovery*, archival version, June 1997, <http://www.misq.org/supplements/>. *Association for Information Systems (AISWorld) Section on Qualitative Research in Information Systems*, updated version, last modified: January 15, 2020 Hentet 05.06.20 fra: [www.qual.auckland.ac.nz](http://www.qual.auckland.ac.nz)

NAV(2018) ‘Digitalisering i NAV’ Hentet fra <https://www.fylkesmannen.no/globalassets/fm-nordland/dokument-fimno/kurs-ogkonferanser/kommunekonferansen/2019/digitalisering-i-nav---simonsen--nav.ppsx>

NAV (2020) Fakta om NAV. Hentet 12.01.20 fra <https://www.nav.no/no/nav-og-samfunn/om-nav/fakta-om-nav/kva-er-nav>

Nguyen, T.H. (2019): «Jeg blir så sliten i tankene». *Innvandrerens bruk av offentlige nettbaserte tjenester*. Master thesis, Dept. of informatics, University of Oslo



NSD (2020a) 'Våre brukere' NSD Hentet 27.05.20 fra <https://nsd.no/personvernombud/brukere.html>

NSD (2020b) 'Samtykke' Hentet 27.05.2020 fra <https://nsd.no/personvernombud/hjelp/samtykke.html>

Olsen, K. 'Roboter som ikke fungerer' (2018) *Kommunal Rapport* Publisert 26.10 2018 Hentet fra 04.04.19 fra: <https://kommunal-rapport.no/meninger/debatt/2018/10/roboter-som-ikke-fungerer>

Regjeringen (2016) 'Digital agenda for Norge' Regjeringen. Hentet 21.01.20 fra <https://www.regjeringen.no/no/aktuelt/digital-agenda-for-norge--ikt-for-en-enklere-hverdag/id2484184/>

Regjeringen(2018) 'Ny personopplysningslov og EUs personvernforordning' Hentet 28.05.20 fra <https://www.regjeringen.no/no/tema/lov-og-rett/innsikt/ny-personopplysningslov/id2592984/>

Ringnes, I. F. (2018). 'Brukerne strever med nav.no'. Hentet 27.05.20, fra <https://memu.no/artikler/brukerne-strever-med-nav-no/>

Simonsen, L.(2019) *Når brukerdialogen automatiseres – hva blir vanskelig? En kvalitativ studie av sekvensen av handlinger mellom menneske og chatbot*. Masteroppgave, Institutt for Informatikk, Universitet i Oslo

Skaarup, S. (2012) "The Mediation of Authority," , EGPA 2012 Conference Det Samfunnsvitenskapelige Fakultet. Syddansk Universitet

Strzelecki, T (2018) 'Chatboten kan snakke med 3000 samtidig' DNB Nyheter Publisert 09.09.18 Hentet fra 04.04.19 fra: <https://www.dnbnyheter.no/privatokonomi/chatboten-kan-snakke-med-3000-samtidig/>

Suchman, L.(2007) *Human-machine reconfigurations: plans and situated actions* New York, NY, USA: Cambridge University Press.

Thakur, A. (2018). How virtual agents work and why you should care. Hentet 04.05.2020, fra <https://www.boost.ai/articles/how-chatbots-work-and-why-you-should-care>

Teknologirådet(2018) Kunstig intelligens Muligheter, utfordringer og en plan for Norge. Rapport. Hentet 20.05.20 fra <https://teknologiradet.no/wp-content/uploads/sites/105/2018/09/Rapport-Kunstig-intelligens-og-maskinlaering-til-nett.pdf>

Tidemann, A. (2018) Dype nevrane nettverk/Dyp læring. Store norske leksikon Hentet 23.03.19 fra [https://snl.no/dyp\\_l%C3%A6ring](https://snl.no/dyp_l%C3%A6ring)

Tidemann, A. Elster, A. (2019) Maskinlæring. Store norske leksikon. Hentet 23.03.20 fra <https://snl.no/maskinl%C3%A6ring>

Tidemann, A. (2020) Kunstig intelligens. Store norske leksikon. Hentet 23.03.20 fra [https://snl.no/kunstig\\_intelligens](https://snl.no/kunstig_intelligens)

Turing, A. (1950) *COMPUTING MACHINERY AND INTELLIGENCE* Hentet fra:  
<https://www.csee.umbc.edu/courses/471/papers/turing.pdf>

Verne, G.:(2014) ‘Two faces of autonomy: Learning from non-users of an e-service’ In:  
*System, signs and actions*, 8 (1).

Verne, G. (2015). «The winners are those who have used the old paper form». On citizens and automated public services (Doktoravhandling, Universitetet i Oslo). Hentet fra  
<https://www.duo.uio.no/handle/10852/46133>

Verne, G. & Bratteteig, T. (2016). Do-it-yourself Services and Work-like Chores: On Civic Duties and Digital Public Services. *Personal Ubiquitous Comput.* 20(4), 517–532.  
doi:10.1007/s00779-016-0936-6

Vågeng, S. (2018) ‘NAV leder an i offentlig digitalisering’ Hentet 21.jan 2020 fra  
<https://memu.no/innspill/nav-leder-an-i-offentlig-digitalisering/>

Weizenbaum, J. (1966). ELIZA---a computer program for the study of natural language communication between man and machine. *Communications of the ACM* , 9 (1), 36-45. [http://www.universelle-automation.de/1966\\_Boston.pdf](http://www.universelle-automation.de/1966_Boston.pdf)

# Vedlegg

## Vedlegg A

### **Samtykkeskjema**

I dette vedlegget vil ett av samtykkeskjemaene som ble benyttet i denne studien bli presentert. Dette samtykkeskjemaet ble benyttet under intervju og observasjon av ansatte i NAV.

## Vil du delta i forskningsprosjektet

### *”Undersøke bruk av chatbot i NAV”?*

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke bruk av chatbot i Arbeids- og velferdsforvaltningen (NAV). I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

#### **Formål**

Formålet med forskningsprosjektet er å undersøke bruk av chatbot i NAV med utgangspunkt i chatbot-loggene. I forbindelse med mine undersøkelser ønsker jeg å intervju og observere deg. Data som samles inn vil bli brukt til å analysere chatboten, og vil ikke selv bli gjenstand for analyse. Det kan også bli aktuelt at innsamlede data vil ligge til grunn for alternative design for elementer i chatboten. Jeg er student ved Institutt for Informatikk ved Universitetet i Oslo. Innsamlede data vil kun bli brukt i forbindelse med gjennomføring av min masteroppgave og til å utvikle eventuelle alternative design for bruk av chatbot hos NAV.

#### **Hvem er ansvarlig for forskningsprosjektet?**

Veileder for prosjektet er Tone Bratteteig, professor ved Institutt for Informatikk, Universitetet i Oslo. Tina Steinthø, masterstudent ved Universitetet i Oslo, gjennomfører prosjektet.

#### **Hvorfor får du spørsmål om å delta?**

Du får spørsmål om å delta fordi du er ansatt hos NAV og bruker systemer vi vil se nærmere på.

#### **Hva innebærer det for deg å delta?**

Å delta i dette forskningsprosjektet innebærer at du deltar i ett eller flere semi-strukturerte intervjuer, eller/og bli observert av meg gjennom din interaksjon med NAV sine systemer. Under intervjuene vil jeg ta notater og gjøre lydopptak ved bruk av en digital lydopptaker. Du vil ikke bli spurt om navn eller andre personopplysninger under intervjuet. Under observasjonen vil jeg ta notater og bilder for å dokumentere aktiviteten. Hverken ditt ansikt eller andre identifiserende elementer vil bli tatt bilde av.

