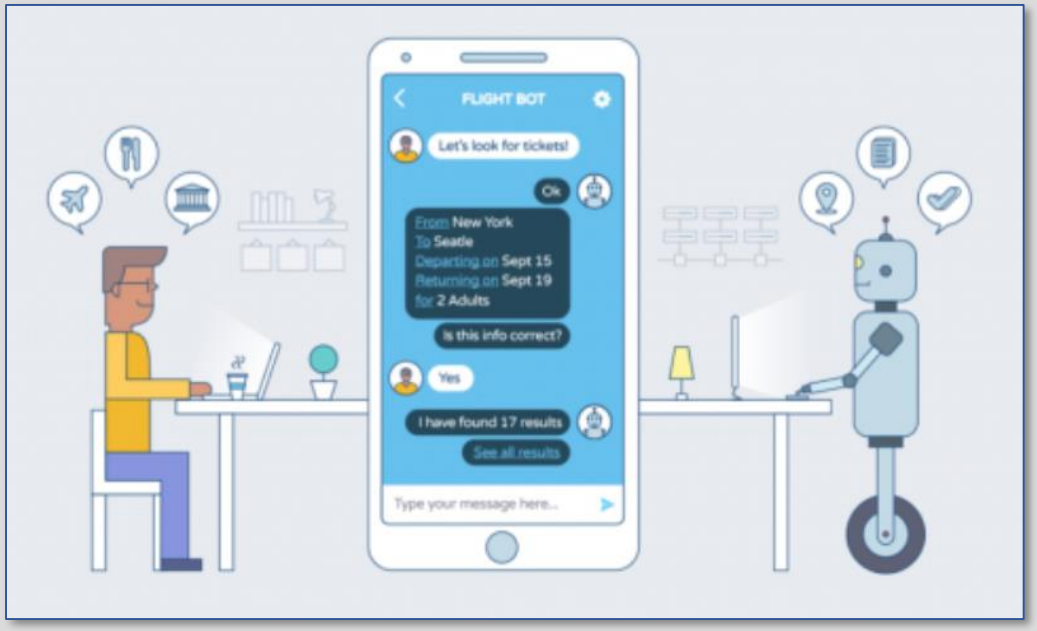
Bruk av chatbots og konversasjonsgrensesnitt i mental helse: screening og sosial støtte

Jo Dugstad Wake
NORCE Norwegian Research Centre
Technology for Health, Medicine and Learning

- Chatbots og konversasjons-grensesnitt
  - Hva er de?
  - Hvordan virker de? Infrastruktur.
  - Chatbot typologi
  - Hva gjøres i helse og mental helse
- 2 designstudier av chatbots
  - Screening av voksne med ADHD
  - Chatbot som peer-support i gruppeterapi



Hva er en chatbot?

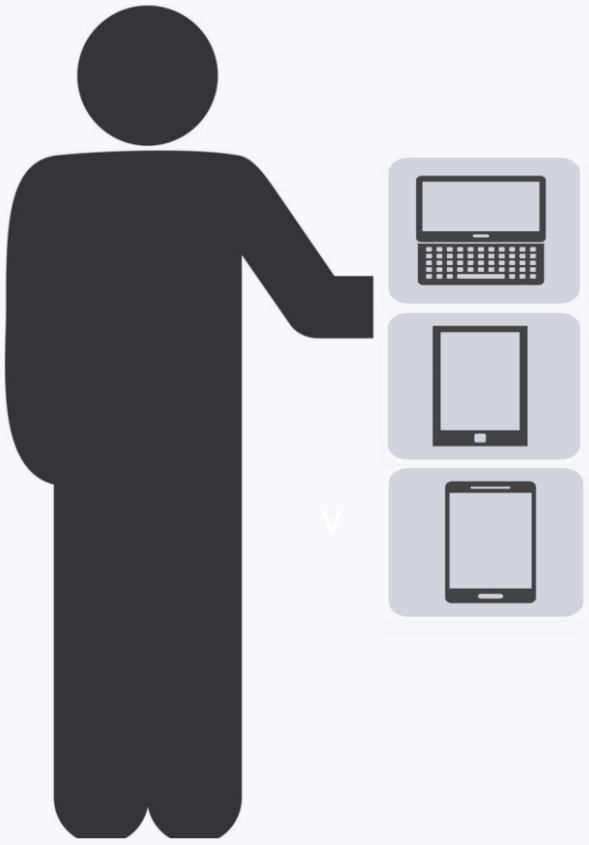
Chatbot: Chat + Robot

Chatbot [1] machine agents that serve as natural language user interfaces to data and services through text or voice

