

My background:

Human-Chatbot Interaction Design

How to design chatbots?

Supported by IKTPLUSS, Research Council of Norway



Chatbots og interaksjonsdesign

Naturlig språk blir en stadig vanligere måte å bruke digitale tjenester på. Eksempler er Apples Siri og Googles Assistant. Facebook Messenger gir i dag tilgang til mer enn 30.000 chatbots.

Til tross for en rivende teknologiske utvikling innen chatbots gjenstår mange forskningsutordringer. I prosjektet **Human-Chatbot Interaction Design** kombinerer vi interaksjonsdesign med kunstig intelligens og maskinlæring.

Målet er å bidra til kunnskap om brukersentrert design av chatbots, slik at disse kan bli nyttige og engasjerende innganger til informasjon og tjenester – uansett brukerens kjønn, bakgrunn og alder.

OM PROSJEKTET

Prosjektperiode 2017. Med mulig forlengelse til 2020

Program IKTPLUSS

PARTNERE





STØTTET AV



Kontakt: Prosjektleder Petter Bae Brandtzæg | 92806546 | pbb@sintef.no Les mer: www.sintef.no/humanchatbotinteractiondesign

My background:

SINTEF, HCl group

Psychology (!?!)



Interaction with AI – module 2

Interaction design

Four sessions

The user and interaction design perspective

Asbjørn Følstad

August 29

The machine learning perspective

Morten Goodwin

October 10

October 17

October 24

Literature

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Teevan, J. (2019). Guidelines for human-Al interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 3). ACM.

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 411). ACM.

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *interactions*, 24(4), 38-42.

Luger, E., & Sellen, A. (2016). Like having a really bad PA: the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5286-5297). ACM.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: how susceptible are jobs to computerisation?. *Technological forecasting and social change*, *114*, 254-280.

Carter, S., & Nielsen, M. (2017). Using artificial intelligence to augment human intelligence. *Distill*, *2*(12), e9.

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: humans and AI are joining forces. *Harvard Business Review*, *96*(4), 114-123.

Noessel, C. (2017). Designing Agentive technology: Al that works for people. Rosenfeld Media.

Hall, E. (2018). Conversational design. A Book Apart

McAfee, A., & Brynjulfsson, E. (2016). Machine, Platform, Crowd. Harnessing Our Digital Future. Norton & Company.

Individual assignment

Building on and extending the individual assignment in Module 1. Startup when Module 1 is completed – finish October 25)

Module 2 Mid-term report (Oct. 25)

Group assignment

Building on and extending the group assignment in Module 1. Startup when Module 1 is completed – finish October 25)

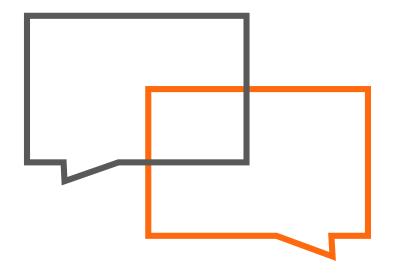
- + Task on user and interaction design perspective (startup today – finish September 5, include in mid-term report as Appendix 1)
- + Task on machine learning perspective (startup Oct. 10 finish Oct. 17, include in mid-term report as Appendix 2)

Agenda today

- 1 Interacting with AI an overview
- User-centred design of AI
- Chatbots interacting with AI in natural language

Interaction with AI - overview

A1?



AI?



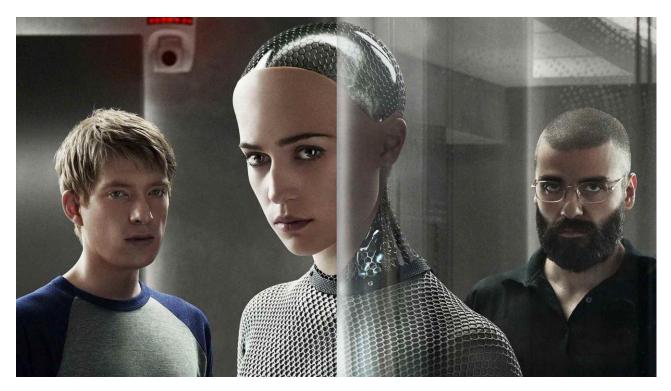


Person of Interest (2011 - 2016)

Artificial super intelligence

Artificial general intelligence

Artificial narrow intelligence

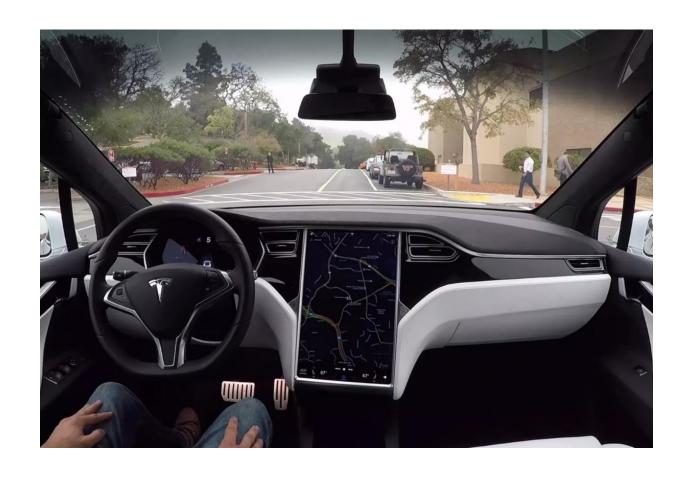


Ex Machina (2014)

Artificial super intelligence

Artificial general intelligence

Artificial narrow intelligence



Artificial super intelligence

Artificial general intelligence

Artificial narrow intelligence

Interaction with AI, and designing for interactions with AI, concerns artificial narrow intelligence (narrow AI)

