

# Prosjektrapport

## Social Health Bots: Chatbot-Nora

### Gruppemedlemmer:

Aleksander Kjelstrup

Andrea Ingvaldsen

Fredrik Lauritsen

Sondre Hamnvik



## Innhold

<b>1 Introduksjon</b>	<b>2</b>
1.1 Utgangspunkt, målgruppe og prosjektgruppe	2
1.2 Forskningsartikler	3
1.3 Learning by doing	3
1.4 Metodologi	3
<b>2 Datainnsamling</b>	<b>4</b>
2.1 Første iterasjon	5
2.1.1 Wit.ai	5
2.1.2 Chatfuel	5
2.1.3 Evaluering	6
2.2 Andre iterasjon	6
2.2.1 Survey	6
2.2.2 Første fokusgruppe	7
2.2.3 Analyse	8
2.2.3.1 Innholdsanalyse	8
2.2.4 Persona	10
2.2.5 Customer Journey Map	10
2.3 Tredje iterasjon	11
2.3.1 Summativ brukertest	12
2.3.2 Grimstad & EHIN	13
<b>3 Veien videre</b>	<b>14</b>
<b>4 Etske aspekter</b>	<b>16</b>
<b>5 Refleksjon over egen prosess</b>	<b>17</b>
<b>6 Konklusjon</b>	<b>17</b>
<b>Kilder</b>	<b>18</b>

# 1 Introduksjon

Social Health Bots er et prosjekt som er utført i samarbeid med SINTEF og vår eksterne veileder Petter Bae Brandtzæg. Prosjektoppgaven har gått ut på å designe en chatbot tilpasset unge mennesker mellom 16-22 år som ønsker svar på spørsmål relatert til psykisk helse. Gjennom prosjektet har vi undersøkt diverse aspekter som plattform, design, interaksjonstype og informasjon. Den endelige prototypen vi har laget eksisterer på Facebook Messenger og er alltid tilgjengelig for målgruppen. Målet er at chatboten skal fungere som et alternativ, slik at de som ikke ønsker å oppsøke psykolog eller helsesøster kan få svar og hjelp av vår “Social Health Bot”.

## 1.1 Utgangspunkt, målgruppe og prosjektgruppe

Prosjektgruppen består av Aleksander Kjelstrup, Andrea Ingvaldsen, Fredrik Lauritsen og Sondre Hamnvik. Vi går alle på andre året på Informatikk: design, bruk og interaksjon. Ingen i prosjektgruppen hadde erfaring med utvikling av chatboter. Vi hadde heller ikke erfaring med unge menneskers psykiske helse, og var derfor klar over at vi hadde et stort prosjekt foran oss. SINTEF har fått midler fra forskningsrådet til å gjennomføre et større prosjekt kalt “Social Health Bots”[1].

Vår oppgave fra Brandtzæg var å bruke data fra Ung.no sine nettsider[2] til å lage en chatbot. Denne skulle senke terskelen for unge med psykiske utfordringer som trenger noen å snakke med, men ikke vil dra til psykolog eller helsesøster.

Etter vår designbrief kom vi fram til problemstillingen som har vært en veileder gjennom prosjektet. Problemstillingen ser slik ut:

***“Hvordan kan vi senke terskelen for å ta i bruk en chatbot for informasjon om psykiske utfordringer på en engasjerende måte”***

Dette har vært et fokus for gruppen gjennom prosjektet; hvordan kan vi gjøre en samtale mellom menneske og maskin interessant nok til at bruker vil få et godt utbytte av å bruke den.

## 1.2 Forskningsartikler

For å få mer kunnskap og en bedre oversikt over domenet chatbots, valgte vi å lese forskningsartiklene som har blitt skrevet av vår eksterne prosjektleder Petter Bae Brandtzæg og Asbjørn Følstad - "*Why people use chatbots*". Ved å lese denne artikkelen fikk vi et innblikk i hva som motiverer en person til å ta i bruk en chatbot. (Brandtzæg, Følstad 2017. s. 2-4) sier at hovedfaktoren for å bruke chatbots er basisbehovene produktivitet og sosial interaksjon. Dette ble svært nyttig for å få en tidlig forståelse av målgruppen, og hva vi ville fokusere videre på i dette prosjektet.

En chatbot er et program som utfører en samtale, gjerne via tekst eller tale. Ifølge (Brandtzæg, Følstad 2017) har teknologien innenfor domenet utviklet seg mye de siste årene ettersom større bedrifter slik som Facebook og Microsoft har investert i teknologien. Dette har ført til at det er rundt 30 000 chatbots på Facebook Messenger alene.

## 1.3 Learning by doing

For å få en praktisk forståelse om hvordan en chatbot fungerer valgte vi å teste ut en chatbot som heter Woebot[3]. Ettersom Woebot er tilknyttet plattformen Facebook Messenger, og prosjektgruppen kjente til det fra før, ble inngangen til chatbot-testing lettere. Ved å bruke Woebot over en periode fikk vi *hands on experience* og innsikt i *look and feel* til en chatbot. Botens funksjon er å være en "terapeut" for humøret til brukeren. Ved hjelp av naturlig språk-prosessering fører chatboten en samtale med brukeren og kommer med tips og råd for å motivere.