#### **Det er frivillig å delta**

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

#### **Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger**

Vi vil kun bruke opplysningene om deg til de formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Kun jeg og min veileder vil ha tilgang til dine personopplysninger. Du vil være anonym både i notater og i den ferdigstilte masteroppgaven. Innsamlet datamateriale vil bli lagret på Universitetet i Oslo sine servere.

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Forskningsprosjektet skal etter planen avsluttes i løpet av juni 2020. Etter prosjektets slutt vil alt råmateriale bli slettet. Opplysningene fra forskningsprosjektet kan brukes til å forske videre på bruk av chatboter og digitalisering. De kan også brukes av NAV til å designe nye chatbot-løsninger.

### **Dine rettigheter**

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

### **Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke. Studiet er meldt til NSD - Norsk senter for forskningsdata AS.

### **Hvor kan jeg finne ut mer?**

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Institutt for Informatikk, Universitetet i Oslo ved Tone Bratteteig, på epost [redacted]
- Tina Steinstø, på epost ([redacted])
- Vårt personvernombud: Maren Magnus Voll, på epost ([personvernombud@uio.no](mailto:personvernombud@uio.no)).
- NSD – Norsk senter for forskningsdata AS, på epost ([personverntjenester@nsd.no](mailto:personverntjenester@nsd.no)) eller telefon: 55 58 21 17.

Med vennlig hilsen

Prosjektansvarlig  
(Forsker/veileder)

Student

---

## Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Undersøke bruk av chatbot i NAV*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i intervju
- å delta i observasjon

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. 19. juni 2020.

---

(Signert av prosjektdeltaker, dato)

## Vedlegg B

### **Vurdering fra NSD**

Dette vedlegget inneholder NSD sin vurdering av behandlingen av personopplysninger i prosjektet. NSD vurderte at behandlingen av personopplysninger ville være i samsvar med personvernlovgivningen.

## **NSD sin vurdering**

### **Prosjekttittel**

Bruk av Chatbot hos NAV

### **Referansenummer**

554735

### **Registrert**

09.09.2019 av Tina Steinstø - tinaste@uio.no

### **Behandlingsansvarlig institusjon**

Universitetet i Oslo / Det matematisk-naturvitenskapelige fakultet / Institutt for informatikk

### **Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)**

Tone Bratteteig, [REDACTED]

### **Type prosjekt**

Studentprosjekt, masterstudium

### **Kontaktinformasjon, student**

Tina Steinstø, [REDACTED]

### **Prosjektperiode**

13.09.2019 - 19.06.2020

### **Status**

16.09.2019 - Vurdert

### **Vurdering (1)**

#### **16.09.2019 - Vurdert**

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 16.09.2019, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

### **MELD VESENTLIGE ENDRINGER**

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du



melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

[https://nsd.no/personvernombud/meld\\_prosjekt/meld\\_endringer.html](https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html)

Du må vente på svar fra NSD før endringen gjennomføres.

## TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 19.06.2020.

## LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

## PERSONVERNPRINSIPPER

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

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

## DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

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

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

#### FØLG DIN INSTITUSJONS RETNINGSLINJER

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

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

#### OPPFØLGING AV PROSJEKTET

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

Lykke til med prosjektet!

Kontaktperson hos NSD: Jørgen Wincentsen

Tlf. Personverntjenester: 55 58 21 17 (tast 1)

## Vedlegg C

### **Intervjuguidene benyttet til denne studien**

## C.1 Intervjuguide – Veiledere i NAV

Intervjuguide for intervju med veiledere i NAV Kontaktsenter Oslo og Vest-Viken.

### Oppvarmingsspørsmål:

- Hvilken bakgrunn har du?
- Hvor lenge har du hatt denne stillingen?
- Hvordan ser en vanlig arbeidsdag ut for deg?

### Hoveddel:

#### Manuelle kanaler vs. chatbot

- Hvem er det som kontakter dere på chat og telefon?
- Hvorfor tar brukere kontakt via manuelle kanaler?
- Hva slags type spørsmål får dere på chat sammenliknet med på telefon?
- Hva slags type hjelp gir dere over telefon sammenliknet med chat?
- Er det mange som lurere på det samme?
- Hva er det brukerne synes er vanskelig? Finnes ofte svaret på nettsidene deres?
- Opplever du at brukerne gjerne ønsker å finne informasjon på egenhånd?
- Er det mange spørsmål dere får via manuelle kanaler som Frida egentlig kunne ha svart på?
- Har antall og type henvendelser på chat og telefon endret seg etter at Frida ble lansert?
- Er det noen type spørsmål du selv synes det kan være vanskelig å besvare, og har du evt. mulighet til å overføre slike samtaler til andre?

#### Språk og språkbruk

- Hvilke språk bruker du på jobb?
- Får du annerledes tilbakemelding fra brukere som ikke kan så godt norsk sammenliknet med norskspråklige brukere?
- Hvordan språkbruk benytter brukerne for å chatte med Frida sammenliknet med det de bruker når de chatter med deg?
- Hvordan omtaler brukerne Frida? Skjønner de at de snakker med en robot?

#### Overgang

- Hva pleier å være grunnen til at chatten blir satt over fra Frida til en veileder?
- Opplever dere at mange vil bli satt over til å chatte med et menneske etter at Frida allerede har gitt dem svar?
- Får du chatten med Frida før du tar over samtalen?
- Hvordan synes du om samspillet mellom hjelp fra maskin og hjelp fra veileder fungerer i dag? Får du noen tilbakemelding fra brukerne vedrørende dette?

#### Utvikling

- Har du vært med på å lage Frida?
- Vet du hvorfor Frida har blitt utviklet?
- Er du med på å forbedre Frida?
- Hva slags kontakt har du med de som jobber med å videreutvikle Frida?
- Hvilke muligheter/begrensninger ser du i denne teknologien?

### Kanalbruk og tilgang til samtaler

- Kan du fortelle litt om de ulike kanalene dere har og hvordan disse brukes?
- Lager dere telefonreferater etter samtaler med brukere? Tar dere opp telefonsamtaler?
- Kan brukere som ringer inn få tilgang til telefonreferater fra samtaler de har hatt med deg?
- Kan brukere laste ned sine chat-samtaler med Frida?
- Kan brukere laste ned sine chat-samtaler med deg?
- Kan du si noe om hva som pleier å være utslagsgivende for ulike brukeres kanalvalg?
- Kan du si noen om hvilke brukere som foretrekker telefon og hvilke brukere som foretrekker chat?

### Analyse

- Vet du om NAV analyserer data fra manuell chat og telefonsamtaler til å forbedre Frida?

Vet du om NAV bruker data fra loggene til Frida, manuell chat eller telefon til å forbedre NAV sine nettsider, digitale søknader eller annet? *Vet du hvem som gjør dette og/eller hvordan dette gjøres det?*

### Tilbakemelding

- Gir brukere tilbakemeldinger på Frida? Hvordan gir de tilbakemeldinger? Hva sier de?
- Får du annerledes tilbakemelding fra brukere som ikke kan så godt norsk sammenliknet med de som kan godt norsk

### **Avslutning:**

- Kan du fortelle litt om dine tanker rundt bruken av Frida?
- Kan du fortelle litt om hvordan du ser for deg NAV om 5 år?
- Er det noe du ønsker å legge til avslutningsvis?