Konversasjonsgrensesnitt (*conversational interface*) [2] viser til interaksjon mellom mennesker og maskiner i form av samtale, hvor maskinen emulerer samtaleformen til et menneske

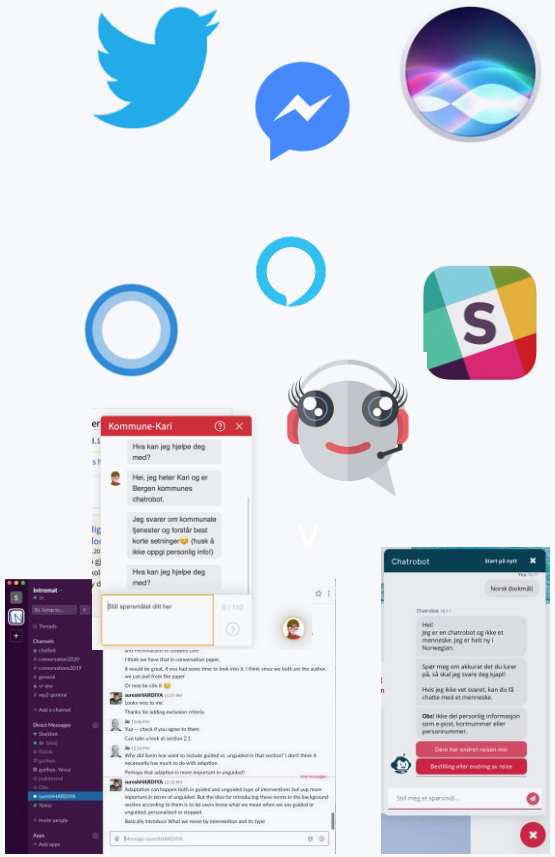
[1] Brandtzæg, P.B. & Følstad, Asbjørn (2018). Chatbots: changing user needs and motivations, Interactions (25)5, 38-43. DOI: 10.1145/3236669
[2] Zue, V. W., and J. R. Glass, "Conversational Interfaces: Advances and Challenges," *Proceedings of the IEEE,* LXXXVIII (2000), 1166–1180.