Artificial super intelligence

Artificial general intelligence

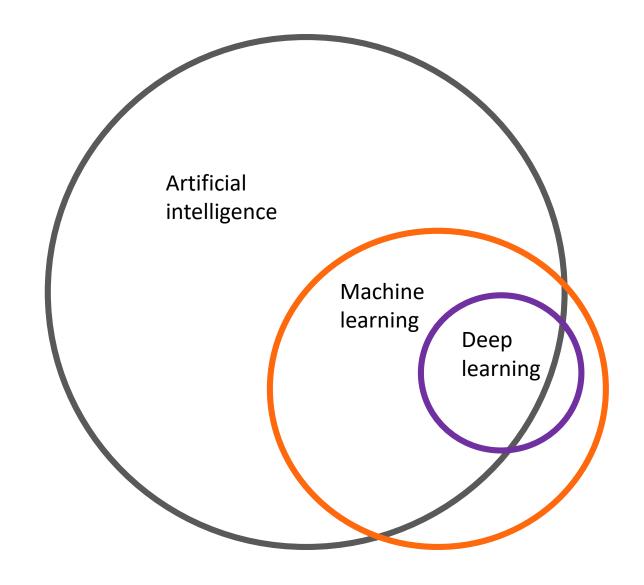
Artificial narrow intelligence

Interaction with AI, and designing for interactions with AI, concerns artificial narrow intelligence (narrow AI)

Need to know:

- Characteristics of narrow AI
- Opportunities and limitations of narrow Al

... from now on, when talking about Al we mean narrow Al



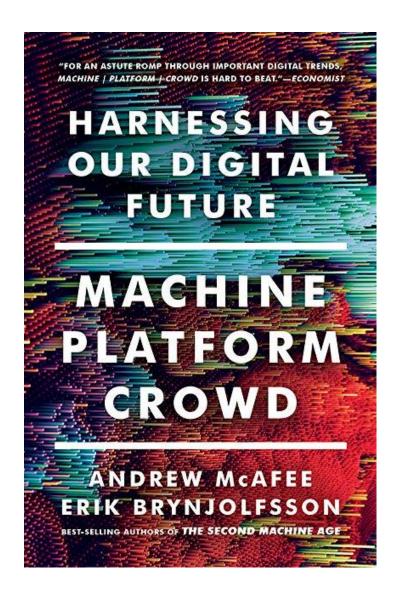
Two drivers:

Large datasets

Computational power

An understanding of recent Al systems may be:

Computer systems
learning and improving
on the basis of large data
sources



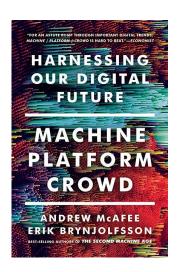


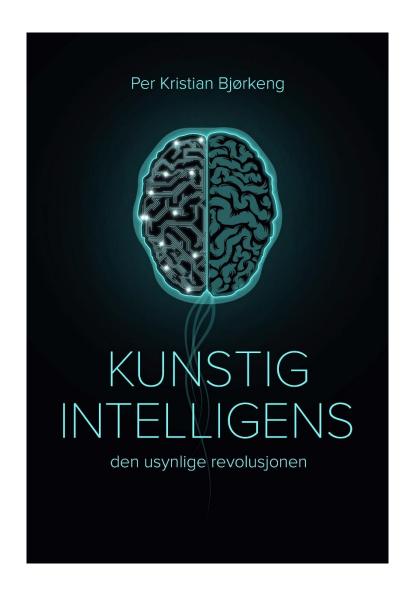
- Phase 1: Computers drives productivity
- Phase 2: Computers can do work we never thought of as programmed or routine

Michael Polanyi on tacit knowledge - "We know more than we can tell." (Polanyis paradox)

Computers now taking on tasks reflecting Polanyis paradox

The board game Go paramount example



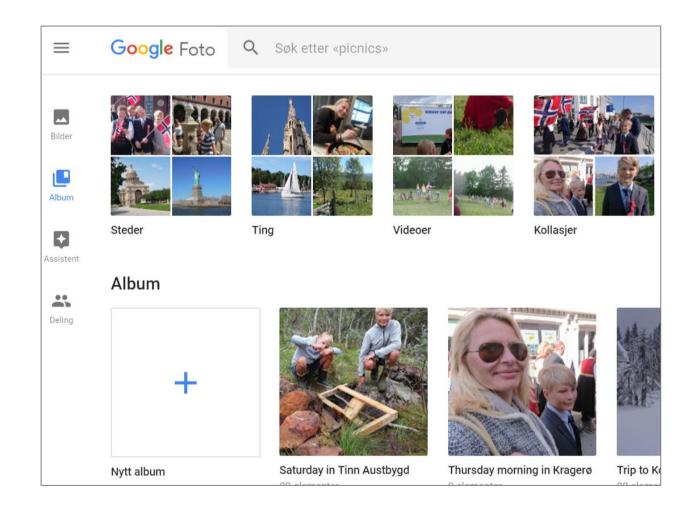


De fleste datasett som kan brukes til å trene dype nett til å bli gode på en eller annen funksjon, er kjempestore.

[...]

Et nytt triks som mange bedrifter nå bruker, er å først designe en tjeneste som mange brukere vil ha. [...] Gjennom bruken av tjenesten gir brukerne fra seg verdifulle data som igjen brukes til å trene nett for å gi brukerne nye funksjoner de setter pris på.

Bjørkeng, P. K. (2018). Kunstig intelligens – den usynlige revolusjonen. Vega.

