

KUNSKAPSKRYSSNING 2019

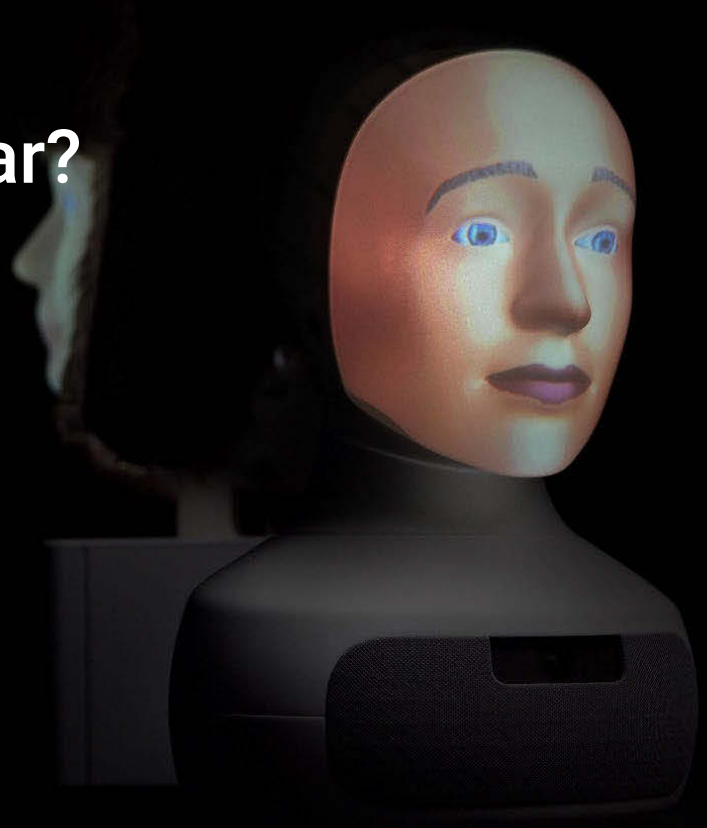


Hur kan vi prata med robotar?