infrastruktur

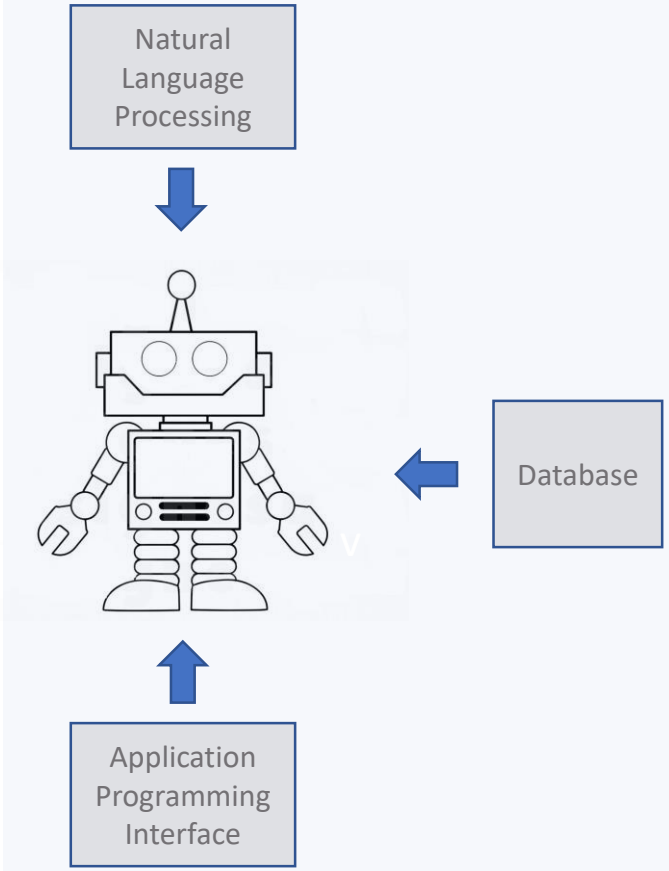


User

Device

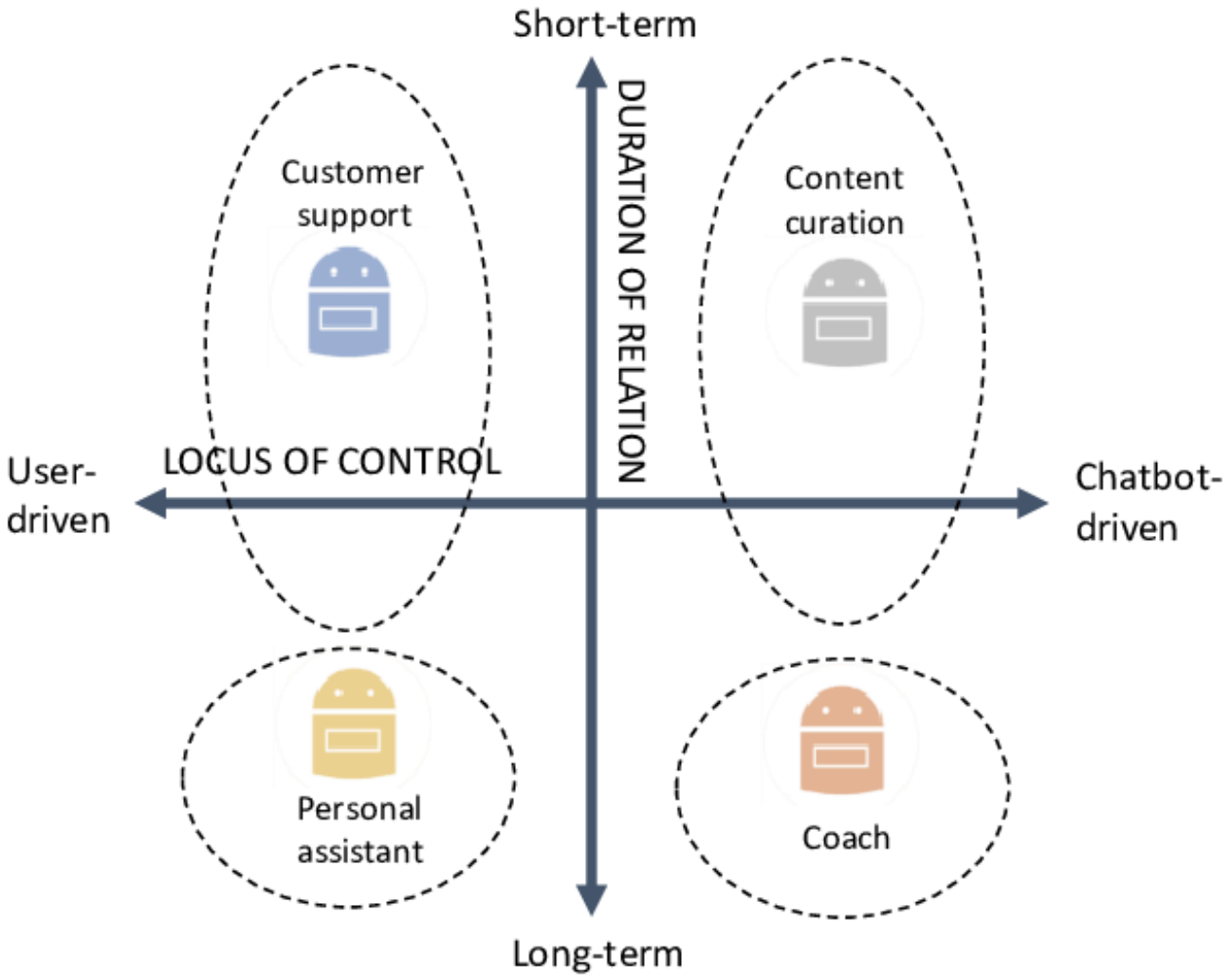


Interface



Application

typologi



Følstad, Asbjørn & Skjuve, Marita & Brandtzaeg, Petter. (2019). Different Chatbots for Different Purposes: Towards a Typology of Chatbots to Understand Interaction Design. 10.1007/978-3-030-17705-8_13.