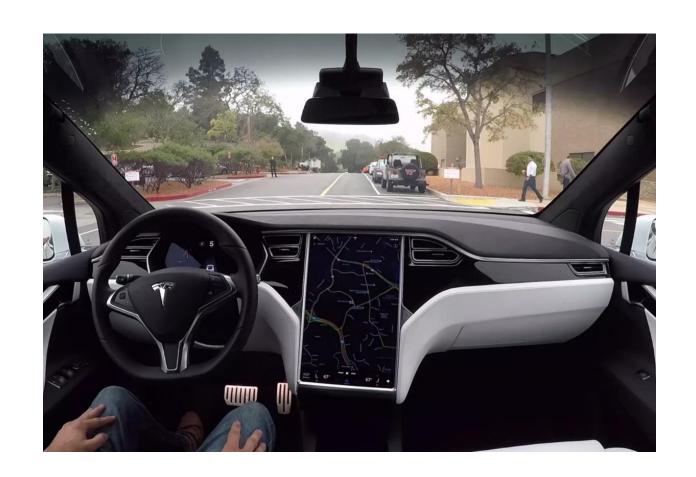


En datafelle er begrepet som brukes om Teslas særegne tilnærming til datainnsamling.

Google-eier Alphabet og deres selskap Waymohar nå brukt ti år bare på å samle inn data om mange nok ulike trafikksituasjoner. (menneskelige sjåfører i selvkjørende biler klare til å ta over)

Elon Musk og Tesla hadde slett ikke 10 år til overs. I stedet etablerte de en datafelle. [...] Her eneste nye Tesla-eier er nå med i dette gigantiske datainnsamlingsprosjektet (shadowmode)

Bjørkeng, P. K. (2018). Kunstig intelligens – den usynlige revolusjonen. Vega.



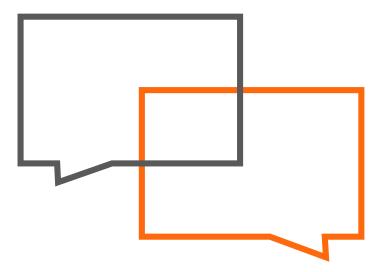
Al-infused systems

"Systems that have features harnessing AI capabilities that are directly exposed to the end user."

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Teevan, J. (2019). Guidelines for human-Al interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 3). ACM.

Al-infused systems

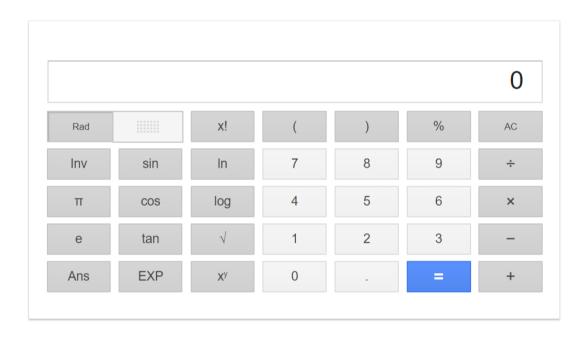
favouriteexamples?



Characteristics of Al-infused systems



VS.



Learning

Improving

Black box | Fuelled by large data sets

Dynamic

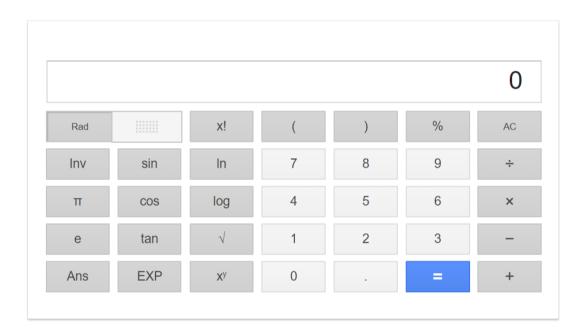
Mistakes inevitable

Opaque

Data gathering through interaction



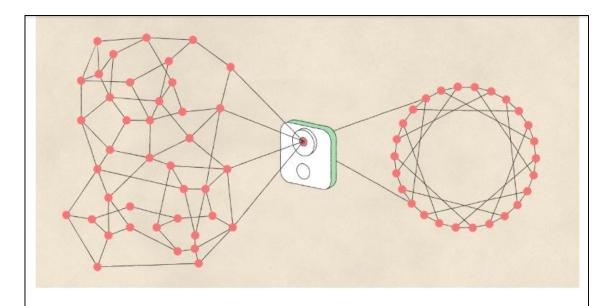
VS.



User-centred design of Al

Why human-centred design of AI?

- Addressing a real human need
- Guiding the intelligence (consider human expertise)
- Building trust



The UX of AI

Using Google Clips to understand how a human-centered design process elevates artificial intelligence

By Josh Lovejoy

As was the case with the mobile revolution, and the web before that, machine learning will cause us to rethink, restructure, and reconsider what's possible in virtually every experience we build. In the Google UX community, we've started an effort called "human-centered machine learning" to help focus and guide that conversation. Using this lens, we look across products to see how machine learning (ML) can stay grounded in human needs while solving for them—in ways that are uniquely possible through ML. Our team at Google works across the company to bring UXers up to speed on core ML concepts,

Guidelines for Human-Al Interaction

Guidelines for Human-Al Interaction

Saleema Amershi, Dan Weld'[†], Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz

Microsoft
Redmond, WA, USA
{samershi, mivorvor, adamfo, benushi, pennycol, jinsuh,
shamsi, pauben, kori, teevan, ruthkg, horvitz}
@microsoft.com

ABSTRACT

Advances in artificial intelligence (AI) frame opportunities and challenges for user interface design. Principles for human-AI interaction have been discussed in the human-computer interaction community for over two decades, but more study and innovation are needed in light of advances in AI and the growing uses of AI technologies in human-facing applications. We propose 18 generally applicable design guidelines for human-AI interaction. These guidelines are validated through multiple rounds of evaluation including a user study with 49 design practitioners who tested the guidelines against 20 popular AI-infused products. The results verify the relevance of the guidelines over a spectrum of interaction scenarios and reveal gaps in our knowledge, highlighting opportunities for further research. Based on the evaluations, we believe the set of design guidelines can serve as a resource to practitioners working on the design of applications and features that harness AI technologies, and to researchers interested in the further development of guidelines for human-AI interaction design.