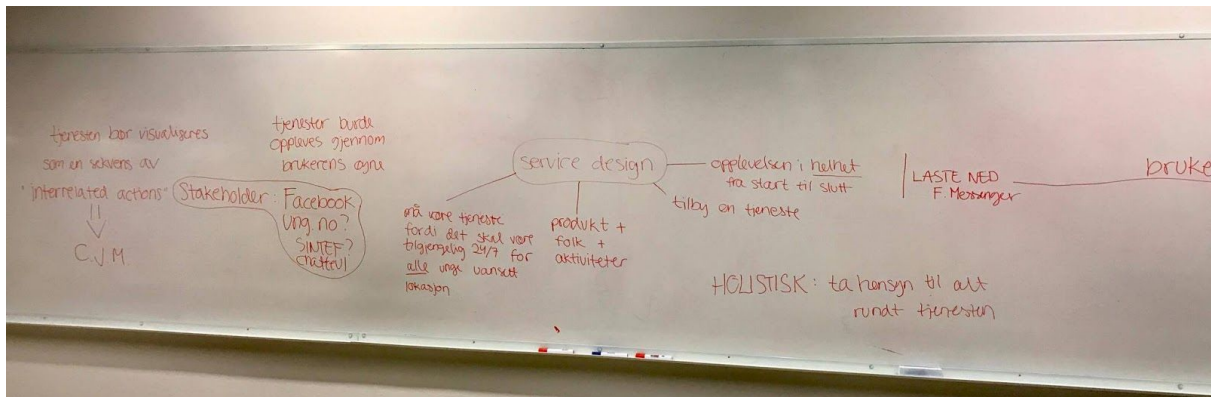
Etter en ukes tid med testing av denne chatboten valgte vi metodologien vi mente var passende for vårt prosjekt.

## 1.4 Metodologi

Interaksjonen og samtalen mellom bruker og chatbot står sentralt i vårt prosjekt. Vi definerer chatbot som en tjeneste for ungdom som ønsker informasjon eller sosial interaksjon når som helst på døgnet. Tjenestedesign ble dermed metodologien vi valgte. Touchpoints er et viktig del av tjenestedesign og defineres som knutepunktet mellom tjeneste og bruker hvor informasjon utveksles. Vi anser prosessene i Facebook Messenger med chatboten til å være våre touchpoints. For å visualisere dette,



ville vi benytte oss av flere metoder innen tjenstedesign. Vi får da muligheten til å fokusere på brukernes interaksjon med touchpointene over tid. Ved å ha en holistisk tilnærming slik som tjenstedesign er, kan vi se brukeropplevelsen og kommunikasjonen i en helhet ved å bruke ulike prototypingsmetoder. Metodene i tjenstedesign lar oss konkretisere og visualisere noe som er *intangible*. Dette hjelper oss med å arbeide mot problemstillingen vår, og å utvikle en tjeneste som er engasjerende for brukerne.



Bilde 1: Brainstorming

## 2 Datainnsamling

Vi startet innsamlingen med å samle inn informasjon rundt chatbots, og knyttet det til vårt prosjekt. Vi fikk en forståelse om at løsningen skal være en chatbot som skal kunne være tilgjengelig for ungdom døgnet rundt, samt hjelpe ungdommene med å finne svar på spørsmål.

For å kunne utvikle chatbot var det viktig at vi forsto verktøyene som var tilgjengelig for å utvikle chatbotene. Vi gjennomførte et intervju med en utvikler fra SINTEF som hadde noe kunnskaper om chatbots, og hjalp oss i gang med den tekniske løsningen. En ting vi ble klar på er at det finnes veldig få utviklerværktøy for chatbots som støtter norsk språk, noe som er en nødvendighet for vårt prosjekt da chatboten skal møte norsk ungdom på best mulig måte. Siden teknologien bak er en viktig del av prosjektet, brukte vi mye tid i starten på å eksperimentere med de forskjellige utviklings-teknologiene for å finne ut hva som fungerte best for oss. To av systemene vi la mest arbeid i var Wit.ai og Chatfuel[4][5]. Disse hadde to forskjellige type tilnærminger til utvikling av chatbots, men med forskjellige styrker og svakheter.

## 2.1 Første iterasjon

I vår første iterasjon var fokuset på teknologien og interaksjon med bruker. Vi utviklet parallelt to prototyper på forskjellige plattformer for å bli enda bedre kjent med utviklings-verktøyene som eksisterte. Den første plattformen vi utforsket var Wit.ai. Dette er en plattform som hjelper oss å tolke brukerens setninger ved hjelp av maskinlæring, og lar oss som programmerere ta alt ansvar for chatte-delen. Den andre plattformen, Chatfuel, benytter seg av drag-and-drop funksjonalitet og nøkkelord for å forstå brukeren.

### 2.1.1 Wit.ai

Wit.ai viste seg å være en teknologi som gir mange muligheter, men som også krever mye arbeid før man kan bruke alle funksjonene den tilbyr. I utgangspunktet er ikke systemet koblet opp mot en chatteplattform, så vi må selv programmere dette hvis vi vil ha vår chatbot tilknyttet Facebook. Wit.ai er en av de få eksisterende maskinlæringsverktøyene som kan behandle norske setninger og det gir oss stor frihet i forhold til prosessering av språk. Likevel viste det seg at det krever mye arbeid for å utvikle en testbar prototype. Vi forsøkte å utvikle en horisontal prototype med Wit.ai for å kartlegge interaksjonsmuligheten i tillegg til kvalitet på språktolkningen. Vi innså at det var mye som gjenstod for å kunne bruke Wit.ai videre i prosjektet da all teknisk løsning krevde mange timer med arbeid bare for å få det grunnleggende på plass.

### 2.1.2 Chatfuel

Den andre prototypen ble utviklet med bruk av Chatfuel. Det er en tjeneste tilknyttet Facebook og er et verktøy laget for utviklere med liten til ingen tidligere erfaring med utvikling av chatbots. Her var systemet allerede koblet opp mot facebook, og vi kunne bruke ferdige verktøy og elementer for å kontrollere brukerens muligheter i en samtale med chatboten. I stedet for at alt er tekstbasert kan man med Chatfuel sette inn knapper som brukeren kan trykke på. Dette er fint hvis man har områder i samtalen hvor det er et begrenset antall valg brukeren kan ta. Hvis brukeren stiller spørsmål er vi avhengig av at de bruker nøkkelord for at systemet skal forstå hva brukeren spør om. Våre erfaringer fra dette systemet er at den har problemer med å fange opp spørsmål den ikke har sett før, og ender ofte opp med å be brukeren forsøke å ordlegge seg på en annen måte. Fordelen med Chatfuel var derimot at det var intuitivt i bruk, og bra for *rapid prototyping*. Dette gir oss mulighet å raskt utvikle den videre som en evolusjonær prototype.