Chatboter i helse og mental helse

Laranjo et al., 2018: 17 studier av 14 unike chatbots for helse; de som ikke handlet om mental helse handlet om: søvnforstyrrelser, Type 2 diabetes, brystkreft, smertehåndtering, høyt blodtrykk, astma og seksuell helse og stoffmisbruk for ungdommer

Abd-alrazaq, et al., 2019: 41 chatboter, brukt til å levere CBT (17), trening (12), screening (10), selv-håndtering (7), rådgivning (5), opplæring (4) og diagnostisering (2)

Lidelser som ble adressert var depresjon (16), autisme (10), PTSD (7), angst (7), stoffmisbruk (5), schizofreni (3), demens (3), fobier (2), stress og spiseforstyrrelser (1)

Chatboter i mental helse

70 % implementert som stand-alone apper, resten som web-grensesnitt
92,5% genererte respons basert på regler eller beslutningstrær, 4 brukte maskinlæring
86% hadde chatbot-drevet interaksjon
83% var tilknyttet avatar

ÍNTROMAT

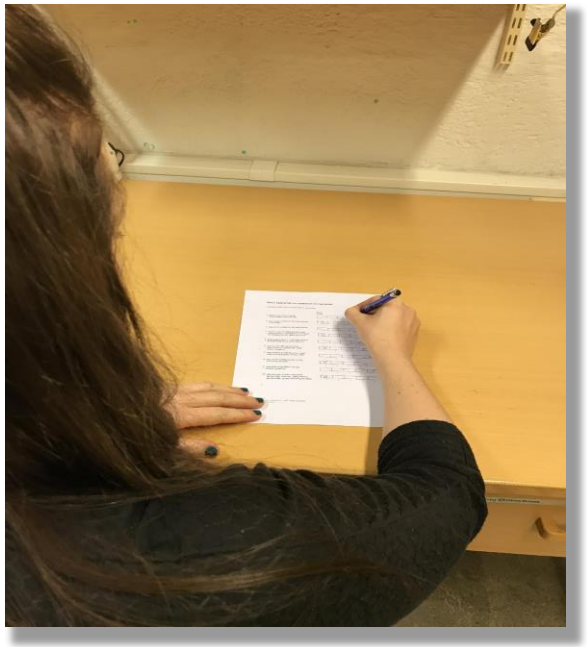
- **INTROducing Mental health through Adaptive Technology**
- Fremskritt i behandling av mental helse gjennom innovative teknologier og behandlinger
- NFR, 48 mill. NOK, 5 år
- Partnere fra akademia, sykehus/klinikker & industri
- Emeistring.no, kognitiv atferdsterapi
- Brukerdrevet forskning
- Designforskning
- Intromat.no

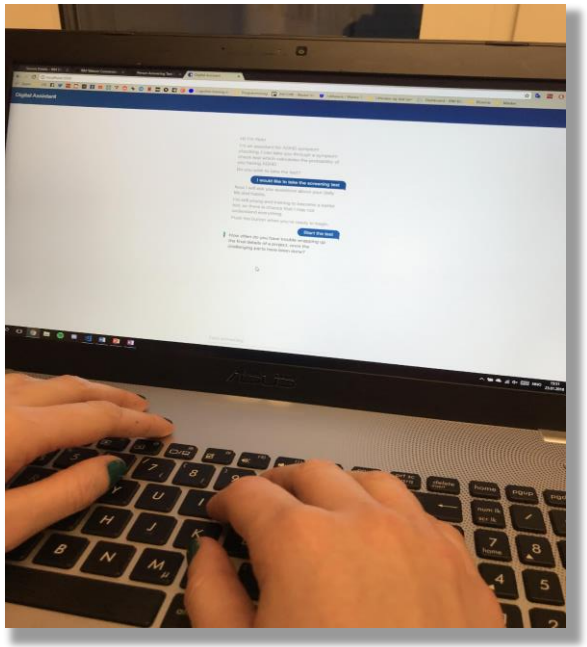
- WPs / arbeidspakker
  - Pasient-monitorering og støttesystem
  - **Pasientbehandlingsmoduler**
  - Klinisk kjerneprosess
  - Klinisk testing
- Caser
  - Motvirkning av tilbakefall i bipolar lidelse
  - Arbeidsfokusert behandling for depresjon hos voksne
  - Tidlig intervensjon og behandling av sosial angst
  - Psykososial støtte for kvinner som har gjennomgått gynekologisk kreft
  - **Kognitiv trening i ADHD**