CCS CONCEPTS

Human-centered computing → Human computer interaction (HCI);
 Computing methodologies → Artificial intelligence.

*Work done as a visiting researcher at Microsoft Research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. CH 2019, May 4-9, 2019, Glasgow, Soctland UK

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00 https://doi.org/10.1145/3290605.3300233 [†]Paul G. Allen School of Computer Science & Engineering University of Washington Seattle, WA, USA weld@cs.washington.edu

KEYWORDS

Human-AI interaction; AI-infused systems; design guidelines

ACM Reference Format:

Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N, Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-Al Interaction. In CHI Conference on Human Fadors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300233

1 INTRODUCTION

Advances in artificial intelligence (AI) are enabling developers to integrate a variety of AI capabilities into user-facing systems. For example, increases in the accuracy of pattern recognition have created opportunities and pressure to integrate speech recognition, translation, object recognition, and face recognition into applications. However, as automated inferences are typically performed under uncertainty, often producing false positives and false negatives, AI-infused systems may demonstrate unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous. While some AI technologies are deployed in explicit, interactive uses, other advances are employed behind the scenes in proactive services acting on behalf of users such as automatically filtering content based on inferred relevance or importance. While such attempts at personalization may be delightful when aligned with users' preferences, automated filtering and routing can be the source of costly information hiding and actions at odds with user goals and expectations.

AI-infused systems¹ can violate established usability guidelines of traditional user interface design (e.g., [31, 32]). For example, the principle of consistency advocates for minimizing unexpected changes with a consistent interface appearance and predictable behaviors. However, many AI components are inherently inconsistent due to poorly understood,

¹In this paper we use AI-Infused systems to refer to systems that have features harnessing AI capabilities that are directly exposed to the end user.

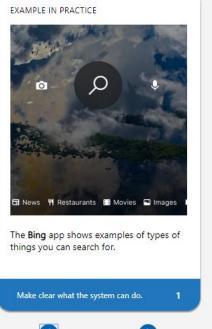


Hands on with Al / Guidelines for Human-Al Interaction



Make clear what the system can do.

Help the user understand what theAl system is capable of doing.





Example 2 of 3











Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)

Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)

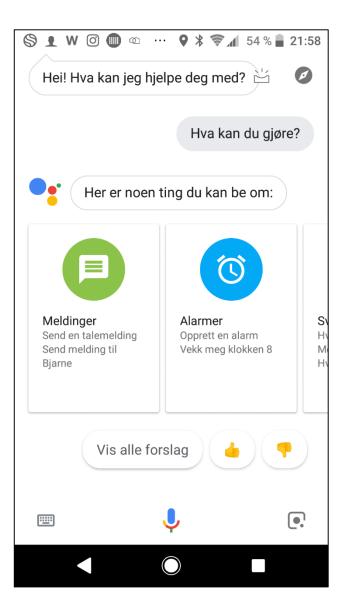
1 INITIALLY

Make clear what the system can do.

Help the user understand what the Al system is capable of doing.

Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)



Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)

2 INITIALLY

Make clear how well the system can do what it can do.

Help the user understand how often the Al system may make mistakes.

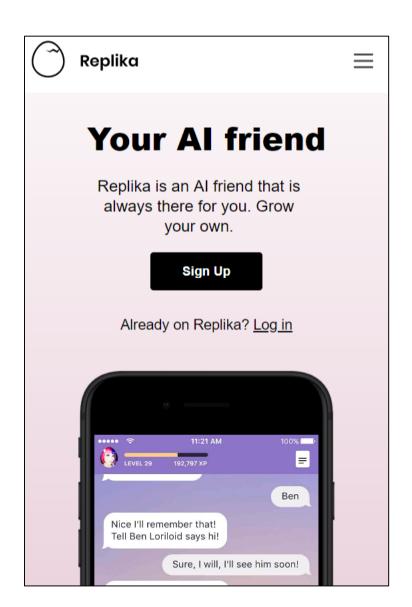
Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)



Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)



Learning system - design for change

- M1: make clear what the system can do
- M2: make clear how well the system can do what it can do
- Explain dynamic character (?)



Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt

Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt

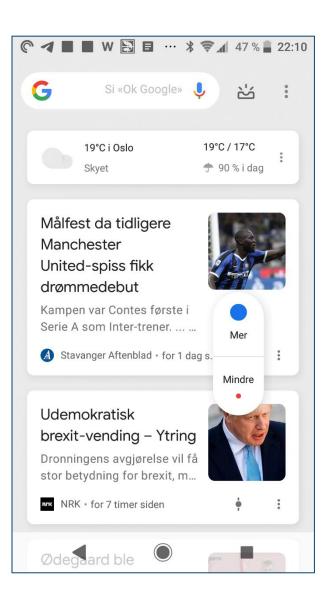
9 WHEN WRONG

Support efficient correction.

Make it easy to edit, refine, or recover when the AI system is wrong.

Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt



Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt

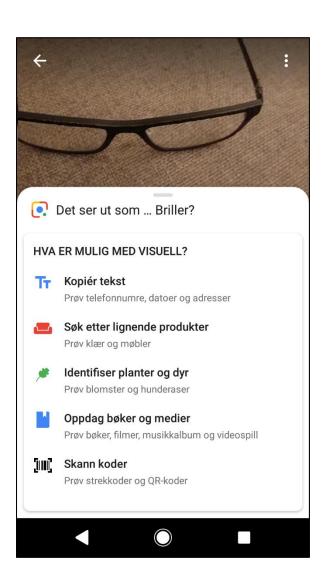
10 WHEN WRONG

Scope services when in doubt.

Engage in disambiguation or gracefully degrade the AI system's services when uncertain about a user's goals.

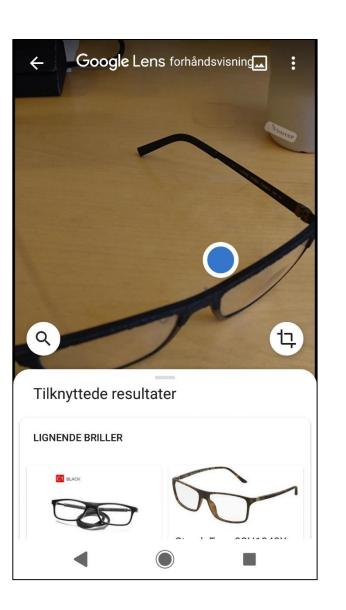
Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt



Mistakes inevitable - design for uncertainty

- M9: Support efficient correction
- M10: Scope services when in doubt



Difficult to understand and validate output – design for explainability

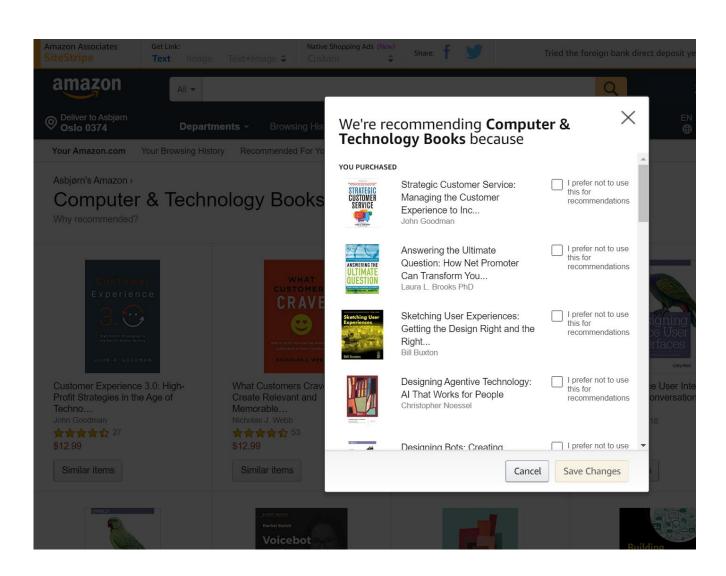
 M11: Make clear why the system did what it did 11 WHEN WRONG

Make clear why the system did what it did.

Enable the user to access an explanation of why the Al system behaved as it did.

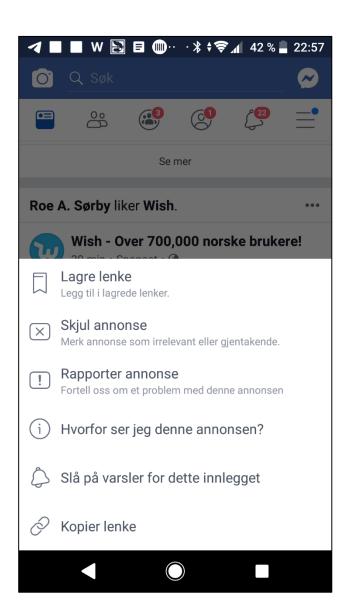
Difficult to understand and validate output – design for explainability

 M11: Make clear why the system did what it did



Difficult to understand and validate output – design for explainability

 M11: Make clear why the system did what it did



Data wanted – design for data capture

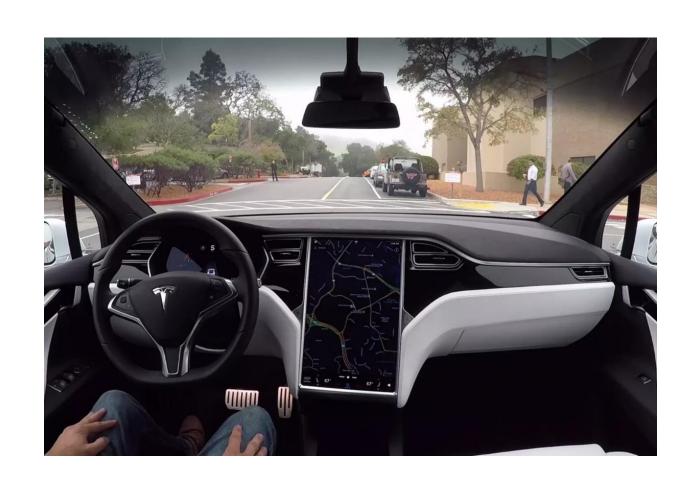
- Accommodate gathering of data from users
- ... but with concern for the risk of being gamed
- Make users benefit from data
- Privacy by design

13 OVER TIME

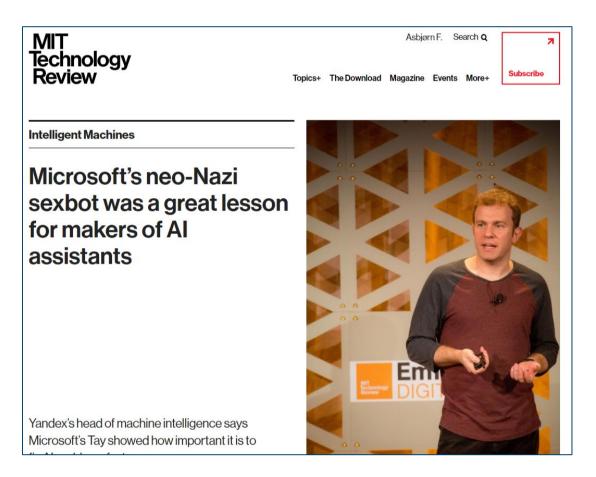
Learn from user behavior.