## C.2 Intervjuguide – Chatbot-trenere i NAV

Intervjuguide for intervju med chatbot-trenere i NAV Kontaktsenter Vestland

### Oppvarming

- Hvilken bakgrunn har du?  
- Har du jobbet som veileder tidligere/i tillegg?
- Hvordan ser en vanlig arbeidsdag ut?
- Hva legger dere i 'trening'?
- Hva er formålet med trening? (Er det et overordnet mål?)
- Hva er formålet med Frida? (Ambisjonsnivå?)

### Utvikling av Frida

- Hadde dere pilot først?
- Kontinuerlig oppdatering eller versjoner?
- Hvordan utvikler dere trestrukturen?
- Hva må du tenke på når du utformer Fridas svar?
- Er teksten hentet fra FAQ?
- Gir Frida annen informasjon enn det som står på nettsiden deres?
- Hva ser dere på som en vellykket samtale?
- Har dere kategorier av mislykkede samtaler?
- Hvordan analyserer Frida teksten?
- Oversetter Frida fra norsk til engelsk før setningen analyseres?
- Hvordan bestemmes det hvilke ord som skal ha og ikke ha betydning?
- Måler dere effekten av de endringene dere gjør? (Færre som snakker med en veileder via chat etc.)
- Hva er dine tanker vedrørende Fridas skalerbarhet? Ser du noen utfordringer ved å utvide Frida for mye?
- Hva synes du er utfordrende i jobben som AI-trener?
- Er det noen etiske aspekter du ønsker å trekke fram?

### Bruk av data/Tilbakemeldinger

- Hvordan får dere vanligvis tilbakemeldinger? Hvordan bestemmer dere hva dere skal prioritere?
- Hvordan "lærer" Frida?
- Hvilke datakilder blir benyttet i utviklingen av Frida?
- Bruker du chatlogger mellom brukere og veiledere i utviklingen av Frida? Hvorfor / hvorfor ikke?
- Kan du fortelle litt om hvordan Yammer-kanalen brukes?
- I hvilken grad er brukere eller har brukerne vært involvert i utviklingen av Frida?
- Så vidt vi har forstått, så skal brukere snart igjen kunne gi Fridas svar tommel opp eller tommel ned. Kan du fortelle litt om dette? Hvorfor ønsker dere å implementere dette?
- Fører dere logg over endringer?

### Ordet 'Ikke'

(Diskutere konkrete eksempler)

- Hvordan jobber dere med dette?
- Finnes det situasjoner der Frida klarer å forstå ordet "ikke", og andre evt. andre situasjoner der Frida ikke klarer det?

### Utfordringer

- Hva er vanskelig med tanke på utvikling av chatboter?
- Er brukere som spør på vegne av andre enn seg selv en utfordring?
- Er det utfordrende med endringer i regelverk, endring i hva barnetrygden er på, osv.?
- Hvordan jobbes det med falske positive og falske negative?

### Bruk av plattformer

- Har dere oversikt over hvor mange som bruker Frida på forskjellige plattformer?
- Hvorfor er den ikke bedre tilpasset smarttelefoner?

### Planer for bruk av chatbot hos NAV

- Hvilken visjon har dere for chatboter i NAV?
- Hvilken ny bruk av chatbot-teknologi jobber dere med?? Hvilke områder av NAV, og hvorfor?
- Er det ønskelig at Frida etter hvert skal kunne overta alle chatter eller skal det fortsette å være en kombinasjon av chatbot og menneskelig chat?
- Hvor mye tid bruker dere på å forbedre Frida?
- Hvordan føler du at bruk av Frida fungerer?

### Bruk av kunstig intelligente systemer

- Bruker dere sentimentanalyse? Hvorfor og hvordan brukes dette?  
(Brukes dette f.eks. til å identifisere humøret til brukerne, for deretter å foreslå om de skal bli satt over til en veileder eller ikke?)
- Hvilket potensial ser dere i teknologien? Kan Frida videreutvikles for å utnytte teknologien ytterligere?

### Avslutning

- Er det noe du ønsker å legge til avslutningsvis?
- Er det noe vi ikke har tatt opp?

## C.2 Intervjuguide – Digital arkitekt i NetNordic

Intervjuguide for intervju med digital arkitekt i NetNordic

### Oppvarming

- Kan du si noe om hvilken stilling du har og hvor lenge du har jobbet her?

### Hoveddel

- Kan du fortelle oss litt om NetNordic, samt sette NetNordic i sammenheng med Boost.ai, Puzzel og NAV? Hvem har ansvar for hva?
- Chatbottrenerne fortalte oss at det var noen studenter som først jobbet med NAVs kontaktsenter-chatbot, Frida. Vet du noe mer om dette? Hvordan og når ble dere involvert i dette arbeidet?
- På hvilken måte tilpasser Boost.ai sin løsning til ulike bransjer? Hva er det Boost.ai leverer og hva er det kunder som NAV selv må gjøre?
- Vet du hvorfor NAV valgte å utvikle en chatbot basert på teknologi fra Boost.ai?
- Ser du noen utfordringer eller muligheter ved å kombinere menneskelig chat og chatbot slik som NAV gjør i sin løsning?
- Vi har forstått at NAV jobber med å lansere en ny versjon av Frida. Er dere involvert i dette? Hvis ja, hvilken rolle har NetNordic i denne prosessen?

### Bruk av chatboter

- Det er flere utfordringer ved bruk av chatbot. Hva tenker du at det er viktig å fokusere på? Bør det f.eks. være mer fokus på å bedre tilpasse chatboten? Mer opplæring av brukere om hvordan chatboten fungerer? Er det noe som er spesielt utfordrende i samspillet mellom teknologi (Boost.ai), implementering (AI-trenere) og bruk (NAV-brukere)? Ser du f.eks. noen utfordringer med at ikke-teknologer trener chatboten eller at brukere ikke forstår hvordan chatboten analyserer setninger?
- NAV har nylig introdusert Frida i AAP-chatten og planlegger å utvide bruk av chatbot ytterligere. Ser du noen utfordringer knyttet til chatbotens skalerbarhet? Hvordan kan disse evt. løses? Eksempelvis lage flere chatboter som kan sette brukere over seg imellom.

### Eksempler

- Kan du se litt på disse eksemplene og si hva du tenker?
- Kan du foreslå noe som kan gjøres for å forbedre interaksjonen mellom chatbot og menneske i disse eksemplene?
- Har dere noe statistikk over chatbot-bruk i NAV?
- Vi skal besøke Skatteetaten for å få en demo av deres chatbot-løsning som også er basert på teknologien til Boost.ai. Er dere involvert i dette? Ser du en fordel i at flere offentlige instanser benytter seg av samme chatbot-leverandør?

### Avslutning

- Hvilke tanker gjør du deg om bruk av chatboter i offentlig sektor? Hvilket potensial ser du i teknologien?
- Er det noe du ønsker å legge til avslutningsvis? Er det noe vi ikke har tatt opp?