Gabriel Skantze

Professor i Talteknologi, KTH

Medgrundare, Furhat Robotics





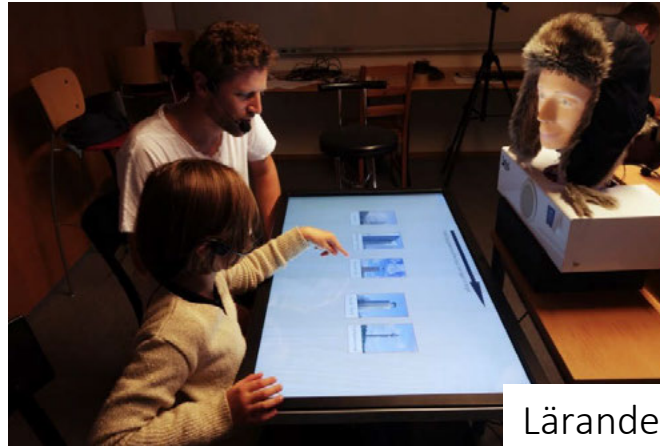
Hälsa



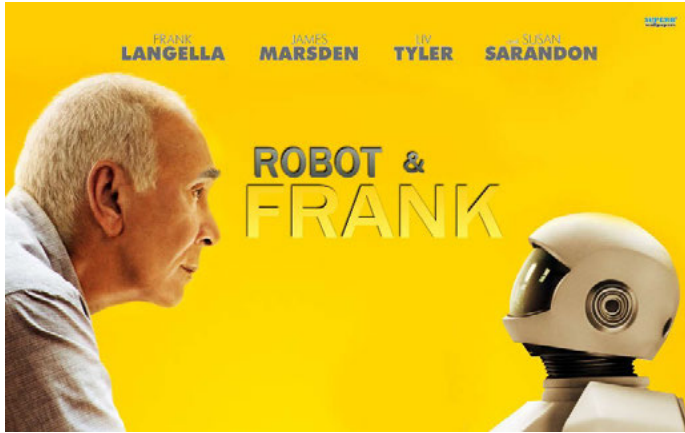
Publika miljöer

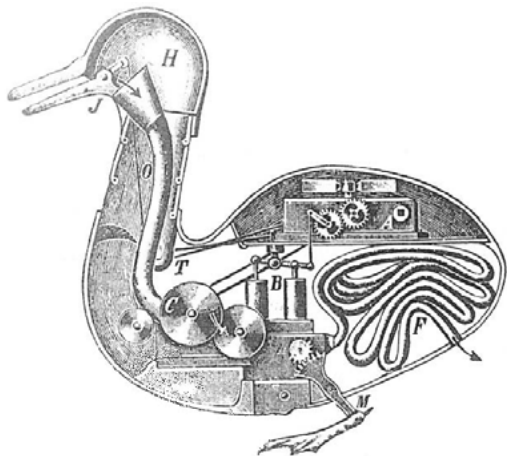


Tillverkning



Lärande



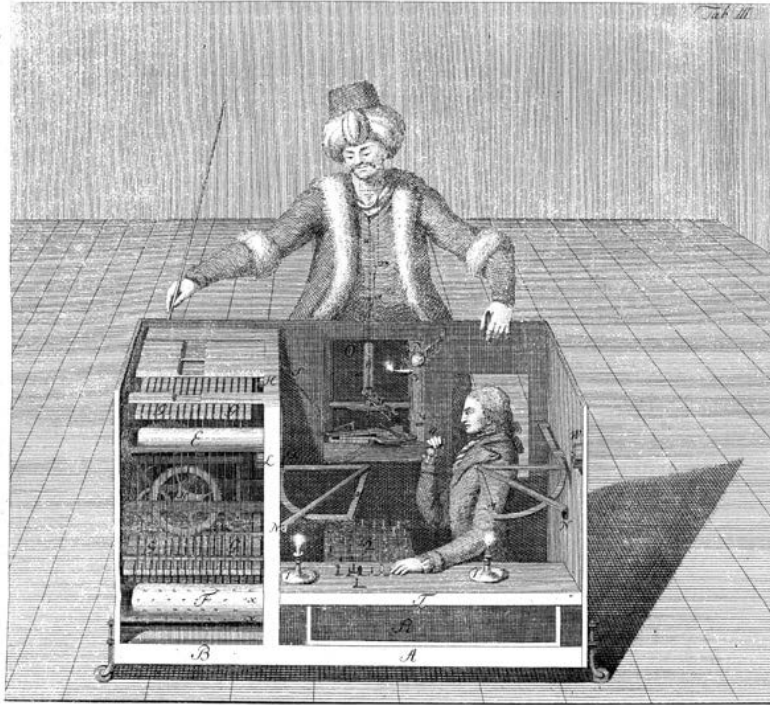


Jacques de Vaucanson (1709-1782)



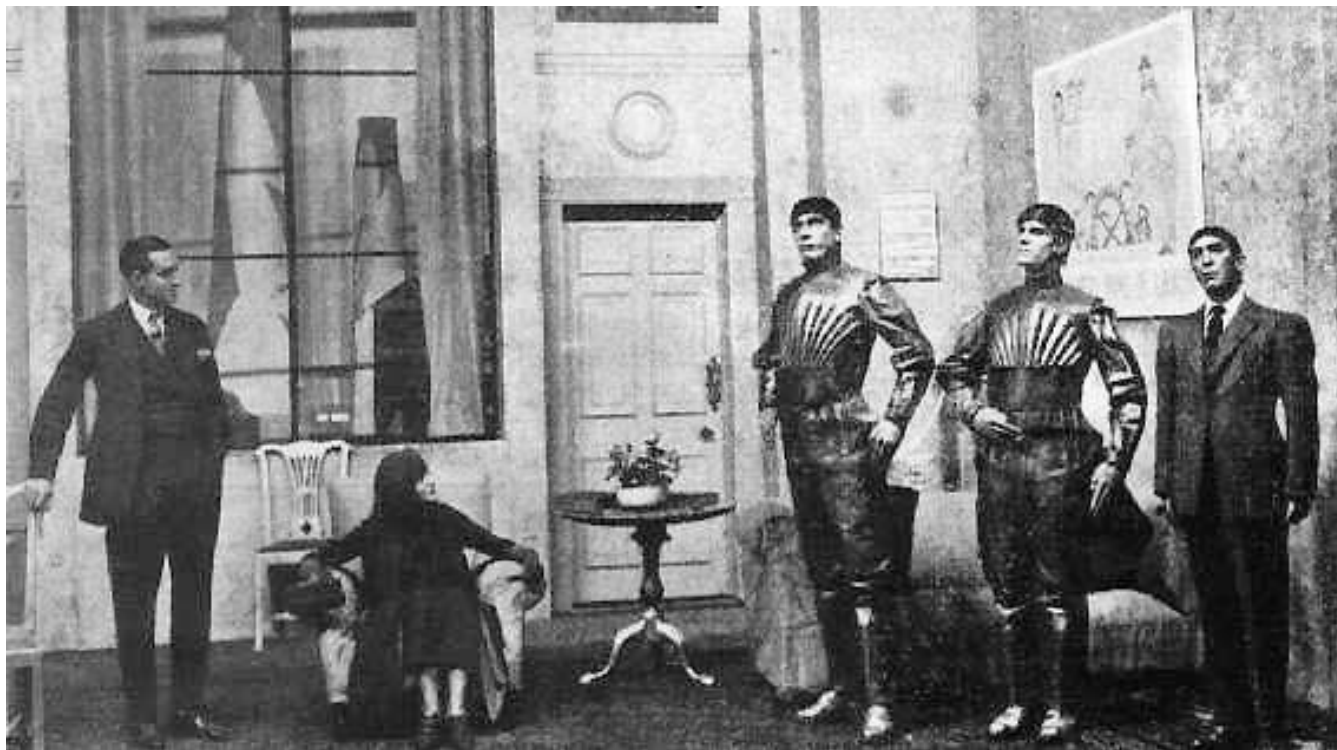
Pierre Jaquet-Droz (1721-1790)