Chatbot i screening for ADHD hos voksne, mål:

i) Gjøre om en papir-basert «screening» - the Adult ADHD Self-Report Scale (ASRS) - til konversasjons-grensesnitt

ii) Sammenligne resultatene ved bruk av chatbot og papirversjon





Case: ADHD and ASRS

- ADHD, Attention Deficiency Hyperactivity Disorder
- 3-5% av barn i Norge, noen varer til voksen alder
- ASRS: Adult ADHD Self-Report Scale
  - Dekker symptomene for ADHD i Diagnostic and Statistical Manual of Mental Disorders (DSM-5)
  - 18 spørsmål, om symptomene uoppmerksomhet / hyperaktivitet
  - Brukere av testen scorer egen atferd på 5—punkts Likert skala fra aldri til veldig ofte
- Resultat: sannsynligheten for ADHD

ADULT ADHD SELF-REPORT SCALE (ASRS-v1.1) SYMPTOM CHECKLIST

Patient Name _____ Today's Date _____

Please answer the questions below, rating yourself on each of the criteria shown using the scale on the right side of the page. As you answer each question, place an X in the box that best describes how you have felt and conducted yourself over the past 6 months. Please give this completed checklist to your healthcare professional to discuss during today's appointment.

	Never	Rarely	Sometimes	Often	Very Often
PART A					
1. How often do you have trouble wrapping up the final details of a project, once the challenging parts have been done?					
2. How often do you have difficulty getting things in order when you have to do a task that requires organization?					
3. How often do you have problems remembering appointments or obligations?					
4. When you have a task that requires a lot of thought, how often do you avoid or delay getting started?					
5. How often do you fidget or squirm with your hands or feet when you have to sit down for a long time?					
6. How often do you feel overly active and compelled to do things, like you were driven by a motor?					
PART B					
7. How often do you make careless mistakes when you have to work on a boring or difficult project?					
8. How often do you have difficulty keeping your attention when you are doing boring or repetitive work?					
9. How often do you have difficulty concentrating on what people say to you, even when they are speaking to you directly?					
10. How often do you misplace or have difficulty finding things at home or at work?					
11. How often are you distracted by activity or noise around you?					
12. How often do you leave your seat in meetings or other situations in which you are expected to remain seated?					
13. How often do you feel restless or fidgety?					
14. How often do you have difficulty unwinding and relaxing when you have time to yourself?					
15. How often do you find yourself talking too much when you are in social situations?					
16. When you're in a conversation, how often do you find yourself finishing the sentences of the people you are talking to, before they can finish them themselves?					
17. How often do you have difficulty waiting your turn in situations when turn taking is required?					
18. How often do you interrupt others when they are busy?					

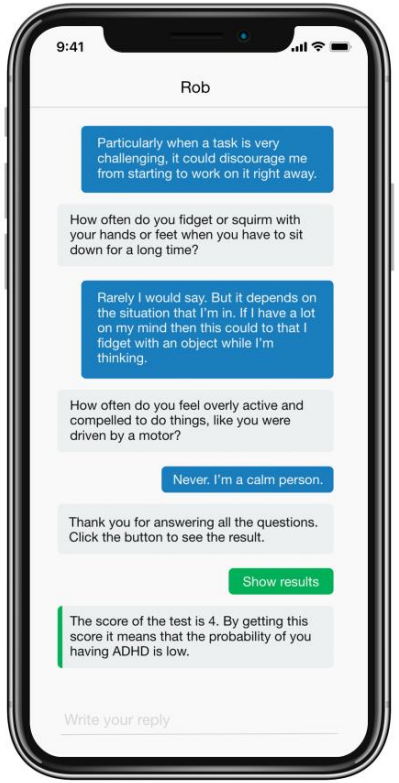
Adapted with permission © 2004 World Health Organization

Forskningdesign og data

- Designforskning (Research through design)
- 11 deltakere, masterstudenter fra Infomedia, UiB
- Ingen krav om diagnose, 3 hadde kjennskap til testen
- To grupper, «within-subjects» design
- Personvern
- Data:
  - Resultatene av screening i de to modalitetene (resultatpar)
  - Semistrukturerte intervjuer om brukeropplevelse
  - Chattelogger