### 2.1.3 Evaluering

Vi valgte å forkaste Wit.ai-prototypen da den fører til at vi må bruke mer tid enn nødvendig på utvikling fremfor å bruke en enklere type chatbot som kan gi oss svar på mange av de samme tingene vi ønsker å forske på. Et endelig produkt hvor man har mer tilgjengelig utviklingstid bør sannsynligvis utvikles med Wit.ai sin språkforståelsesteknologi da det virker som det beste alternativet som eksisterer nå for å forstå norske brukere.

## 2.2 Andre iterasjon

### 2.2.1 Survey

Vi bestemte oss tidlig for å utforske survey som metode. Her kunne vi samle inn store mengder data på kort tid, uten store kostnader. Vi hadde en bred målgruppe og ønsket derfor et generelt overblikk i starten av datainnsamlingsprosessen. Deretter utviklet vi en *persona* (Preece, J. et al. 2015. s. 357-9), som hjalp oss med å visualisere hvilke mennesker som var av interesse for studiet vårt. Vi ønsket i en kort tidsperiode å legge ut surveyen på Ung.no slik at vi kunne nå ut til en andel av målgruppen. På denne måten fikk vi brukt metoden random sampling for å nå en delmengde av alle mulige deltakere i populasjonen. Dette medfører også at deltakerne velges tilfeldig, uavhengig av hverandre, og uten mulighet for å påvirke hverandres valg.

Vi anså Ung.no sitt publikum i tiden surveyen lå ute til å være vårt “sample”. Surveyen som gruppen ønsket å legge ut var utarbeidet basert på kunnskapen som hadde blitt tilegnet oss innledningsvis i prosjektet. Vi formulerte et overordnet mål for surveyen:

*“Hensikten med datainnsamlingen er å skaffe informasjon om hvordan målgruppen (16-26 år) ønsker å interagere med en helse-chatbot.”*

Da det ble gjennomført pilot-test av surveyen forstod vi at vi hadde gjort en del feil i utarbeidelsen. Vi hadde gjort en antakelse om at målgruppen allerede visste hva en chatbot var. Surveyen inneholdt flere spørsmål som både kunne være vage og vanskelige å svare på for personer som ikke hadde noen kjennskap til chatbots fra før av. Dataen ville derfor blitt sterkt preget av bias. Svarene vi får kunne dermed ført til at sluttproduktet vårt hadde lav grad av validitet.

Gruppen ble også bevisst på en etisk problemstilling i datainnsamlingen. Målgruppen består blant annet av unge mennesker under 18 år, og for at resultatet av dataanalysen skulle bli representativ i



henhold til målgruppen, trengte vi svar fra alle aldre. Dataen vi ønsket å samle inn er også noe sensitiv. Loven om sensitive personopplysninger § 9 sier at kravene i § 8 om samtykke må være oppfylt[6]. Det er derfor påkrevd å be om samtykke fra deltakerne hvis man skal samle og lagre sensitiv informasjon. For alle deltakerne under 18 år måtte vi da hatt foreldrenes samtykke. Dette, kombinert med biasen i svarene, gjorde at vi valgte å forkaste survey. Likevel var dette en lærerik prosess som gjorde oss mer bevisst på etiske hensyn og problemstillinger under videre datainnsamling i prosjektet.

### 2.2.2 Første fokusgruppe

Vi bestemte oss for å arrangere en fokusgruppe for å oppnå en dypere forståelse av hva målgruppen vet om chatbots. Vi ønsket å samle inn kvalitativ data som kunne hjelpe oss med overblikket vi følte vi trengte. I forkant av fokusgruppen diskuterte vi valg av deltakere, og hva vi ønsket å finne ut av. Project Social Health Bots sin målgruppe var i utgangspunktet 16-26 år, men etter andre møte med eksterne veileder Brandtzæg, ble gruppen enige om å innsnevre den inn til 16-22, ettersom gruppen anså dette som å være en mer passende målgruppe å jobbe med i dette prosjektet. Vi kom i kontakt med seks ungdommer, hvor alle hadde passende egenskaper for studiet vårt. Det var tre gutter på 16, 20 og 22 år, og tre jenter på 17, 18 og 19 år. Alle hadde et aktivt sosialt liv, og alle kunne i større eller mindre grad relatere seg til psykiske utfordringer i hverdagen. Alle hadde smarttelefoner med tilgang til sosiale chatte-plattformer. Det ble så utviklet et samtykkeerklæringskjema som deltakerne skrev under på. For deltakerne under 18 år, sendte vi ut skjemaet i forkant og mottok underskrift av vedkommendes foresatte.

Målet var å finne ut hva ungdommen faktisk vet om chatbots og hvordan de ønsket å bruke det. Vi forsøkte denne gangen å ikke gjøre noen antakelser om deltakernes kjennskap til domenet. Det første vi gjorde, var derfor å be deltakerne skrive ned alt de kom på i forbindelse med ordet “chatbot” på en lapp. Det viste seg at fire av seks ikke visste hva en chatbot var. I plenum så ble det diskutert hva alle hadde skrevet ned. Ut fra ordene på lappene ble det valgt ut tre ord som vi mente representerte deltakernes mentale modell av en chatbot. Disse var “**robot**”, “**chatte**” og “**samtaleemne**”. Videre hadde vi noen forhåndsdefinerte spørsmål vi ønsket å snakke rundt. Deltakerne hadde nå en bedre forståelse av hva en chatbot er, og vi klarte å skape gode diskusjoner rundt forventninger til en chatbot, interaksjonsform og hvordan den bør fremstå for at brukerne skal ha en god brukeropplevelse. Det var to fra prosjektet som deltok på fokusgruppen. En hadde fokus på samtalen og spørsmålene, mens den andre tok notater og gjorde lydopptak fra det som ble sagt.