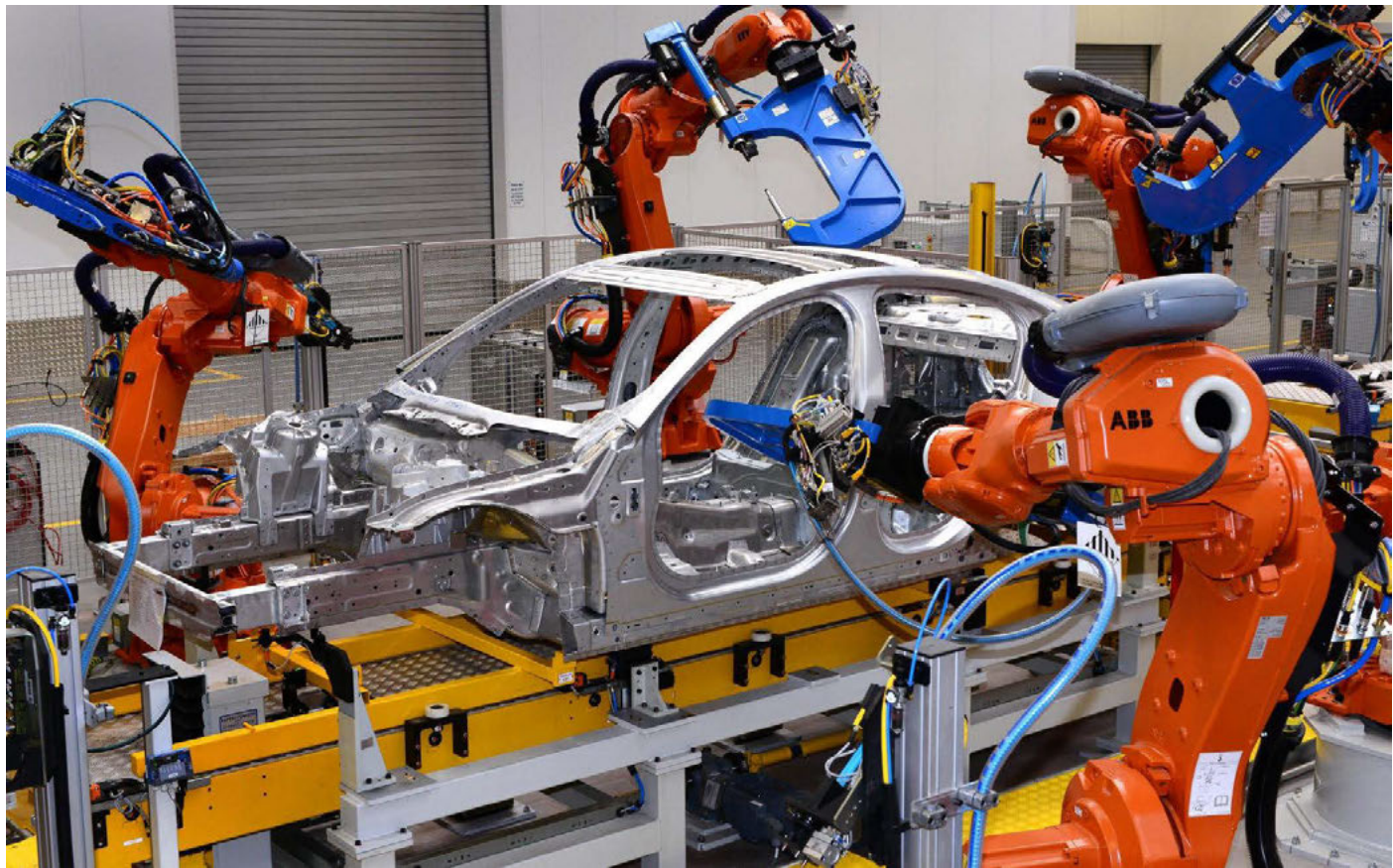


"The Mechanical Turk"
Wolfgang von Kempelen (1770)