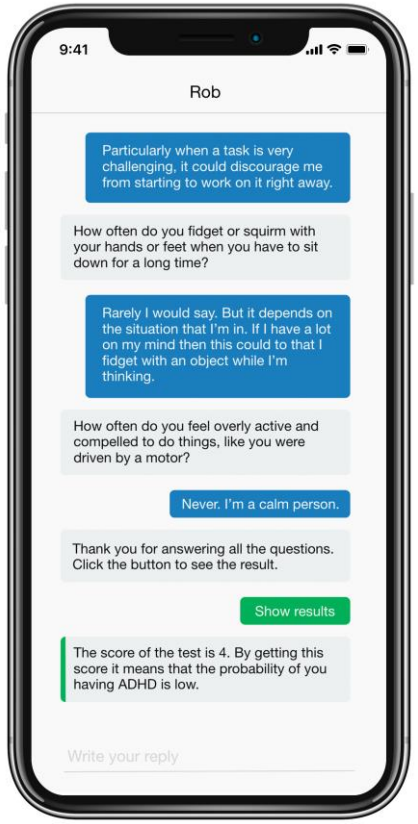
Chatboten ROB, oversikt

- Presenterer ASRS i konversasjonsgrensesnitt
- IBM Watson Assistant “intents and entities”
- Input: tekst og knapper
- Strukturert dialog
- Synonymer fra synonymordbok
- “Fallback” beskjeder (konversasjonsredning)



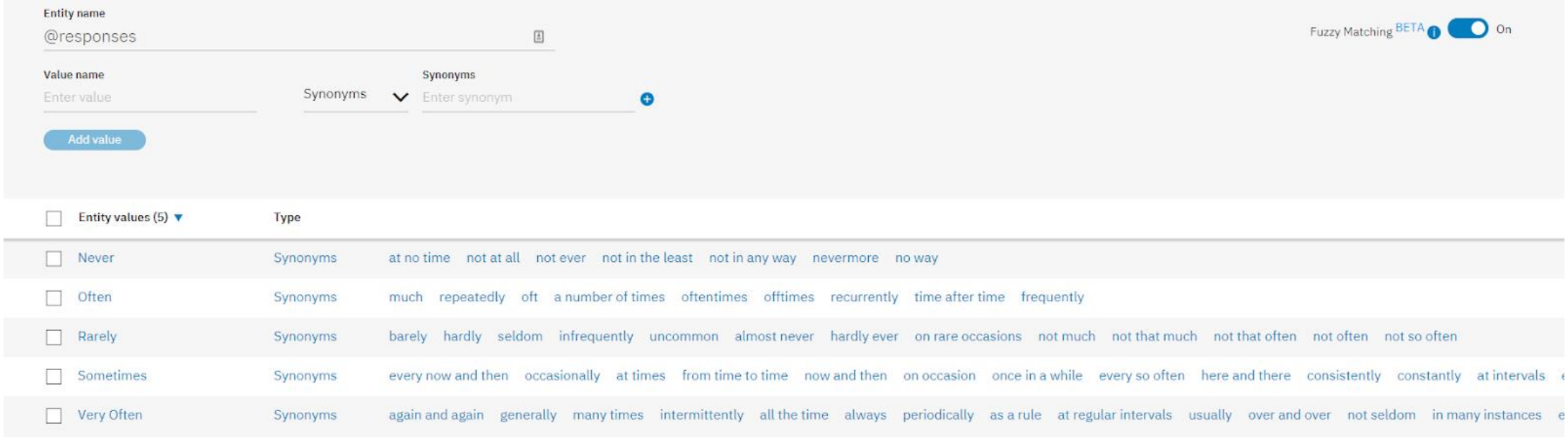
ROB, intents

- **greeting**: hilser brukeren når denne skriver en hilsen
- **goodbye**: avslutter konversasjonen
- **screening**: starter screeningen
- **information**: utløses hvis brukeren spør om ADHD



ROB, dialogstruktur og synonymer

- Strukturert dialog, initiert av ROB
- Frekvensfraser med synonymer
- «Fallback» meldinger



Resultater for de to modalitetene

Table 1. Results across modalities.