## 2.2.3 Analyse

### 2.2.3.1 Innholdsanalyse

Vi transkriberte den delen av opptaket hvor deltakerne svarte på de forhåndsdefinerte spørsmålene. Deretter ønsket vi å gjøre *emergent coding* (Lazar, J., Feng, H. F., & Hocheiser, H. 2010. s. 289). uten forhåndsdefinerte kode-kategorier. Vi startet med å gjøre *åpen coding* (Lazar, J. et al. 2010. s. 284, 291) på én del av datasettet hver. Samtidig utarbeidet vi en liste med kode-kategorier basert på egen tolkning. Deretter samlet prosjektgruppen seg for å sammenligne og diskutere forskjellene i kategoriene. Vi valgte de kode-kategoriene det var størst enighet om, for så å kode resten av datasettet med den konsoliderte kategori-listen. Denne prosessen sikret oss høyere *inter-reliabilitet* (Lazar, J. et al. 2010. s. 296) ved at vi var fire kodere som kodet hele datasettet. Vi var *subjektive kodere* ettersom vi kodet datasett til eget prosjekt. Vi anser stabiliteten (*intra-reliabiliteten*) å være ok, da gruppen mente de interessante aspektene i datasettet kom tydelig frem, og det var stor enighet om hva de var. Likevel kan mye bias oppstå som følge av de individuelle forskjellene hos deltakerne i fokusgruppen, noe vi ønsket å være bevisst på. Deltakerne vil tolke spørsmålene basert på hvem de er og hvor de kommer fra- som igjen vil prege gruppedynamikken. I en ideell situasjon ville vi arrangert flere fokusgrupper. En med 16-18 åringer, en med 19-20 åringer og en med 21-22 åringer. Disse aldersgruppene representerer mennesker som er på et lignende stadiet i livet. Dynamikken og samtalen kunne da blitt bedre ved at deltakerne hadde følt seg mer tilpass. En 22-åring kan virke stor og voksen for en 16-åring, og 16-åringen vil kanskje unngå å være 100% ærlig. Dette kan gjøre at man går glipp av viktige innspill og tanker. Den eldre delen av målgruppen risikerer å bli overrepresentert, mens den yngre blir underrepresentert.

Vi brukte *affinity diagram* (Lazar, J. et al. 2010. s. 208, 210). som teknikk for å visualisere og sortere de tre kode-kategoriene vi hadde blitt enige om. Disse var: **Venn**, **robot-følelse** og **trygghet**. Dette er *in-vivo koder* (Lazar, J. et al. 2010. s. 291) vi lånte direkte fra deltakernes utsagn, da vi syntes det var gode begreper å jobbe videre med. “Venn” sier noe om hvordan chatboten bør **fremstå**, “robot-følelse” sier noe om hva vi skal **unngå**, mens “trygghet” sier noe om hva vi skal **oppnå**.

Frustrasjoner og potensielle frustrasjoner markerte vi med røde lapper.

Et av funnene fra denne analysen var at brukerne syntes Facebook var en plattform hvor brukerne følte seg trygge nok til å være ærlige om sensitive temaer. Likevel er dette noe man må vurdere videre i utviklingsprosessen, spesielt med tanke på at all informasjon som går over Facebooks går innom deres servere.



Prosjektgruppen tok utgangspunkt i det som ble lært fra innholdsanalysen og ønsket å gjøre metoder fra *experience prototyping* (Preece, J., Rogers, Y., Sharp, H. 2015. s. 399) for å prototype opplevelsen rundt bruk av tjenesten. Vi ønsket å simulere forholdet mellom chatbot og bruker for å oppnå bedre innsikt i hvordan brukeren kan oppleve systemet i bruk. Opplevelse er et subjektivt fenomen som varierer basert på bakgrunn, miljø og dagsform så informasjonen vi tilegnet oss her er ikke en fasit på hvordan det føles, men en pekepinn for å hjelpe vår forståelse av brukerens følelser. Vi brukte teknikken *service safari* for å visualisere brukerens reise gjennom tjenesten og oppnå større innsikt i brukeropplevelsen fra en brukers perspektiv. Videre hadde vi også et rollespill på tavla, dette hjalp oss å formulere spørsmål og svar vi kunne ta med i prototypen. Dette gjorde oss bevisst på setningsflyten og språket i samtalen.

Erfaringen fra disse teknikkene ble brukt for å utvikle et flytdiagram med hvordan samtaleflyten i en samtale kan være, fra introduksjon til avslutning. Fokuset lå på starten av samtalen og at brukerne til enhver tid skulle forstå deres muligheter. Vi ønsket en vennlig og personlig introduksjon for å unngå “robot-følelsen” som deltakerne fra fokusgruppen anså som noe negativt. Deretter bestemte vi oss for å gi brukerne mulighet til å velge mellom “bare stille spørsmål” eller å starte en samtale. Her var ønske at de som bare hadde et spørsmål, skulle få raskt svar på det, mens de som hadde behov for noen å snakke med, skulle få det.