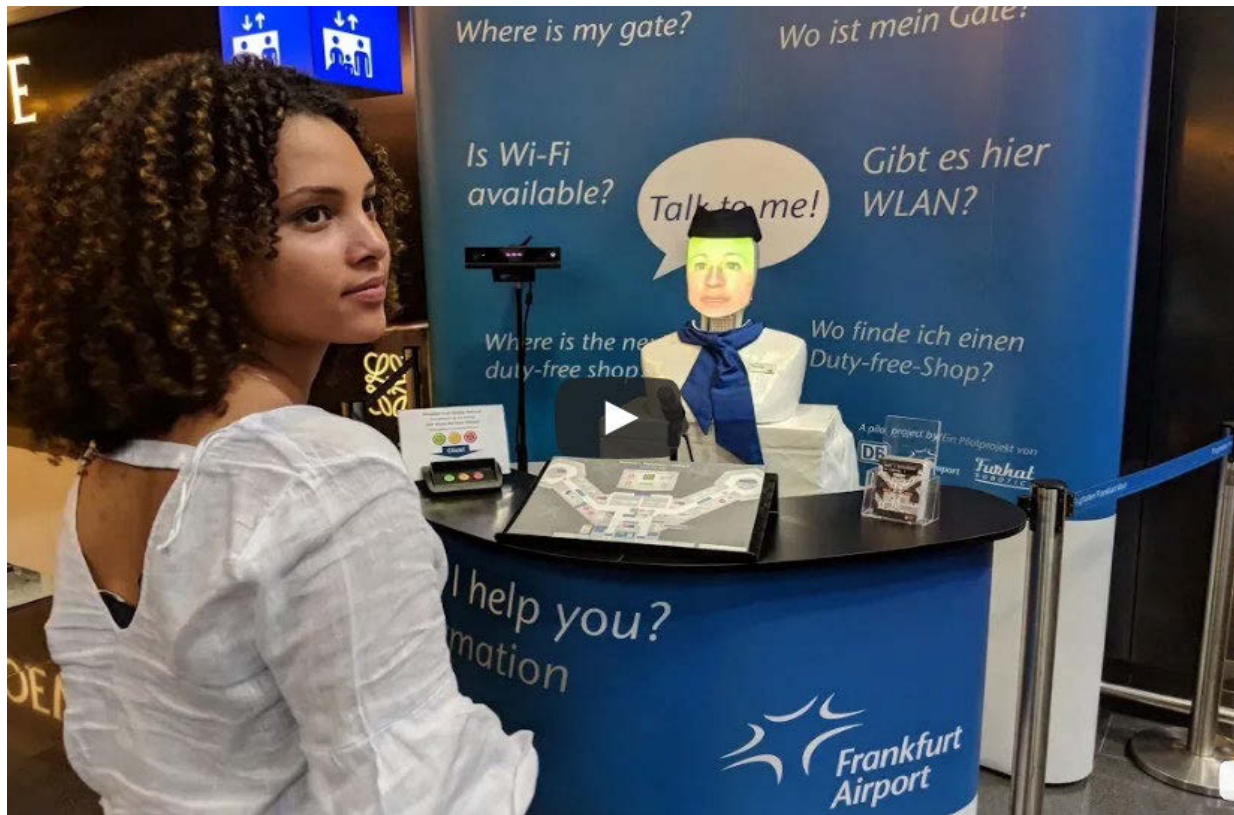
Rossum's Universal Robots av Karel Čapek (1920)









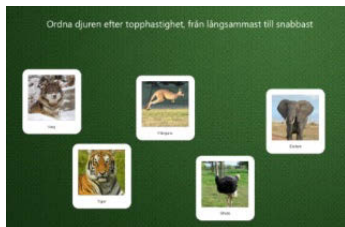
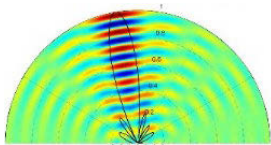