## Vedlegg D

### **Artikkel**

‘ “I’m disabled and married to a foreign single mother”  
Public service chatbot’s advice on citizens’ complex  
lives’

Skrevet av Linett Simonsen, Tina Steinstø, Guri Verne  
og Tone Bratteteig

# “I’m disabled and married to a foreign single mother”. Public service chatbot’s advice on citizens’ complex lives

Linett Simonsen<sup>1</sup>, Tina Steinstø<sup>1</sup>, Guri Verne<sup>1</sup> and Tone Bratteteig<sup>1</sup>

<sup>1</sup>University of Oslo, Gaustadalléen 23B, 0373 Oslo, Norway  
{linetts, tinaste, guribv, tone}@ifi.uio.no

**Abstract.** This paper describes a study of citizens’ chats with a chatbot of a public agency. We have analyzed chat logs and identified citizens’ lack of domain knowledge as a source of inadequate or failed chatbot responses. We identify three types of lack of domain knowledge: lack of the right vocabulary, uncertainty if a regulation fits the citizen’s situation or the “shape sorting box” problem, or citizen’s misunderstanding the regulations. The most serious failure is when a misunderstanding is not detected and corrected during the chat. The chatbot we studied is not able to make sense of badly formed questions from citizens. As implications for design we suggest making the chatbot limitations visible by not presenting it as a human-like avatar with a name. We also suggest to enable domain knowledge learning through its conversations.

**Keywords:** chatbot, domain knowledge, digitalization.

## 1 Introduction

Governmental digitalization includes the digitalization of public administration as well as the digitalization of public services, which is where the citizens meet this change. Digitalized public services are often self-service solutions where the citizens have to carry out parts of the service, hence they need to know how to use the self-service solution. Use of electronic services could decrease the need for “administrative literacy” but will increase the need for computer skills [1]. The government in Norway is therefore concerned with the “digital competence” of the population since low or no digital competence in (parts of) the population may make the digital public services inaccessible to these citizens [2]. A digital divide rooted in inaccessible digital services may have severe socio-economic consequences for the affected individuals and for society [3]. Our concern in this paper is grounded in recent studies claiming that a major reason for unsuccessful or failed use of public self-services is the lack of domain knowledge [4-7]. These studies emphasize that an important service from public call centers and staffed offices is the ability of public administration advisors to “translate” questions from citizens into the right categories for the public “machinery” to process their inquiry. Being able to use the self-services requires some understanding of the public service system [5]. In this paper we explore how a digitalized public service handles this problem by reporting from an empirical study of a particular public service where a chatbot has been designed to be the first public encounter for a citizen who seeks information and help [8]. The study gives an in-depth description of how the chatbot functions in practice through an analysis of a set of real chatbot

conversations. The paper identifies problems occurring in practice that can act as a basis for designing improved digitalized public services.

The paper first presents the public service arrangements in Norway as well as the particular service where the chatbot has been introduced. We add a brief section about chatbots in general and this one in particular before we go on to tell about our study. We have analyzed logs from chatbot chats with citizens, and the next section describes three chats each illustrating a different type of problem concerned with lack of domain knowledge. The last section discusses what sort of domain knowledge the inquiries from the public may require; our analysis of the logs is confirmed by interview data. We also discuss how chatbots may handle the problems we have identified and hint to some implications for design. The last section concludes the paper.

## **2 Background: Digital public services in the welfare state**

The public service we have studied is a part of a welfare state, which offers a range of benefits and services for citizens in various life situations, e.g., retirement pension, unemployment allowance, healthcare services and sick leave benefits, a right to education, and various forms of family benefits. The benefits are allowances to the citizen based on legally grounded arrangements. In this paper, we concentrate on family benefits, because this is the services that the chatbot we have studied is made for. If a child is born in Norway, the mother of the child will automatically receive a birth allowance. Approx. two months after the child is born, she will start receiving a monthly child benefit (until the child is 18 years old). Single mothers and fathers, who live alone with children, may also be entitled to extended child benefit and infant supplement.

### **2.1 The public service**

It is the Norway's Labour and Welfare Administration (LWA) that manages the services and benefits on behalf of the welfare state, altogether approx. 60 services and benefits. LWA meets the citizens through their 456 local offices, their Contact Center (CC) or through their digital services. The CC covers all services for the whole country and is responsible for contact with citizens through telephone, chat, and social media presence. Every year the center receives about 4.500.000 telephone calls, 700.000 written documents via net-based services (logged-in), and approx. 140.000 inquiries via chat and Facebook. The CC is not involved in the LWA's case handling: their employees only answer questions about rules and regulations, case handling response times, and payment dates. In addition, they can inform the citizen about status in her/his own case and teach citizens to use the web pages and self-service solutions.

As a part of its digitalization the LWA has expanded its digital communication and reduced the opening hours of the local offices. A chatbot has been introduced in the handling of inquiries about family benefits. A citizen who initiates a chat with LWA will first meet the

chatbot before the chat can be transferred to a chat with a human. The chatbot is based on Machine Learning technology (ML) and LWA hopes that this technology combined with the large amounts of data captured by the public sector may enable automated services offering just-in-time welfare services to citizens without their active participation. The current chatbot is, however, not personal in that it does not require the citizen to log in or give any personal information.

Chatbots are expected to reduce the load on manned services, in particular calls to the call centers, by taking over the task of responding to questions from the citizens. Studying use of ML in the public sector is important for understanding how such technology functions in the public encounter [8].

## **2.2 Digitalization of the public encounter**

The case handler or advisor, the street level bureaucrat who communicates directly with a citizen, represents the welfare state towards the citizens [9]. The digitalization of civil servants' work has implied that their contact with citizens and clients have changed from the "street level" contact [9], where they meet in person, to the "screen level", where data is entered into a screen, to the "system level", where decisions are made by a computer [10]. Several studies on channel choices and how they affect the relationship between the citizens and the government have been reported. Both Skaarup's study [11] of how different channels affect the citizens' feeling of control, and Ebbers et al.'s study [12] of how different channels are suitable for different purposes, show that the types of channel available can make a significant difference to the citizens. However, clients are increasingly being invited or forced to communicate with their advisors through online channels, but they also continue to use the traditional channels after adopting new digital service channels [13]. As digital technologies are used for larger parts of the communication between the client and the agency, new challenges arise. New types of skills are required from the citizen when manual services are replaced by electronic services.

There are a number of studies about what frontline service workers do and know, and many of them report that frontline service workers like the "street-level bureaucrats" [9] and the call center operators are good at "translating" callers' inquiries to internally specific categories and formats [14-16]. Several studies report that call center operators carry out knowledgeable and even emotional work in their response to a caller [17-23]. Flexible communication and emotional work are important for the operators to create a good interaction with the caller [17] and the ability to read a caller's emotional state can also be used for interpreting the emergency of the call [21].

There is a large variety of problems and issues that lead a citizen to call a public agency. Earlier studies have shown that many callers know the rules and regulations but ask if the rules apply to their particular situation; what Verne [6] calls "the shapessorting box", i.e., fitting a real-life problem into one of the categories formed by the rules and regulations of

the public agency. The callers need help to interpret the rules and understand how they fit their own life situation [6]. However, many callers show very little domain knowledge [5, 6]. The call advisors often need to disentangle the callers' problems into smaller steps that the callers can take themselves [24]. A study of call center operators in a public agency shows that the operators responded differently to different callers depending on how they interpreted the caller's digital competence and domain knowledge: A caller appearing helpless received more guidance and practical help than callers who expressed themselves in more competent ways [25].