Samtale-funksjonaliteten ble nedprioritert og ikke tatt med i denne prototypen ettersom dette ville krevd mye mer tid enn det vi hadde til rådighet. Likevel ser vi for oss en endelig løsning basert på Cognitiv Behavior Therapy i en liknende løsning som Woebot tilbyr. Vi ønsket derfor å la dette være en mulighet da deltakerne fra fokusgruppen mente dette ville dekke flere behov og tilby bedre oppfølging av psykisk helse. Vi ønsket da å utvikle en bot som kunne gi kvalitetssikrede svar på spørsmål relatert til psykisk helse. Målet var at brukerne skulle føle seg trygge. Det handler om at personopplysninger og personlig data ikke skal kunne misbrukes, og at informasjonen gitt av boten er oppriktig og til å stole på. Vi fokuserte derfor på spørsmål-delen, og hvordan vi kunne gi effektive, veiledende svar med en vennlig assosiasjon. Vi lagde tre hovedkategorier av svar, slik at vi kunne være litt genrelle med de kodeordene vi puttet i hver kategori. Disse ble: rus, psykiske utfordringer og mobbing.



implementert, ønsker vi kun å visualisere veien brukeren går for å stille et spørsmål. For å oppnå høyere *granularitet*, delte vi opp den mest sentrale fasen som er å snakke med chatboten. Ved å dele opp prosessene og identifisere underliggende aktiviteter, blir det også lettere å se nøyaktig hvor problemene oppstår.

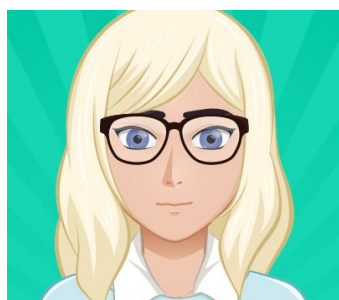
CUSTOMER JOURNEY MAP

	➔	➔	➔	➔	➔	➔
<b>AKTIVITET</b>	Oppdage chatboten	Starte samtale	Velge mellom spørsmål eller samtale	Stille spørsmål	Stille flere spørsmål eller ikke	Velge mellom å starte samtale eller avslutte
<b>TOUCHPOINTS</b>	Bekjente/web/ sosiale medier	Facebook Messenger	Facebook Messenger	Facebook Messenger	Facebook Messenger	Facebook Messenger
<b>FØLELSER</b>	Nysgjerrig/ forventningsfull	Nysgjerrig	Tenkende	Forventningsfull	Tenkende	Fornøyd
<b>NY INFORMASJON</b>		Introduksjon av chatboten og valgmulighetene				Informasjon om samtale-delen
<b>PAIN POINTS</b>	Finner ikke chatboten/har ikke tilgang på internett	Liker ikke introduksjonen	Vet ikke hva valgene innebærer	Vet ikke hva man skal skrive/ ugylding input	Fikk ikke svar på det man lurte på	
<b>BRUKERNES MULIGHETER</b>	Søke etter chatboten på internet	Starte en samtale eller avslutte	Velge mellom spørsmål eller samtale	Fri tekst	Velge ja eller nei	Samtale eller avslutte

Bilde 4: Customer Journey Map

Vi valgte å inkludere følgende aspekter: *touchpoints*, *følelser*, *ny informasjon*, *pain points* og *brukernes muligheter*. Touchpointet brukeren interagerer med i nærmest alle aktivitetene, er Facebook Messenger. Vi lærte fra fokusgruppen at dette er plattformen målgruppen er mest komfortabel på. Vi ville også å inkludere brukernes følelser om alt gikk som det skulle, samt pain points som beskriver hva som kan gå galt og skape frustrasjon. Disse genererte vi fra affinity diagrammets røde lapper. Vi ønsket å fokusere på designprinsippet *feedback* (Preece, J. et al. 2015. s. 26-27). , og var derfor opptatt av at brukeren til enhver tid skal vite hva som har blitt gjort, og hvordan han/hun kan fortsette. Vi ville derfor inkludere brukernes muligheter. For å passe på at førstegangsbrukere ikke får for mye informasjon å forholde seg til, inkluderte vi aspektet “ny informasjon”.

## 2.3 Tredje iterasjon



Fra fokusgruppen lærte vi at brukere helst vil ha en bot som takler naturlig språk, slik at en samtale kan bli så ekte som overhodet mulig. Her ble det valgt å bruke en kombinasjon av nøkkelord og flytdiagram i stedet siden en chatbot som bruker naturlig språk vil ta for lang tid å utvikle med tanke på tidsrammen i prosjektet.

Bilde 5: Avatar til chatbot - “Nora”



Mot slutten av denne innspurten valgte vi å lage en avatar til chatboten vår, slik at den kunne gi en mer menneskelig følelse. Her ga vi boten navnet “Nora”.

Vi valgte å personifisere Nora med både navn og profilbilde ettersom Nora skulle eksistere på Facebook Messenger, og vi mente at opplevelsen av en chat ville føltes ufullstendig uten navn og bilde. Utformingen av avataren ble gjort sammen med resten av gruppen hvor planen var å skape en imøtekommende og hjelpende karakter med et lite snev helsesøster i seg. Brukerne skulle føle seg trygge og velkomne til å prate med boten.

Med tanke på målgruppen og at informasjonen de ser etter gjelder psykisk helse så er det viktig å tenke på hvordan man skal presentere det. Hvis en samtale hadde bare kopiert masse informasjon og bare spyttet det ut i en tekstboble etter en annen, så kunne bruker føle seg overveldet med informasjon og samtalen ville føles unaturlig og kunstig. Målet vårt var å gjøre en samtale så personlig så mulig og få boten til å virke nesten menneskelig.

### 2.3.1 Summativ brukertest

Som en siste datainnsamling ønsket vi å teste vår høyoppløselige prototype med brukere. Dette gjorde vi i forkant av Grimstad-workshopen og EHIN-konferansen, slik at domene-ekspertene kunne hjelpe oss med de viktigste detaljene. Vi gjennomførte en *summativ* (Lazar, J. et al. 2010. s. 260) brukertest med fokus på brukbarhets faktorer, nyttighetsfaktorer og emosjonelle faktorer. Målet var å kunne si noe om den helhetlige brukeropplevelsen til deltakerne.