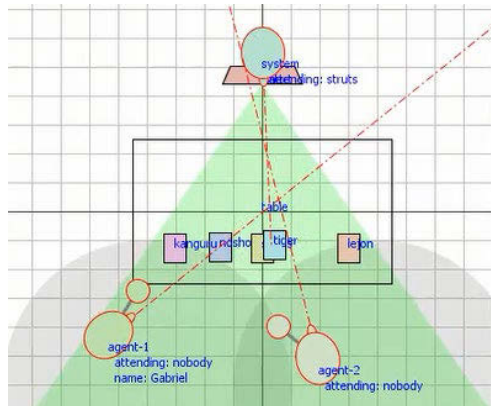
En social robot måste kunna:

- Se möjliga samtalspartner
- Tolka deras ansiktsuttryck, blickar och gester
- Förstå om de vill prata med roboten
- Förstå när roboten blir tilltalad
- Förstå vad som sägs
- Veta vad den ska säga
- Veta när den ska prata
- Syntetisera talet (med rätt uttal och prosodi)
- Veta var den ska titta
- Skapa rätt ansiktsuttryck och läpprörelser

Sensorer



Situationsmodell



Var är användarna?
Vart riktar de sin uppmärksamhet?
Vem är det som pratar?
Var finns det föremål?

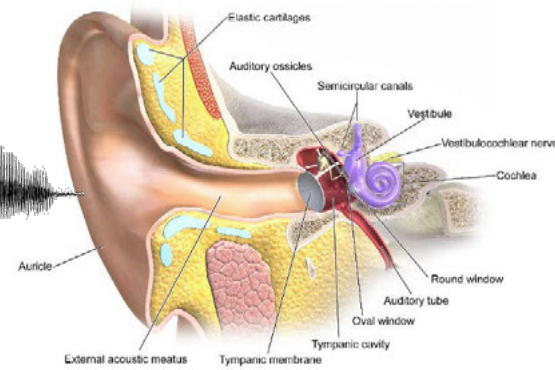
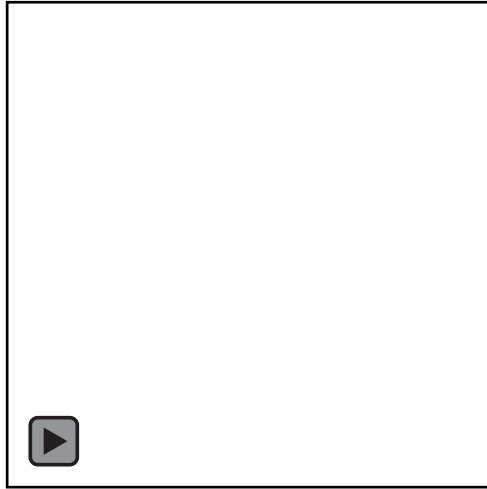
Beslutsmodell



Var ska roboten titta?
När ska roboten prata?
Vad ska roboten säga?
Hur ska den uttrycka sig?



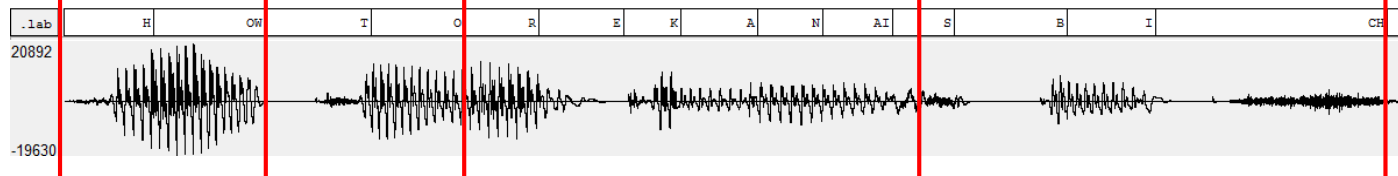
Varför är talat språk svårt för datorer?