### **3 Chatbots**

A chatbot is an automatic system to which a user can text or "chat" in real time over the Internet [26]. Today, chatbots are able to communicate with people in natural language [27]. From the very first chatbot Eliza [28, 29], most chatbots have a limited vocabulary and repertoire based on preset keywords. The Eliza program first analyzed the syntax of the user's statement then formed a response from simple pattern recognition in addition to replacing keywords. Even if chatbots are able to "seemingly reply in intelligent ways to many queries, they still cannot understand a basic sentence in natural language" [30, 31]. Moreover, Hill et al [31] found that users adapt their language by using simpler messages when interacting with a chatbot: the users they studied were very aware that they were communicating with a robot.

Today most chatbots are based on Machine Learning (ML), which is a special area within AI where statistical methods are used for identifying patterns in large data sets, to recognise patterns and produce results [32, 33]. ML is different from traditional programming where the machine performs predetermined operations. By being exposed to a number of examples or tasks with given answers, i.e., correct data, the machine is "trained". "Learning" is central in ML: the machine has to "be trained" by a set of training data and tested with a different set of test data [30, 32, 34]. The quality of the ML depends of the quality of the training data and the test data as well as how relevant these data sets are for the real data that the ML will meet when in operation. The algorithms will in principle give better results as the ML learns by processing more data. However, the development of a ML-based system like a chatbot needs to be monitored by people. A well-known example of a chatbot failure is the Microsoft chatbot Tay, which was designed to "learn" through dialogues with Twitter users and did not have security mechanisms to avoid "false teaching". Within 24 hours Tay got out of control and had to be removed from the open net due to "trolling" from the Twitter users [35].

#### **3.1 The LWA chatbot**

The LWA chatbot is named Anna (a pseudonym). Anna went public in the fall of 2018 and answered chats from approximately 400 users each day in the beginning of 2019. 40 % of the users got an answer in this first encounter, 40 % were transferred to chat with a human

advisor, and 20 % went to another channel (e.g., telephone) to communicate with LWA. Anna is not expected to respond to very complicated questions, as neither the technology nor the users are considered “good enough” yet.

Anna is based on ML technology and the aim is that it will learn and improve its responses through interactions with users. The ML analyses the users’ questions to recognize what they are asking for. Once identified, the ML selects the corresponding response. ML is only used to identify questions: the answer the users receive to a question has been designed to adhere to the laws and regulations for that question, theme or keyword. The concrete formulation of the answers is designed by the advisors, who know what kind of replies the callers need to receive on particular questions.

In response to a question from a user, Anna pre-processes the text and classifies it as belonging to a category of questions, where all questions will match a predefined general answer. The ML performs several steps starting with correcting spelling errors and wrong words used. Then the ML identifies the question by trying to identify what the user wants to know by asking this question: the “intention” of this input. Based on the input the ML calculates which intention seems most probable. The system will match this intention with the set of predefined intentions. If the system finds the user’s intention to be similar enough to a predefined intention, the predefined answer to this intention (i.e., question) will be given as the chat response. If the system does not find a sufficiently similar predefined intention, a standard fall-back message will be given, such as “Please rephrase your question in simpler terms”.

Anna is constantly trained. LWA’s chatbot trainers are constantly monitoring Anna, using reports generated by Anna in addition to reviewing chat conversations and feedback on Anna’s performance from advisors. The chatbot trainers create training data based on this and run training sessions to improve Anna’s responses. Training the chatbot means that more questions are classified, and new answers are added to the vocabulary of the system.

#### **4 Method: studying chatbot logs**

The main material for this interpretive qualitative case study [36] is a segment of chat logs retrieved from chats between users and the LWA’s chatbot Anna from more than 6 months in 2019, each week including approx. 3000 chats. For this study we extracted approximately 8000 unique chat logs from four of these weeks: two weeks in May and two weeks in September (more detailed studies in [37, 38]). We studied documents: LWA’s plans and reports as well as the chatbot developers’ documentation, and we interviewed and observed advisors and chatbot trainers at two CC’s units. We carried out four semi-structured interviews with three different advisors in addition to observing them as they chatted with citizens. This gave us insight into typical questions and issues that the citizens ask about. We had three semi-structured interviews with two chatbot trainers working with Anna improvements, and observed them while they worked.

We took a hermeneutic approach [36] to analyzing the data, focusing on making sense of the log data. We used Suchman's analytical framework [39] as a starting point for interpreting and understanding chat logs where difficulties had occurred. The framework (see Figure 1) distinguishes between the user's "actions available to the machine" (i.e. input) and actions not available to the machine, and the machine's response ("effects available to the user") as well as the machine's rationale or inner workings. In this paper we are concerned with the machine's responses to the users' actions and how Suchman's framework allows us to include the user's life circumstances, both those available and not available to the machine.

THE USER		THE MACHINE	
I	II	III	IV
Actions not available to the machine	Actions available to the machine	Effects available to the user	Rationale

Fig. 1. Suchman's analytical framework for studying and analyzing human-machine interaction [39].

We started by analyzing each chat to see if it was successful or not. We followed up the unsuccessful ones with a more detailed analysis of the questions and responses where we coded the challenges they illustrated into different categories. In this analysis we looked for communication breakdowns and used this as a basis for understanding how the chatbot worked. Examples that we considered difficult were discussed with chatbot trainers and advisors in the LWA. In the examples given below, the questions and answers are translated by us, taking care to preserve the nuances of the legal terminology and the user's original (mis)spelling.

## 5 Chatting with Anna

In the material we have analyzed we have found many chat logs that appear to have been successful: The user does not follow up with more questions, and our interpretation of Anna's answer indicates that it is correct and relevant. LWA's criteria for a successful answer from Anna is a "sufficient" answer, where the chatbot gives relevant and correct information as a response to the user input.

Some of the chat logs indicate an unsuccessful chat session. In some obvious cases, the user leaves the chat after swearing or expressing dissatisfaction with Anna's responses, e.g., by writing "This is hopeless" before leaving. In other chats, the user does not express dissatisfaction but leaves the chat abruptly. In this section we present and discuss three examples of chat conversations that illustrate different challenges to the use of chatbots in a legal setting. We focus on lack of domain knowledge as a reason for lack of success, i.e., chat conversations where a user applies the wrong word or inconsistent terms or shows lack

of knowledge about the rules and regulations. In the following three subsections, we present different chat sessions that illustrate how weak domain knowledge of the user leads to challenges in getting good answers.

### 5.1 Lack of key concept

In this category, the user formulates a question where the wrong terms are used. In this example the user asks a question about the amount of child benefit and writes the word “*child bennefit*” (translated from Norwegian with the misspelling preserved). Despite the spelling mistake, Anna recognizes that the question is about child benefits. Anna does, however, not recognize that the user asks about the amount.

**Anna:** My name is Anna and I am a chatbot. **Anna:** I am best adapted to answering short questions. Remember to include what type of benefit you are asking about.  
**Anna:** How can I help you?  
**Anna:** For English chat, click here (*button*)  
**User:** What is the child bennefit\* this year?  
**Anna:** If you have a child under the age of 18 living with you in Norway, you may be entitled to child benefit.  
**Anna:** You can receive child benefit from the month after the child is born.  
**Anna:** If you are entitled to child benefit at a later point, you may receive child benefit from the month after the terms have been met.  
**Anna:** Child benefit will be paid up until the month before the child turns 18. If the right to child benefit ceases earlier than this, you will receive child benefit up until the month the terms ceases to apply.  
 • Who is entitled to child benefit? (*button*)  
 • Extended child benefit (*button*)

- Payment (*button*)
- Changes in family situation (*button*)
- Apply for child benefit (*button*)

**User:** What is the rate of the child benefit?  
**Anna:** Ordinary child benefit is 1054 NOK monthly per child.  
**Anna:** Extended child benefit is 1054 NOK and will be in addition to the child benefit you will receive for children living with you.  
**Anna:** Infant supplement is 660 NOK  
 • Extended child benefit (*button*)  
 • Infant supplement (*button*) **User:** Thank you, Anna  
 ☺ **Anna:** You are welcome!  
 Conversation continues ...