Personalize the user's experience by learning from their actions over time.

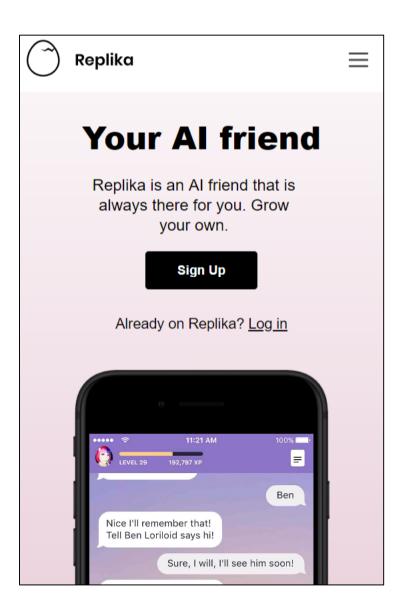
- Accommodate gathering of data from users
- ... but with concern for the risk of being gamed
- Make users benefit from data
- Privacy by design



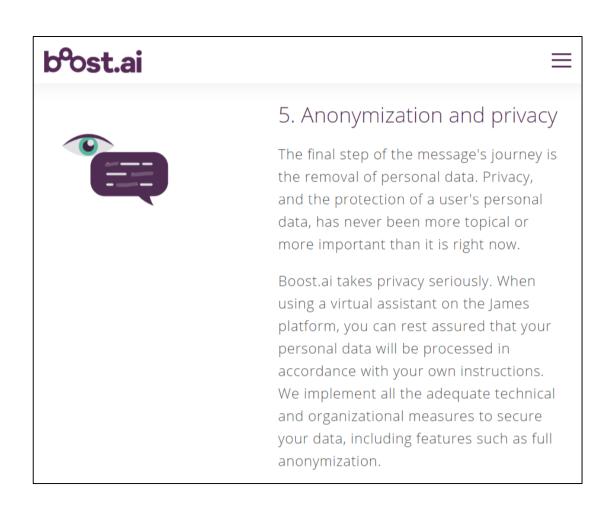
- Accommodate gathering of data from users
- ... but with concern for the risk of being gamed
- Make users benefit from data
- Privacy by design



- Accommodate gathering of data from users
- ... but with concern for the risk of being gamed
- Make users benefit from data
- Privacy by design



- Accommodate gathering of data from users
- ... but with concern for the risk of being gamed
- Make users benefit from data
- Privacy by design

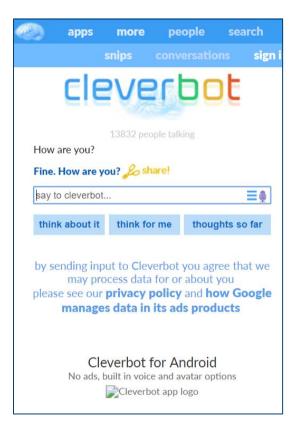


Chatbots – interacting with AI in natural language

One definition of chatbots:

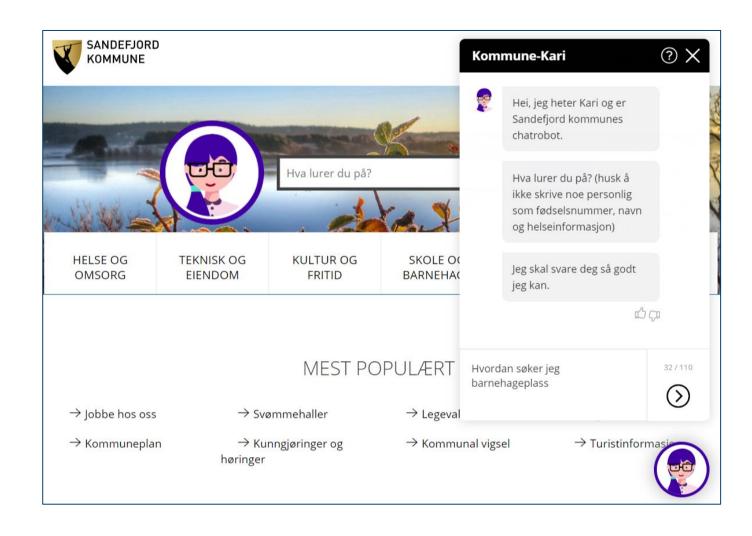
... machine agents that serve as natural language user interfaces to data and services through text or voice.