Partic.	Time Ch.	Time Pap.	Sum Ch.	Sum Pap.	Deviance
0	05:30	03:00	18	21	-3
1	04:30	02:45	40	38	2
2	06:22	01:56	24	19	5
3	03:12	01:45	35	31	4
4	03:29	01:41	23	14	9
5	10:08	02:10	33	28	5
6	02:20	02:30	6	8	-2
7	09:22	04:10	23	23	0
8	02:37	02:06	22	23	-1
9	05:04	02:26	35	30	5
10	04:15	02:28	20	21	-1

Resultat, responspar

- Av 197 responspar var 137 identiske
- 60 hadde ulikheter i svarene
- Gjennomsnittlig avvik i de parene var 1.01 (et av parene hadde diskrepans på 2)

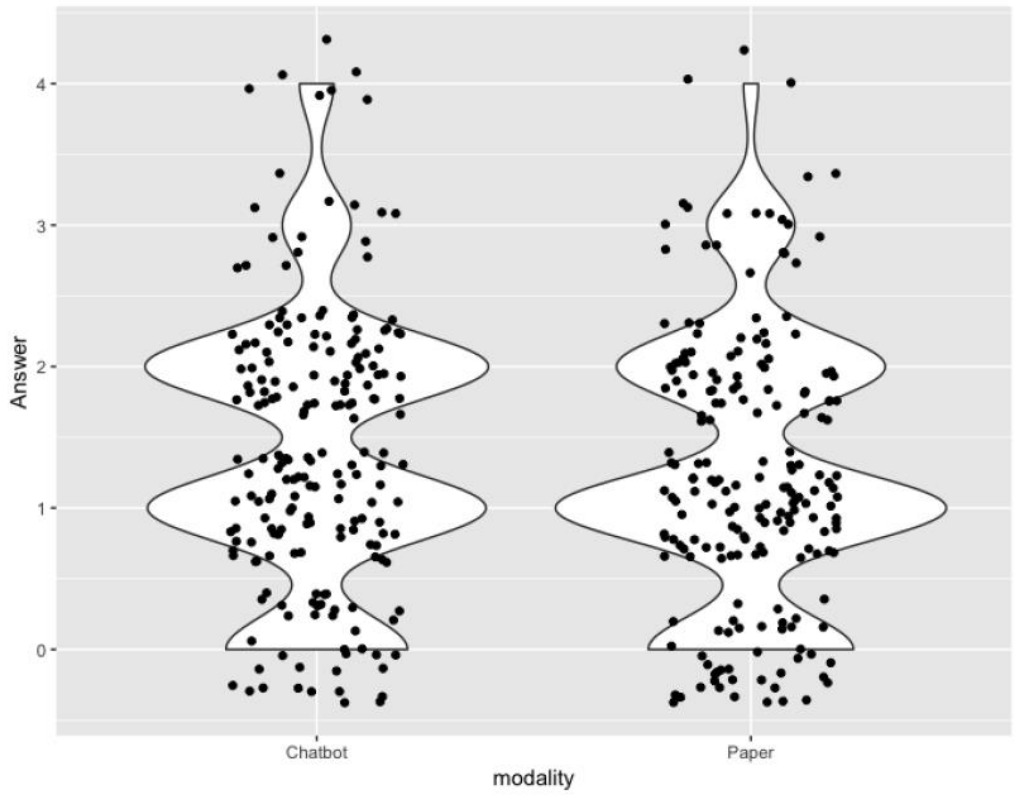


Fig. 2. Scatterplot of answers to all questions across modalities.

Resultater, chattelogg

- Respons-typer hos brukerne:
  1. Enkeltorduttalelser
  2. Setninger som inneholdt frekvens-basert uttalelse
  3. Setninger som la til ekstra innhold eller kontekst

- Examples:
  - Spørsmål: “Hvor ofte distraheres du av bråk eller aktivitet i omgivelsene?"
    - «Ikke veldig ofte.»
    - ”Noen ganger blir jeg distrahert av bråk eller aktivitet i omgivelsene."
    - ”Det kommer an på bråket, men noen ganger. Konversasjoner er ofte fokuskillere."