\*misspelled word  
 (Conversation ID: 617631 (masked code))

This user seems to understand that the first term s/he uses does not give the wanted answer and tries to rephrase the question using the term “rate”. The user gets a different answer, which s/he seems to be satisfied with.

### 5.2 The “shape-sorting” box

In this category, Anna cannot link the user’s question to the relevant services and benefits this user asks for. The result is that Anna does not recognize what the user asks for, and therefore cannot help.

**Anna:** My name is Anna and I am a chatbot  
**Anna:** I am best adapted to answering short questions. Remember to include what type of benefit you are asking about.  
**Anna:** How can I help you?  
**Anna:** For English chat, click here (*button*)  
**User:** Going to have a child, what do I need to apply for?

**Anna:** I’m sorry, but I don’t understand what you are asking.  
**Anna:** I will understand it better if you write short questions and which benefit it is regarding.  
**Anna:** Can you ask the question again?

- Need more help (*button*)
- No (*button*)



**User:** How do I receive money when I am expecting a child?

**Anna:** Which day the payment arrives, depends on which benefit (it regards).

**Anna:** You can see your payments from LWA at “Your payments” at <LWA.website>, but they do not show up before they have arrived in your account. <lists info about payments>

(Conversation ID: 679785)

The question formulated by the user in this example indicates weak domain knowledge, as the user does not know what the different services and benefits are called, or who is entitled to what. We call this a problem with “the shape-sorting box”: matching a citizen’s life situation to the various services and benefits.

The user leaves the chat without asking further questions. If this user had asked explicitly to be transferred to chat with an advisor, it would have been visible in the chat log. S/he may have changed the contact to another channel or given up without the chat log giving any indication.

### 5.3 A logical/legal flaw

In this category, the terms used indicate a weak understanding of the rules and regulations for LWA’s benefits and services. In this chat, the user asks the question “can a young disabled person married to a foreign single mother receive child benefit and child supplement”. The question contains the legally inconsistent phrase “married to a ... single mother”: the rules are that if you are married you are not categorized as a single parent. Anna's response concerns benefit for single parents, which does not match the user's stated life situation. The user follows up with asking whether a “foreign single mother” can receive child benefit and child supplement. Anna does not recognize that the user expresses a legal inconsistency and responds by providing more information about the rights of single parents

**Anna:** My name is Anna and I am a chatbot. **Anna:** I am best adapted to answering short questions.

Remember to include what type of benefit you are asking about.

**Anna:** How can I help you?

**Anna:** For English chat, click here (*button*)

**User:** can a young disabled person married to a foreign single mother receive child benefit and child supplement

**Anna:** If you are a single mother or father, you may be entitled to extended child benefit if you are living alone with children.

**Anna:** The extended child benefit is child benefit for one more child than what you actually are living together with.

**Anna:** You have to apply for an extended child benefit, it will not be automatically granted. • Who is entitled to child benefit? (*button*)

<lists info about extended benefit>

**User:** Entitled to extended child benefit? (*action*)

**Anna:** If you are a single mother or father you can receive extended benefit if you are living alone with the child and:

- you are separated, divorced or widowed

<continues to list criteria for extended benefit>

**Anna:** If you do not have common children with the person you are living with, we do not need documentation that you are not living together. It could be:

- Declaration of breakup
- Separation permit

**User:** When does the right to extended child benefit expire? (*button*)

• Apply for extended child benefit (*button*) • Shared child benefit (*button*) **User:** When does the right to extended child benefit cease? (*action*)

**Anna:** Extended child benefit will cease if:

- You marry
- You have children with a new cohabitant
- You have been cohabitant with a person you do not have joint children with for twelve months
- You move together with the father of the child

**User:** can a foreign single mother receive child benefit and child supplement

**Anna:** If you are a single mother or father, you may be entitled to extended child benefit if you are living alone with children. **Anna:** <lists criteria for extended benefit> **User:** Payment (*action*)

**Anna:** Child benefit will be paid out the last day of the week each month. Child benefit is not liable to tax.

**Anna:** Remember that you can see your payments from LWA by logging into 'Your LWA'. **Anna:** At <LWA.website> you can find an overview of payment dates for child benefit.

<lists payment info>

**User:** Can foreign single mother receive child supplement

**Anna:** In some cases, the mother may receive the entire period of child benefit

**Anna:** She can be alone with the care or live together with the child's other parent that have not yet served up the right to child benefit.

- I am caring for the child alone (*button*)
- I live together with the father/co-mother (*button*)
- I mean child benefit (*button*)
- I mean transitional benefit (*button*)

**User:** I mean child benefit (*action*)

**Anna:** If you are a single mother or father, you may be entitled to extended child benefit if you are living alone with children.

**Anna:** The extended child benefit is child benefit for one more child than what you actually are living together with.

**Anna:** You have to apply for an extended child benefit, it will not be automatically granted.

<lists info about extended benefit>

(Conversation ID: 670860)

Nowhere in this conversation does Anna explain that you cannot be categorized as a single mother if you are married. As the user continues to ask about benefits for single mothers, this misunderstanding is not detected, and Anna continues to answer about benefits for single mothers without explaining that the mother in question may not be eligible for such benefits. The user selects the button “I mean child benefit” and is given a response that Anna has already provided earlier in the conversation. Since Anna's reply contains the buttons “I mean child benefit” and “I mean transitional benefit”, it may look as if the system recognizes the user’s weak domain knowledge. However, the answers the user receives may be completely irrelevant if the mother in question actually is married. In the worst case, this user is misled by the chatbot into believing that the mother he is married to is eligible for benefits for single mothers, which she is not.

## **6 Discussion**

### **6.1 Chatbots as a frontline in public services**

Previous research about call centers has shown that the operators do more than merely answer questions about facts and information. The operators’ expertise and knowledge about their work add value to their clients’ inquiries [19]. Whalen et al. [23] describe operators’ work by the seemingly contradictory notions of improvisation and choreography, indicating that their work routines are craft-like practices where the operator adjusts his/her response to the situation-at-hand: the actual, concrete inquiry from the client. Verne’s study of call center operators in a public agency showed that operators helped callers to formulate their inquiries in the relevant and correct terms as well as providing problem-solving steps that the caller could perform to solve the problem [25].

There is no creative translation capacity in a chatbot: the chatbot only compares the words given to it through the user’s input and matches these to the predefined keywords – the only pre-processing done by the chatbot is spell check. The cases above show that the chatbot often does not forgive spelling errors or use of wrong words by clients. This makes the chatbot a very strict first encounter with the welfare agency for many citizens. The chatbot becomes a representative for LWA, hence the welfare state may appear rigid when it does not provide information that the citizen is not aware of that s/he should ask for.