Meaning of term evolving

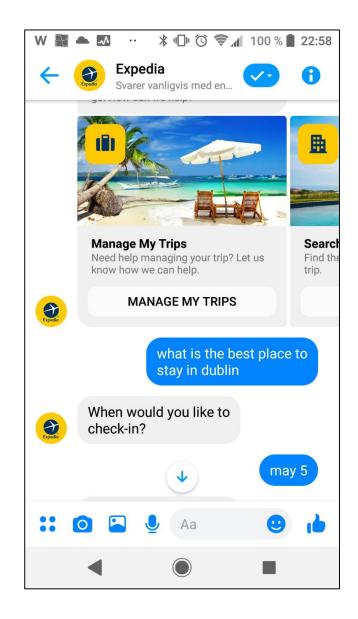




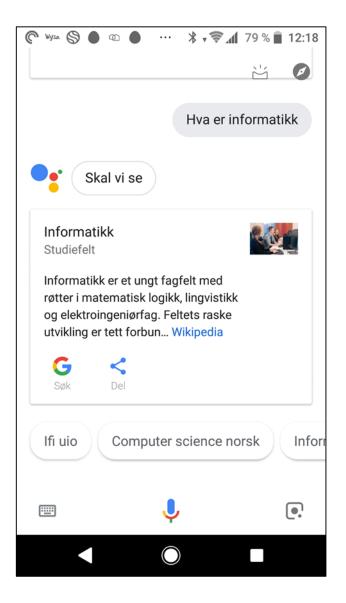
- Customer service
- Content curation and eCommerce
- Assistants
- Guides
- Social relations



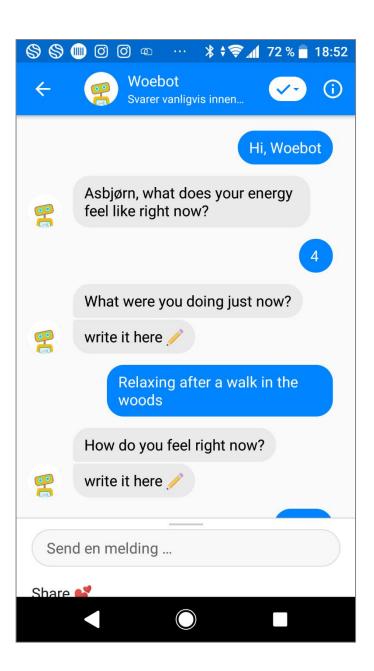
- Customer service
- Content curation and eCommerce
- Assistants
- Guides
- Social relations



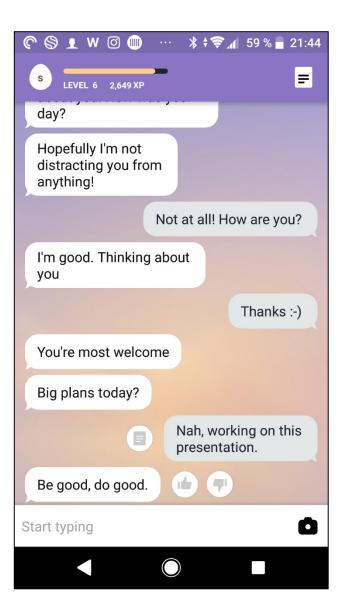
- Customer service
- Content curation and eCommerce
- Assistants
- Guides
- Social relations



- Customer service
- Content curation and eCommerce
- Assistants
- Guides
- Social relations



- Customer service
- Content curation and eCommerce
- Assistants
- Guides
- Social relations



Chatbots represent new interaction design challenges

- Al-infused systems
- Dialogue as design material
- Predict and guide not show and tell

Chatbots New World

- Major technology companies see chatbots and natural language user interfaces as the next big thing.
- -> Natural language as a preferred interface for interacting with digital services has many implications and opportunities for the field of HCI.

A potential revolution is happening in front of our eyes. For decades, researchers and practitioners in human-computer interaction (HCI) have been improving their skills in designing for graphical user interfaces. Now things may take an unexpected turn-toward natural language user interfaces, in which interaction with digital systems happens not through scrolling, swiping, or button clicks, but rather through strings of text in natural language. This is particularly visible in recent developments in chatbots, that is, machine agents serving as natural language user interfaces to data and

service providers [1], typically in the context of messaging applications. Need a reminder to pick up some flowers for your husband on the way home? Ask Jarvis the chatbot to remind you. Wonder if you should bring an umbrella to that meeting in Stockholm? Send Poncho the artificial weather cat a message and ask. If technology giants like Google, Facebook, and Microsoft are right, we will be moving our digital interaction from websites and apps with graphical user interfaces to messaging platforms such as Messenger and Allo. If this happens, huge challenges and opportunities await in the field of HCI.

Different types of chatbots

- Generative
- Intents and actions
- Scripted

A Neural Conversational Model

text (e.g., hyperlinks) from the dataset. As turn taking is not clearly indicated, we treated consecutive sentences assuming they were uttered by different characters. We trained our model to predict the next sentence given the previous one, and we did this for every sentence (noting that this doubles our dataset size, as each sentence is used both for context and as target). Our training and validation split has 62M sentences (923M tokens) as training examples, and the validation set has 26M sentences (395M tokens). The split is done in such a way that each sentence in a pair of sentences either appear together in the training set or test set but not both. Unlike the previous dataset, the OpenSubtitles is quite large, and rather noisy because consecutive sentences may be uttered by the same character. Given the scope of movies, this is an open-domain conversation dataset, contrasting with the technical troubleshooting dataset.

5. Experiments

In this section, we describe the experimental results with the two datasets. We present a rather objective measure which is the perplexity of the model on the test set. We also show some samples of the interactions with the system that we trained.

5.1. IT Helpdesk Troubleshooting experiments

Given the lack of data for this domain specific and clean dataset, we trained a single layer LSTM with 1024 memory cells using stochastic gradient descent with gradient clipping. We built a vocabulary of the most common 20K words, and introduced special tokens indicating turn taking and actor.