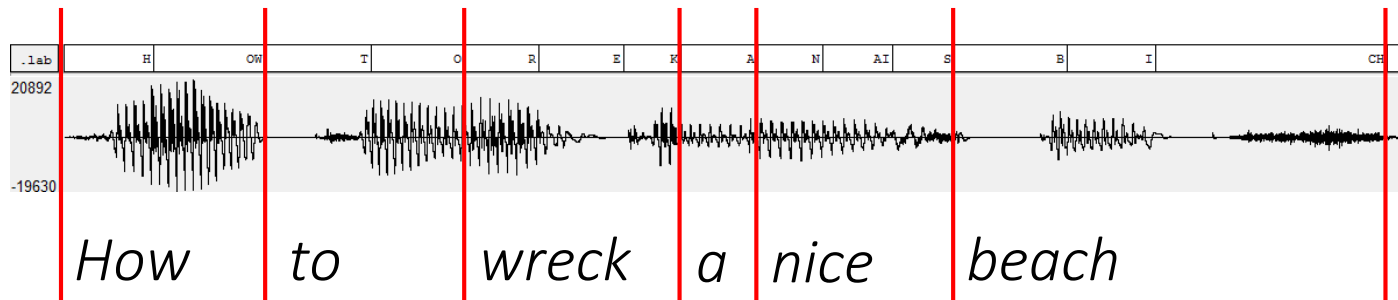


Skrivet språk	Talat språk
Använt sedan 5000 år	Använt minst 100.000 år
Ord, bokstäver, mellanrum, skiljetecken	Fonem och prosodi
Asynkron kommunikation	Kommunikation i realtid
Syntaktiskt välformat	Disfluent (Uppprepningar, tvekljud, avbrutna ord, etc)
Uteslutande symboliskt (<i>vad</i> vi säger)	Icke-symboliska komponenter (<i>hur</i> vi talar: prosodi, skratt, andning, etc)