### **6.2 Domain Knowledge**

Our study shows different ways that the lack of domain knowledge affects the citizens conversations with the welfare state through chatbots. We have identified three types of domain knowledge that may result in unsuccessful communication with chatbots like Anna if missing.

The first type of domain knowledge is the language: the correct terms. The user in our first case misspells a (correct) term and also uses an everyday word (“amount”) that Anna does not recognize. Lack of knowledge about the correct terms is a well-documented problem that goes

deeper as knowledge about the correct terms often depends on knowledge about the phenomenon that the terms refer to, i.e., the rules and regulations of the public service. Our interviews with advisors confirmed that many citizens contact them using everyday language. “It is easier for them to use the chat and say ‘Hi, I am new in this. And I wonder about so-and-so’. Maybe they don’t know where to start. These are the people I think contact us on the chat,” said one of the advisors. Another advisor confirmed this by stating that “My experience is that many people ask like ‘Hi, I plan to have kids, my due date is soon, my due date will be like so-and-so. What should I take care of? How do I apply?’” The tax call advisors in Verne’s study [6] recognize weak domain knowledge by wrong or inconsistent terms used by the callers. A study among immigrants in Norway showed that even if they had lived in Norway for several decades and spoke Norwegian well, the specific terms of LWA were difficult to understand and master because the terms had no equivalents in their mother tongue [40]. This also holds for native Norwegians without competence in LWA’s services: the advisors told that they “normally use an oral language style to make it easier to understand for the users” in the chats and that they “break the information down to make it more understandable”. Another advisor told us that “it is very rare that you can use the standard answers”, so the answers have to be adjusted. “I assess every chat and adjust the information not too much and not too little to the concrete chat.” Moreover, the advisors emphasized that the users ask in many different ways: “There are incredibly many nuances and incredibly many ways of solving things”.

The second type of domain knowledge also seems to be a language problem, but we categorize it as a “shape sorting box” problem. The user asks if s/he is eligible for a particular benefit. This problem is discussed by Verne [6]. Her study of the tax authorities’ call center concluded that many of the callers asked how their particular life situation fitted with the rules – even callers with good knowledge about the rules wanted to have their interpretation of the laws and regulations confirmed from the official view of the tax authorities. This was supported by the advisors: “Very often they want a confirmation because it is complicated regulations and many nuances” and “the users contact us to make sure they have done things right”. The advisors also said that “much of the information [the users ask for] you can find at the <LWA.website>, but we understand that they chat with us. Because it is a lot of information. And they have read it, but they have not really understood it properly”. They also say that “The application [for benefits] has become a lot easier, but people don’t stop contacting the LWA”. One advisor explained that “A challenge in LWA is that things overlap a lot. There are no strict borders, parental things flow into [other topics] ... There are a lot of terminologies that are similar, but not quite the same, but almost the same: child benefit, child allowance ...”. Anna is not able to answer questions about the “shape sorting box” because the chatbot only answers general questions, referring to general rules and regulations. Anna helps people find general information but cannot relate to the actual situation of a person.

The third type of domain knowledge is a misunderstanding of the rules that looks like a “shape sorting box” question. However, as the chatbot does not detect this misunderstanding, there is a danger of mis-informing the caller, who may get an unpleasant surprise if s/he has calculated with a benefit it later turns out that s/he is not entitled to. The advisors told us that

“Many people do not understand what is written at <LWA.website>”. Combined with what the advisors told about adjustment of their language to the individual user and the many nuances in both questions and answers, we find that the human chat offers what call centers normally do: “I answer on their level”.

### **6.3 Implications for design**

In this study, we have found several challenges in citizens’ chat with Anna that suggest shifting the focus in chatbot design from a focus on the chatbot end of the interaction to taking the whole chat conversation into consideration (in line with [41, 26, 43]). Our analysis of the chatbot’s responses to the users’ questions (cf. Figure 1, [39]) has of course made use of our (human) interpretation of the content of the questions and what they mean as well as the assumed value and appropriateness of the chatbot’s answers. Based on this analysis we suggest some implications for the design of the chatbot end of those conversations. As users’ lack of domain knowledge may cause miscommunication and communication breakdowns, we encourage chatbot designers to explore the possibility of helping users gain domain knowledge, for instance by providing cues or basic domain information as part of the chatbot communication.

Our analysis indicates that users’ expectations were out of step with the practical realities of chatbot capabilities, causing frustration and dissatisfaction on the user side. One reason for the high expectations may be that LWA’s chatbot Anna is presented with a human name and a human-like avatar. The chatbot logs we studied contained anthropomorphic trigger responses such as “Aww, now you are making me blush” and “I understand very well what you are saying”. Our analysis suggests that not all users understand that they communicate with a computer, and we suggest designing the chatbot in a way that does not try to impersonate a human, but clearly reveal its non-human qualities. Programmed trigger responses may set unrealistic expectations framing the ongoing user experience [17, 27, 42], hence we suggest a careful approach when implementing such features.

Reflecting on the introduction of Anna, one of the advisors say that “It has been a pronounced change – there are fewer conversations on [the human] chat but I do not know if more people get help”. As of today, LWA’s chatbot can only be used as a nonlogged in service. However, it is possible to imagine a future where users can communicate with the chatbot as logged in users. In this case the system may have access to personal information about the users, e.g., if they are married or have children. This may affect how the chatbot system will be able to help users in their particular situations. The chatbot can, e.g., check if the marital status mentioned in the question is in line with what is registered in their databases and respond appropriately.

The fact that the chat is the first encounter with the welfare state for the citizen makes it important that the chatbot’s answers are to be trusted. We saw a small sign that the trust is fragile in a chat conversation with a human, where the user asks a question and the human chat advisor recommends a form and provides the link to the form. The user responds by asking “Is the form you linked to the right one?” The advisor had to explain the reasons for

recommending this particular form. Chatbot encounters that make users lose trust in the welfare state may be a serious cost of the benefits of digitalization.

## 7 Concluding remarks

In this paper we have explored how a digitalized public service in the form of a chatbot handles citizens' inquiries about family benefits. By analyzing logs from chatbot chats with citizens we have shown that chats where citizens lack domain knowledge may result in inadequate responses by the chatbot. Lack of domain knowledge is exemplified as lack of the right vocabulary, uncertainty if a regulation fits the citizen's situation – the “shape sorting box” problem, or misunderstanding the regulations. We find that the lack of adequate response from the chatbot is particularly problematic if a misunderstanding that the citizen has is not detected and addressed during the chat. Our studies of real chats with a chatbot show that the chatbot is not able to interpret and “translate” badly formed questions from citizens into the right categories for further advancement of their inquiry. We suggest making the limitations of the chatbot's capabilities more visible by presenting it as a computer rather than a human-like avatar as well as expanding its ability to explain and teach citizens relevant domain knowledge through its conversations. Lack of domain knowledge makes the user unaware of inconsistencies and limitations in the chatbot's answers.

## Acknowledgements

We thank LWA Contact Center and its leader for giving us access to the chat logs and for allowing us to interview the advisors and chatbot trainers.