Vi kom i kontakt med fem deltakere med omtrent samme ferdighetsnivå knyttet til chatbots. Det er viktig for resultatet med et likt utgangspunkt når deltakerne skal utføre den samme oppgaven. En domeneekspert ville kunne utføre en oppgave raskt, mens en som aldri har prøvd en chatbot før, vil muligens behøve litt tid til å tenke. Etter å ha vært i kontakt med flere ungdommer angående chatbots, var inntrykket at de visste relativt lite. Deltakerne vi valgte hadde derfor aldri prøvd en chatbot før, men visste hva det var. Dette var to gutter på 17 og 19 år, og tre jenter på 16, 20 og 21 år. Vi utførte samme prosedyre med samtykkeerklæringsskjema som i fokusgruppen. Vi definerte en oppgave deltakeren skulle løse individuelt én gang, og observerte mens de utførte den.

---

*Start en samtale med chatboten Nora, og still spørsmålet: “Hva gjør jeg når noen blir mobbet?” Når du har fått svar, still et nytt spørsmål: “Hvordan vet jeg at jeg er gravid?”*

---

Vi tok tiden for å kunne si noe om effektiviteten, og når en deltaker var ferdig stilte vi vedkommende tre spørsmål: **Fikk du svar på spørsmålene? Hvordan var kommunikasjonen med chatboten? Hvilke følelser kjente du på når du gjennomførte oppgaven?**

Alle deltakerne fikk svar på begge spørsmålene innen 2,5 minutter. De beskrev kommunikasjonen som enkel, rask og hyggelig. Av følelser knyttet til prosessen nevnte de blant annet tilfredshet, nysgjerrighet og glede. En av deltakerne sa at chatbotens svar var litt lange. Dette gjorde at vedkommendes opplevelse av Nora ble litt upersonlig. Dette er et aspekt vi har vært bevisst på gjennom flere iterasjoner. Vi var klar over at dette fortsatt kunne være et problem, men følte vi hadde gjort så godt vi kunne med den tiden og kunnskapen vi hadde til rådighet. I en ideell situasjon ville vi ønsket enda mindre robot-følelse for brukerne av Nora.

### 2.3.2 Grimstad & EHIN

Vi hadde tidligere blitt spurt av Brandtzæg om vi ønsket å være med på et workshop om chatbots i Grimstad på Universitet i Agder. Hvor vi skulle delta og holde en presentasjon om vår tilnærming og løsning på chatbots, dette var et godt utbytte for oss siden vi fikk et større innblikk i hvordan chatbots fungerer, hvordan de jobber, hvem bruker det og hvorfor. Her var det representanter fra forskjellige bedrifter, forskningsgrupper og andre ansatte ved universiteter som deltok for å vise frem nye ideer, prototyper og ny forskning innenfor chatbots-området.

Uken etter turen til Grimstad deltok gruppen på EHIN (E-helse i Norge) konferansen i Oslo Spektrum sammen med to andre chatboter som hadde blitt utviklet i samarbeid med SINTEF. Dette var en to dagers konferanse i Oslo spektrum hvor mange bedrifter fra Norge og utlandet som jobbet med e-helse befant seg. Med prototypen vår på stand med Social Health Bots fikk vi muligheten til å vise frem prototypen og ta til oss tilbakemeldinger, som igjen var et veldig godt utbytte. Her ble det tydelig at noen aspekter ved prototypen vår ikke var enkel å forstå, blant annet på hvilke tidspunkt en skulle skrive til boten og når en skulle trykke på knapper. Det ble også observert at flere brukere ikke skjønnte at boten kun tok enkeltord som nøkler til samtalen, noe som fort førte til forvirring blant de som prøvde den ut. Etter EHIN så ble det ble en større forståelse i gruppen om hvordan brukere kom chatboten i møte, hvordan brukerens mentale modell av chatboten er og forskjeller fra implementasjon modellen.



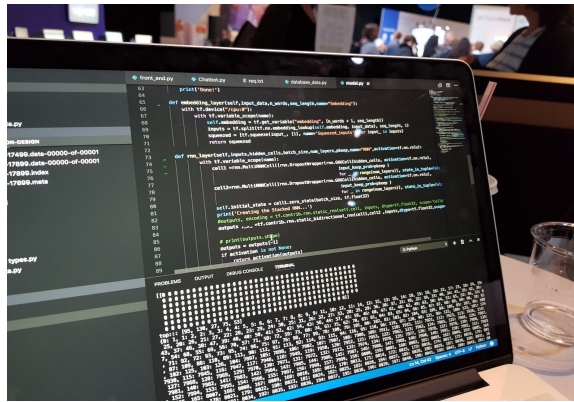
Bilde 6: Div. bilder fra EHIN

Vi fikk også mer domenekunnskap når vi sett på andre chatboter som var utviklet i samarbeid med Social Health Bots-prosjektet. Disse chatbotene hadde også noen mangler, da chatboten fra Grimstad viste seg å kunne takle spøringer på naturlig språk, men manglet evnen til å returnere et riktig svar. Den klarte kun å komme inn på et tema det ble spurt om, selv da så kunne den være langt unna det man egentlig så etter. Vi hadde også mange samtaler med personer som hadde ledet andre prosjekter med chatboter, og det var er gjennomgående stort problem å få en chatbot til å forstå hva en norsk bruker spør om. En stor grunn til dette er at det ikke finnes nok brukere som tester chatbots på norsk, og dermed får ikke maskinlærings-algortimene nok data til å kunne forstå norsk språk.

### 3 Veien videre

Vi har siden starten av prosjektet sett for oss hvordan funksjonaliteten og samtalen til en chatbot kan være. Et naturlig skritt videre i prosessen vår med prototype-boten Nora ville vært derfor vært å gi chatboten flere menneskelige egenskaper. Slik som bedre flyt i samtaler og mer korrekte svar. Etter å lest om og testet flere chatboter under prosjektet, har vi forandret vår tidligere mentale modell noe. Hva som er faktisk er mulig per i dag, samt de problemene vi i fremtiden vil støte på er noe av de tingene vi har erfart. En videre utvikling av Nora med naturlig språk vi gi oss både tekniske, etiske og sosiologiske utfordringer.