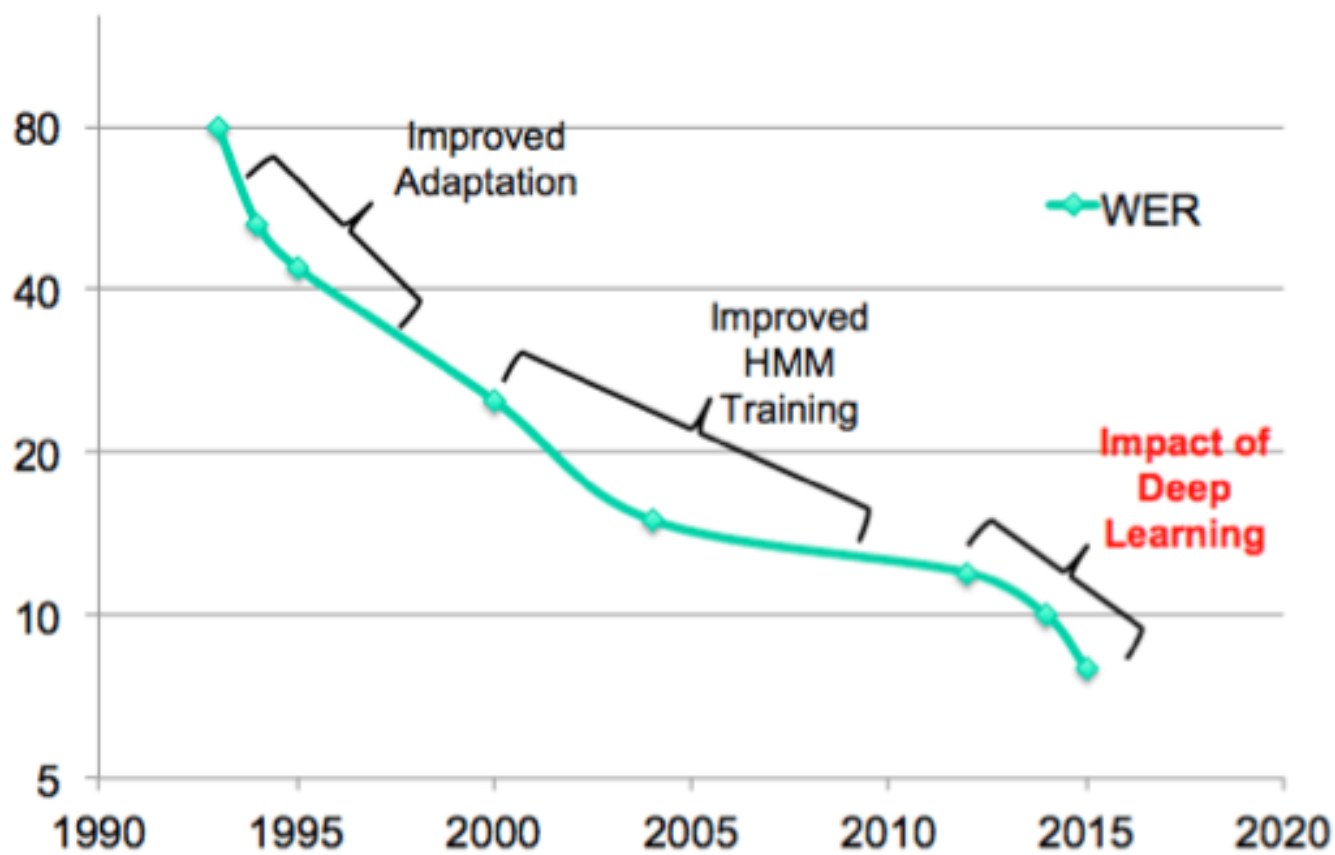
How to recognize speech

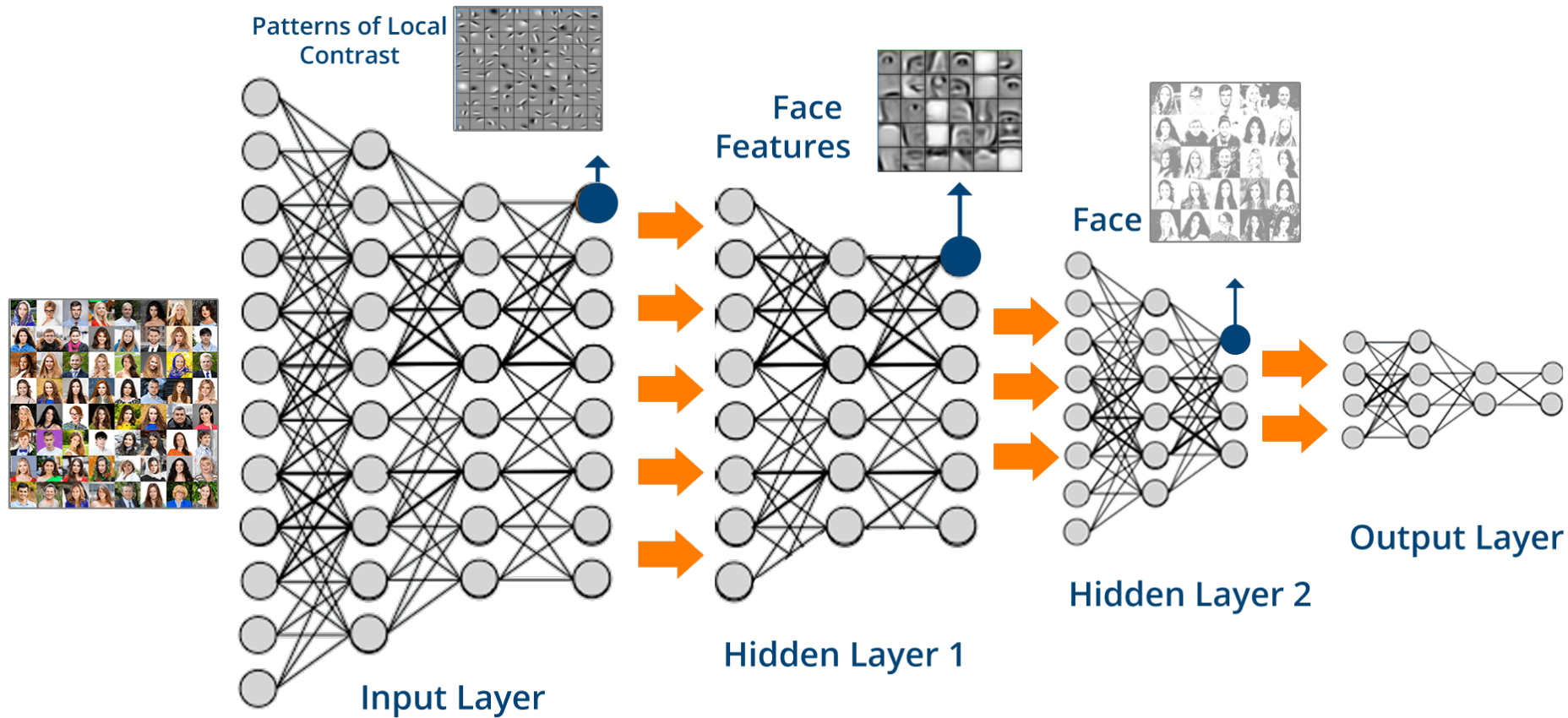


How to wreck a nice beach



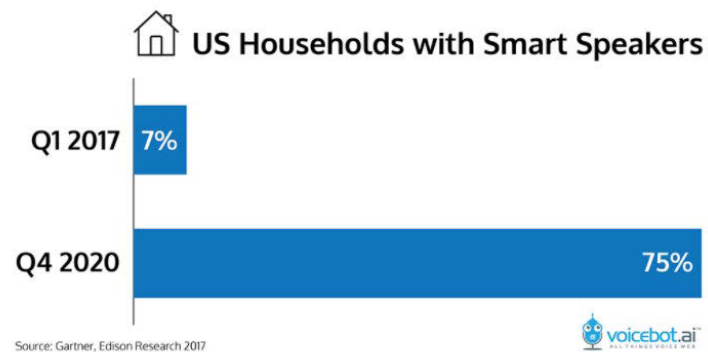
Word Error Rate (WER)