Resultat, intervjuer

- Satte pris på muligheten til å reflektere mer generelt rundt spørsmålene
- Satte pris på muligheten til å gi mer kontekst
- Mangelen på »svar til svar» eller anerkjennelse av respons, føltes «kaldt» sammenlignet med menneskelig kommunikasjon

Konklusjon og begrensninger

- Begrensninger:
  - Lite utvalg, begrensninger i forklaringskraften
  - Potensiell læringseffekt, som følge av forskningsdesign
  - «Novelty effect» eller nyhets-effekt av konversasjonsgrensesnitt

- Konklusjoner
  - Chatbots kan være en nyttig tilvekst til papirbaserte screeninger i mental helse, med økte muligheter for refleksjon
  - Chatbots kan integreres med andre plattformer, spres enklere, så lenge personvernet til brukeren ivaretas

Chatbot som peer support

- De fleste chatboter støtter dyadisk interaksjon, eller interaksjon mellom en bruker og chatboten
- Seering et al., 2019: 103 av 130 dyadiske, 13 multiparti
- Det er et uutforsket potensiale i multiparti-chatboter; hvordan de kan mediere konversasjon i grupper av mennesker
- Peer support: mennesker i samme situasjon som støtter hverandre
- Online peer support: interaksjonen støttet av IKT
- Utfordring: frafall, motvirkes av mentoring and monitoring

- Mål: Hvordan kan chatboter legge til rette for peer support i online terapi-grupper?

Metode

- Designforskning: research through design
- Wizard of Oz: 2 varianter
- Terabot (Therapy Chatbot), co-designet med kliniske psykologer med erfaring fra gruppeterapi for ADHD + designstrategier for ADHD
- Kontekst: online gruppe-terapi for voksne med ADHD
- To «trials»:
  - Wizard of Oz 1, uten krav om diagnose, 9 deltakere
  - Wizard of Oz 2, med voksne med ADHD, (6) 4 (2) deltakere
  - Trial struktur: introduksjon til «the stop excercise», evaluering av øvelsen
  - Data: observasjoner, chat-logger, spørreskjema + evaluering av trial 2

Resultat, spørreskjema

Statement	Scores	Trial 1 (N=18)	Trial 2 (N=6)	Mean (N=24)
To what extent did you experience contact with the other participants in the chat?	1 = very little, 7 = very large	5,34	6	5,67
To what degree did you find it useful to discuss 'The Stop Exercise' with people in the same situation?	1 = very useless, 7 = very useful	6	5,25	5,63
What was your impression of Terabot (the chatbot)?	1 = very negative, 7 = very positive	6,5	4,75	5,63
How disturbing did you experience Terabot, the chatbot, in the conversation?	1 = very disturbing, 7 = not disturbing at all	5,78	4,88	5,33
How useful did you find Terabot, the chatbot, for the conversation?	1 = very useless, 7 = very useful	6,62	5,42	6,02
How did you experience the distribution between who spoke/chatted?	1 = very little, 7 = very large	6,56	6,13	6,35
To what degree did you experience to say what you wanted to?	1 = very little, 7 = very large	6,28	6	6,14

Table 1. Mean questionnaire results Trial 1 and 2

Resultat

- Positivt press for å gjennomføre oppgaven
- «Peer» support:

  “I also thought it was nice to have peer support. I often experience that people without ADHD can sometimes make me feel like they are there to educate. I can often experience that other people are “proper" and that they want to educate us. That's a bit condescending, and you feel like you have to compensate in a way for how you behave“
- Scheduling / planlegging av møtetidspunkt, hvordan fange opp og designe rundt at deltakere kan være forsinket

Takk for oppmerksomheten!

Havik, R., Wake, J.D., Flobak, E., Lundervold, A., Guribye, F.: A Conversational Interface for Self-screening for ADHD in Adults. In: Bodrunova, S.S., Koltsova, O., Flstad, A., Halpin, H., Kolozaridi, P., Yuldashev, L., Smoliarova, A., Niedermayer, H. (eds.) Internet Science. INSCI 2018. Lecture Notes in Computer Science, vol 11551. pp. 133-144. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-17705-8 12

Oda Elise Nordberg, Jo Dugstad Wake, Emilie Sektnan Nordby, Eivind Flobak, Tine Nordgreen, Suresh Kumar Mukhiya, and Frode Guribye. (2019). Designing chatbots for guiding online peer support conversations for adults with ADHD. Conversations 2019, 3rd international workshop on chatbot research. November 19-20, Amsterdam, The Netherlands.

Jo Dugstad Wake
NORCE Norwegian Research Centre
Technology for Health, Medicine and Learning