Under vår stand på EHIN fikk vi mer innsikt om hvordan kunstig intelligens kan bli brukt i en chatbot. Studenter fra UiA som har jobbet med AI i det samme prosjektet som oss kunne gi oss mer innsikt i hvordan det vil fungere med en chatbot, samt de problemene dem har med boten.



Bilde 7: Kode fra UiA prototype med kunstig intelligens

Man trenger over middels teknisk innsikt og forståelse hvordan kunstig intelligens virker for å implementere AI til en chatbot. Bare det å trene opp en bot vil ta flere måneder, basert på hvor mye data den har å jobbe med. Et eksempel på dette var at det tok rundt 12 timer å kjøre gjennom *en* treningsrunde med samlingen av data som prototypen til UiA hadde å jobbe med. Det ble altså bekreftet at å bygge en chatbot basert på kunstig intelligens var svært omfattende og tidkrevende.

Noe annet som kan føre til å korte ned levetiden til vår løsning er sosiale forventning til en chatbot. Det stilles høyere krav til kommunikasjon med en “maskin” enn et menneske. Om et menneske gir feil svar så er det forståelig for den som stiller spørsmålet, det er menneskelig å feile. Mens den generelle forventningen til en maskin er at den skal svare rett hver gang de blir stilt ett spørsmål. Hvis en chatbot ikke klarer å oppføre seg som det forventes av brukere, så vil brukere bli irriterte og droppe å benytte tjenesten. Dette gjør det spesielt vanskelig når temaet Nora skal svare på er sensitive spørsmål fra mennesker med mulige psykisk utfordringer. Derfor kan det være viktig å “trå varsomt” når vi tester med brukere. *Hva om vi ikke kan få Nora til å svare rett hver gang, hvor mye blir kredibiliteten svekket? Og hvor lang levetid har chatboten dersom vi ikke kan svare rett 100% av gangene?* Dette er spørsmål vi ville måtte spørre oss selv som designere og tatt noen grep for å få en klarhet i dette.

Hva er det rette designet for Nora sin framtid; er det flytdiagram versjonen vi har prototypet, eller en eventuell chatbot med kunstig intelligens som har en mye større sjanse for å gi feil svar til brukere. En omfattende A - B testing med en flytdiagram og en naturlig språk chatbot prototype vil kunne hjulpet oss med å få en bredere oversikt over det valget. Det kan være at det enkleste valget, flytdiagram drevet chatbot, er det beste valget per dags dato.





Chatbot har virkelig tatt verden med storm siden Facebook and Microsoft har startet å investere store summer i dem for noen år tilbake. Ny funksjonalitet og algoritmer kommer stadig, dette gir oss også nye muligheter å utvikle og designe chatbots. Det virker som chatbot er kommet for å bli.

## 4 Etiske aspekter

Vi ble bevisst over de etiske aspektene vi måtte tenke over tidlig i prosjektet. Alleredet da vi hadde vårt første møte med vår eksterne veileder så var det et fokus for samtalen. Det er en liten del av målgruppen som er under 18 og dette krevde mer forsiktighet fra vår side når det gjaldt å samle inn data. Videre så er alle temaene som prototypen tar opp veldig sensitive å snakke om, derfor er det viktig å gi brukere et trygt og sikkert miljø som de kan bruke. Noe vi også må tenke på er hvordan fremmede aktører kan prøve å ta til seg informasjon som blir skrevet til chatboten, spesielt med tanke på hvordan bedrifter som Facebook skreddersyr reklamer ut i fra hva en bruker snakker eller skriver om. Hvis slik sensitiv informasjon hadde kommet avveie hadde det ødelagt troverdigheten til chatboten vår og muligens flere fremtidige chatbots.

Vårt ansvar i dette vil være å opptre som en moralsk aktør, ved å sende ut skjemaer til de som ikke er myndige, sette oss inn i lovene rundt personopplysninger og sørge for at data ikke kommer på avveie. Et problem som kan oppstå med denne plattformen for psykisk hjelp er hvor mye en chatbot faktisk kan hjelpe, denne boten kan ikke hjelpe deg fysisk. Den kan kun videreføre informasjon, dette er et viktig punkt i prosjektet. Denne chatboten er ment for å styrke opp tilgangen til sikker informasjon og sikre inngangen til mer omfattende støtte.

Vår tekniske løsning må også stå ansvarlig for informasjonen den gir til brukeren.

Betydningen av informasjonen fra ung.no kan ha store invirkninger og konsekvenser for hva brukern velger å gjøre med de svarene han eller hun får fra boten. Derfor er det viktig at vi presenterer de riktige svarene men samtidig også henviser brukeren til å oppsøke riktig helsepersonell om nødvending. Hva og når det blir riktig å henviser til helsepersonell kan få betydelige følger og er svært viktig å få rett.

Kommunikasjonen med bruker er også noe som vi måtte ta hensyn til. Hvilke situasjon brukeren er i når han eller hun tar kontakt med chatboten kan endre valg og språkmønster i en samtale. Det er viktig for oss å kunne møte brukere i ulike situasjoner, med nøytralt språk og valgmuligheter.

Samtidig som vi vil at Nora skal fremstå som en person er det viktig at vi er tydelige på at samtalen er faktisk mellom bruker og en bot, og ikke et menneske. Dette for å unngå misforståelser om hva vår

prototype kan faktisk tilby, og for at brukeren ikke skal basere sine handlinger fullstendig på svarene fra Nora.