## References

1. Grönlund, Å., Hatakka, M., Ask, A.: Inclusion in the E-Service Society – Investigating Administrative Literacy Requirements for Using E-Services. In: *Electronic Government: 6th International Conference (EGOV 2007)*, 4656/2007. Springer Berlin, pp. 216–227, (2007).
2. Carretero, S., Vuorikari, R., Punie, Y.: *The Digital Competence Framework for Citizens*. Publications Office of the European Union (2017).
3. Van Deursen, A. J., Van Dijk, J.A.: The First-level Digital Divide Shifts from Inequalities in Physical Access to Inequalities in Material Access. In: *New Media & Society* 21 (2), pp. 354–75 (2019).
4. Breit, E., Salomon, R.: Making the Technological Transition – Citizens' Encounters with Digital Pension Services. In: *Social Policy & Administration*, 49, pp. 299–315 (2014).
5. Madsen, C. Ø., Kræmmergaard, P.: The efficiency of freedom: Single parents' domestication of mandatory e-government channels. In: *Government Info. Quarterly*, 32 (4), pp. 380-388 (2015).
6. Verne, G.: The winners are those who have used the old paper form. On citizens and automated public services. Doctoral dissertation, Dept. of Informatics, Univ. of Oslo, (2015).
7. Verne, G., Bratteteig, T.: Do-it-yourself services and work-like chores: on civic duties and digital public services. In: *Personal Ubiquitous Comput.*, 20 (4), pp. 517–532 (2016)
8. Lindgren, I., Madsen C.Ø., Hofmann, S., Melin, U.: Close Encounters of the Digital Kind: A Research Agenda for the Digitalization of Public Services. In: *Government Information Quarterly* 36(3), pp. 427–36 (2019).
9. Lipsky, M.: *Street-Level Bureaucracy. Dilemmas of the Individual in Public Service*. Russel Sage (2010).
10. Bovens, M., Zouridis, S.: From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. In: *Public Administration Review*, 62, pp. 174–184 (2002).
11. Skaarup, S.: The Mediation of Authority. In: *scss.tcd.ie*, pp. 1–32 (2012).

12. Ebbers, W.E., Pieterse, W.J., Noordman, H.N.: Electronic government: Rethinking channel management strategies. In: *Government Info. Quarterly*, 25 (2), pp. 181–201, (2008).
13. Madsen C. Ø., Hofmann S., Pieterse W.: Channel Choice Complications. In: *International Conference on Electronic Government 2019* Springer, pp. 139–151 (2019).
14. Wynn, E.: Office conversation as an information medium. Doctoral dissertation, University of California, Berkeley, Department of Anthropology (1979).
15. Suchman, L., Wynn, E.: Procedures and problems in the office. In: *Office: Technology and People 2*, pp. 133–154 (1984).
16. Bowker, G. C., Star, S.L.: *Sorting Things out: Classification and Its Consequences*. Cambridge, Massachusetts: MIT (1999).
17. Maass, S., Rommes, E.: Uncovering the Invisible: Gender-Sensitive Analysis of Call Center Work and Software, Gender Designs IT: Construction and Deconstruction of Information Society Technologies. In: Zorn, I. et al. (eds.) *Gender Designs IT*. Wiesbaden: VS (2007).
18. Martin, D., O'Neill, J., Randall, D., Rouncefield, M.: How Can I Help You? Call Centres, Classification Work and Coordination. In: *Journal of CSCW* 16 (3), pp. 231–264. (2007)
19. Muller, M. J.: Invisible Work of Telephone Operators: An Ethnographic Analysis. In: *Journal of CSCW* 8 (1-2), pp. 31–61 (1999).
20. Nyberg, D.: Computers, Customer Service Operatives and Cyborgs: Intra-actions in Call Centres. In: *Organization Studies* 30 (11), pp. 1181–1199 (2009).
21. Svensson, M.: Routes, Routines and Emotions in Decision Making of Emergency Call Takers. Doctoral Dissertation, School of Management, Blekinge Institute of Technology (2012).
22. Tjora, A.H.: The technological mediation of the nursing-medical boundary. In: *Sociology of Health & Illness* 22 (6), pp. 721–741 (2000).
23. Whalen, J., Whalen, M., Henderson, K.: Improvisational choreography in teleservice work. In: *The British Journal of Sociology* 53 (2), pp. 239–258 (2002).
24. Bratteteig, T., Verne, G.: Conditions for Autonomy in the Information Society; Disentangling as a Public Service. In: *Scand. Journal of Information Systems* 24, pp. 51–71 (2012).
25. Verne, G.: Two faces of autonomy: Learning from non-users of an e-service. In: *System, signs and actions*, 8 (1), pp. 6–24 (2014).
26. Brandtzaeg, P., Følstad, A.: Chatbots: Changing User Needs and Motivations. In: *Interactions*, 25 (5), pp. 38–43, ACM (2018).
27. Liao, Q.V., Hussain, M.M., Chandar, P., Davis, M., Khazaen., Y., Crasso, M.P., Wang, D., Muller, M., Shami, N. S., Geyer, W.: All Work and No Play? Conversations with a Question-and-Answer Chatbot in the Wild. In: *CHI 2018*, pp. 3.1-3.13, Montreal, QC, Canada (2018).
28. Weizenbaum, J.: *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: Freeman (1976).
29. Shevat, A. *Designing bots: creating conversational experience*. Beijing (2017).
30. Holmquist, L. E.: Intelligence on Tap: Artificial Intelligence as a New Design Material. In: *interactions*, 24 (4), pp. 28–33, ACM (2017).
31. Hill, J., Ford, W.R., Farreras, I. G.: Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. In: *Computers in Human Behavior*, 49, pp. 245–250 (2015).
32. Jordan, M.I., Mitchell, T.M.: Machine learning: Trends, perspectives, and prospects. In: *Science*, 349 (6245), pp. 255–260 (2015).
33. Russell, S., Norvig, P.: *Artificial Intelligence: A modern approach*, Pearson, Boston, (2010).
34. Broussard, M.: *Artificial unintelligence: how computers misunderstand the world*. Cambridge, MA, MIT Press (2018).
35. Hendler, J., Mulvehill, A.: *Social Machines: The Coming Collision of Artificial Intelligence, Social Networking, and Humanity (1st)*. Berkeley, CA, USA: Apress (2016).
36. Myers, M.D.: *Qualitative Research in Business & Management*, 2<sup>nd</sup> edition, Sage Publications (2013).
37. Simonsen, L.: Når brukerdialogen automatiseres – hva blir vanskelig? En kvalitativ studie av sekvensen av handlinger mellom menneske og chatbot. Master thesis, Dept. Of Informatics, University of Oslo (2019).
38. Steinstrøm, T.: Bruk av chatbot i praksis. En kvalitativ studie av utfordringer ved bruk av chatbot i offentlige tjenester. Master thesis, Dept. Of Informatics, University of Oslo (2020).
39. Suchman, L. *Human-machine Reconfigurations: Plans and Situated Actions*. 2nd ed. Cambridge: Cambridge UP (2007).
40. Nguyen, T.H.: «Jeg blir så sliten i tankene». Innvandreres bruk av offentlige nettbaserte tjenester. Master thesis, Dept. of informatics, University of Oslo (2019).
41. Hageback, N.: *The Virtual Mind: Designing the Logic to Approximate Human Thinking*. Boca Raton, FL: CRC Press (2017).
42. Luger, E., Sellen, A.: Like Having a Really Bad PA: The Gulf Between User Expectation and Experience of Conversational Agents. In: *CHI 2016*, pp. 5286–5297 (2016).
43. Følstad, A., Brandtzaeg, P. B.: Chatbots and the New World of HCI. In: *interactions*, 24 (4), pp. 38–42 (2017).