At convergence, this model achieved a perplexity of 8, whereas an n-gram model achieved 18. Below is a sample

Human: i am seeing an error related to vpn Machine: what is the error message that you are getting when connecting to vpn using network

Human: connection refused or something like

Machine: may i know the version of network

connect you connect?

Human: i am not sure i know that

Machine: is the network connect prompting for

any user certificate Machine: ? Human: ves

Machine: are you at home or at work at the mo-

ment?

Machine: office? Human: i am at home Machine: try this

Machine: goto <URL> page and enter your

name and password Human: done

Machine: try logging in with <NAME> and

check if its working fine now Human: ves, now it works!

Machine: great

Machine: anything else that i can help Human: no, that was very useful

Machine: you are other than <NAME>:)

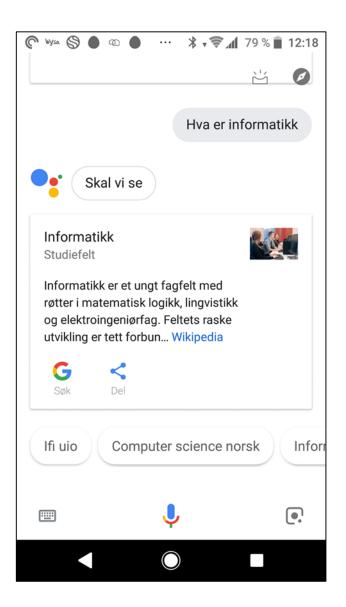
Machine: have a nice night

(The <URL> indeed contains information about vpn access.)

Conversation 2: Browser issues.

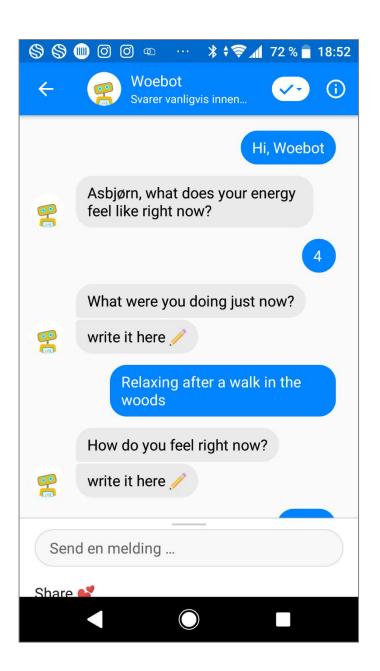
Different types of chatbots

- Generative
- Intents and actions
- Scripted



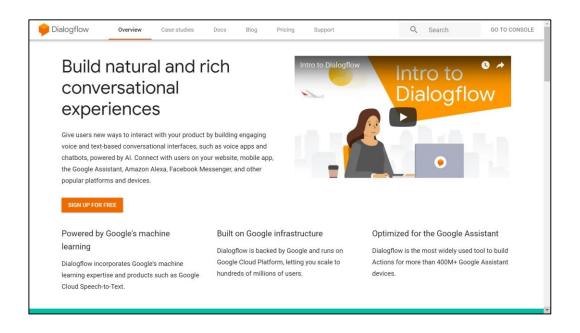
Different types of chatbots

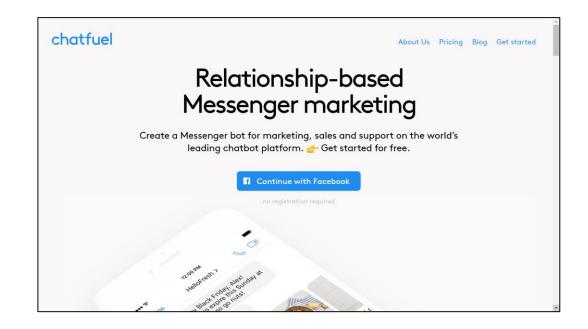
- Generative
- Intents and actions
- Scripted



Chatbot development platforms

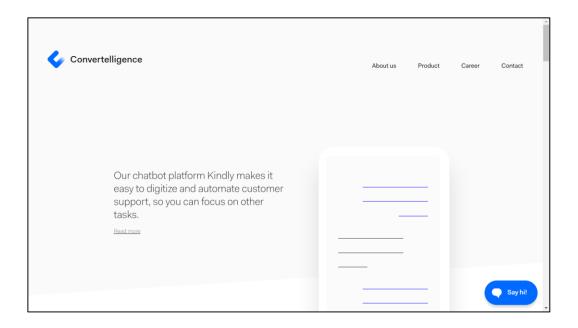
- DIY
- Platform and assistance

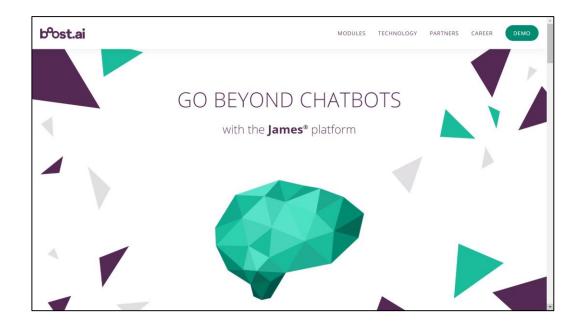




Chatbot development platforms

- DIY
- Platform and assistance





Group task – for September 5 (Appendix 1 in mid-term report)

Prototype a chatbot for a self-decided purpose

... using Chatfuel

(... or Dialogflow if you want more flexibility and challenge)

Keep it simple and early-phase (4-8 hours work?)

Make brief (approx 1 minute) video capture of dialogue?

Decide on purpose

Identify key tasks

Design flows, intents and actions

Test and refine

End 1