## 5 Refleksjon over egen prosess

Når vi skal reflektere over egen prosess i dette prosjektet, vil vi begynne med surveyen. Vi satte i gang litt for raskt, uten å være særlig bevisst på aspekter som etikk og validitet i dataene vi samlet inn. Gjennom pilottesten ble det tydelig at spørsmålene var vanskelige å svare på, og dessuten at mange ikke vet hva chatbots er. Vi forstod at datasettet ville blitt sterkt preget av bias.

Læringskurven ble bratt, og vi forstod hvor mye vi måtte ta hensyn til før vi kunne sette i gang på nytt. Vi diskuterte hvordan vi kunne samle verdifull data uten å gjøre noen antakelser om målgruppen, og gikk for å arrangere en fokusgruppe. Her fikk vi masse kvalitativ data som hjalp oss med å forstå brukerne mye bedre. På dette tidspunktet hadde vi et mye bedre utgangspunkt for å begynne å jobbe med domene. Likevel ville vi i en ideell situasjon delt fokusgruppen opp i aldersgrupper, men en fokusgruppe per aldersgruppe. Dette fordi det er stor forskjell på en 16-åring og en 22-åring, og disse kan påvirke hverandre i en sosial setting.

På workshopen i Grimstad og EHIN-konferansen var vi så heldige å få teste prototypen med domeneeksperter. Dette gav oss utrolig mye, både i form av tips til hvordan vi kunne gjøre prototypen bedre, og lærdom om domenet. Vi fikk innblikk i hvordan man utvikler og lærer opp chatbots med *artificial intelligence*. Vi snakket med mennesker som jobber med chatbots på heltid, noe som både vekket nysgjerrigheten og gav masse inspirasjon.

Når vi reflekterer over valg av metodologi, kan vi si at valget ble tatt litt for raskt. Vi hadde ikke satt oss godt nok inn i de forskjellige metodologiene og metodene de innebar. Likevel er vi fremdeles enige om at tjenestedesign var en grei metodologi for prosjektet. Metodene vi valgte å bruke har gitt oss verdifull innsikt i brukernes ståsted og domenet i helhet. Det har også gitt oss et resultat både vi, og brukeren er fornøyd med.

## 6 Konklusjon

Chatbot som psykologisk hjelpemiddel er et stort og nytt forskningsområde som det satses endel på i Norge i dag. (Fitzpatrick, K. K., Darcy, A., Vierhile, M. 2017) viser til at det absolutt er mulig å få



øke den psykiske helsen til personer ved å bruke en chatbot. Likevel kan det være store utfordringer med å få chatboter til å forstå norske brukerne. Forhåpentligvis vil de bli flinkere etterhvert på grunn av maskinlæring

En automatisert helse-chatbot kan gi store besparelser på helsesektoren, men en fullverdig erstatning for ekte helsepersonell kan en chatbot aldri bli. Chatbot kan avlaste ofte stilte spørsmål som helsepersonell ikke har tid til å svare på og være tilgjengelig tjuetimer i døgnet.

Vår prototype har begrensninger med tanke på funksjonalitet og antall spørsmål den kan svare på.

Derfor er ikke Nora egnet i en virkelig situasjon hvor bruker har et reelt spørsmål om helse. Nora er basert på et lite omfang av brukertester. Et større antall tester og analyser hadde vært nødvendig før vi kunne ta den i bruk i en natural setting. Men prototypen er *proof of concept* på at det lar seg gjøre.

## Kilder

[1] Petter Bae Brandtzæg, Social Health Bots (2017)

<https://www.sintef.no/en/projects/socialhealthbots/>

[2] Ung.no “Hjelp” (2017)

<https://www.ung.no/hjelp/>

[3] Woebot Labs Inc. (2017)

<https://woebot.io/>

[4] Wit.ai, Inc (2017)

<https://wit.ai/>

[5] 200 Labs, Inc (2017)

<https://chatfuel.com/>

[6] Lov om behandling av personopplysninger (personopplysningsloven) (2001)

<https://lovdata.no/dokument/NL/lov/2000-04-14-31>

Brandtzæg, P. B., Følstad, A. *Why people use chatbots.* (2017)

[https://www.researchgate.net/publication/318776998\\_Why\\_people\\_use\\_chatbots\\_Paper\\_accepted\\_for\\_presentation\\_at\\_the\\_4th\\_International\\_Conference\\_on\\_Internet\\_Science\\_22-24\\_November\\_2017\\_Thessaloniki\\_Greece](https://www.researchgate.net/publication/318776998_Why_people_use_chatbots_Paper_accepted_for_presentation_at_the_4th_International_Conference_on_Internet_Science_22-24_November_2017_Thessaloniki_Greece)



Brandtzæg, P.B., Følstad, A. *Chatbots and the new world of HCI* (2017)

<https://dl.acm.org/citation.cfm?id=3085558>

Fitzpatrick, K. K., Darcy, A., Vierhile, M. (2017)

*Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial*

<https://mental.jmir.org/2017/2/e19/>

Suhas Govind Joshi, Tjenestesign (2016)

[http://www.uio.no/studier/emner/matnat/ifi/INF1500/h16/undervisningsmateriale/forelesning\\_160927.pdf](http://www.uio.no/studier/emner/matnat/ifi/INF1500/h16/undervisningsmateriale/forelesning_160927.pdf)

Marit Aalen, Etikk i psykisk helsearbeid (2016)

<https://skriftserien.hioa.no/index.php/skriftserien/article/download/56/55>

Copyright for “Social Health Bots logo” - <https://logomakr.com>

Lazar, J., Feng, H. F., & Hocheiser, H. (2010). *Research Methods in Human-Computer Interaction (1st edition)*. United Kingdom, John Wiley & Sons Ltd

Preece, J., Rogers, Y., Sharp, H. (2015). *Interaction Design: Beyond human-computer interaction (4th edition)*. Great Britain, Bell & Bain Ltd