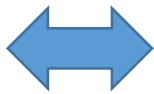
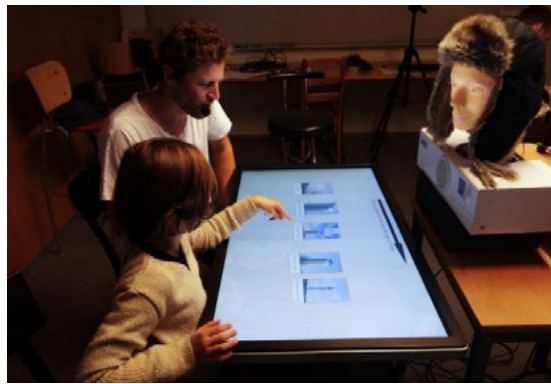
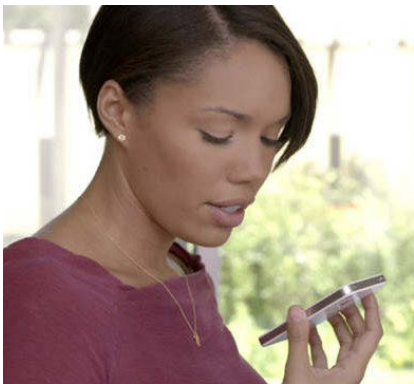




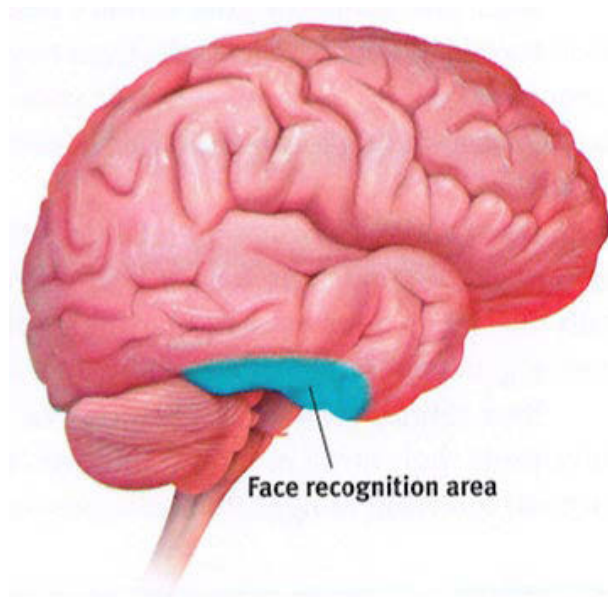
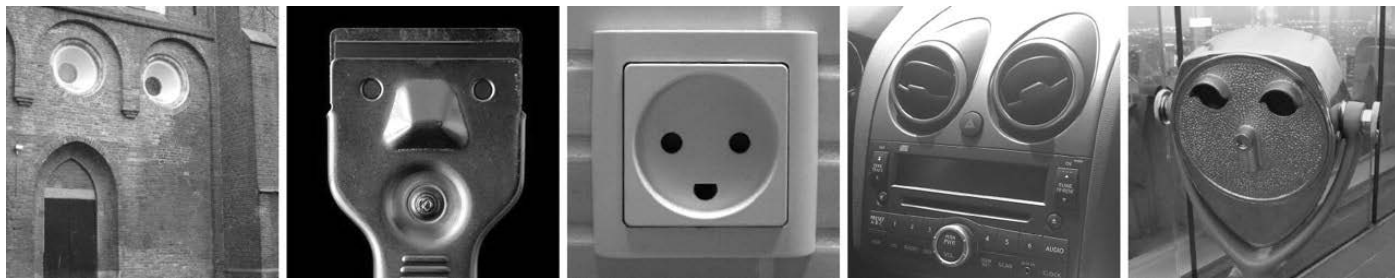


Smarta högtalare











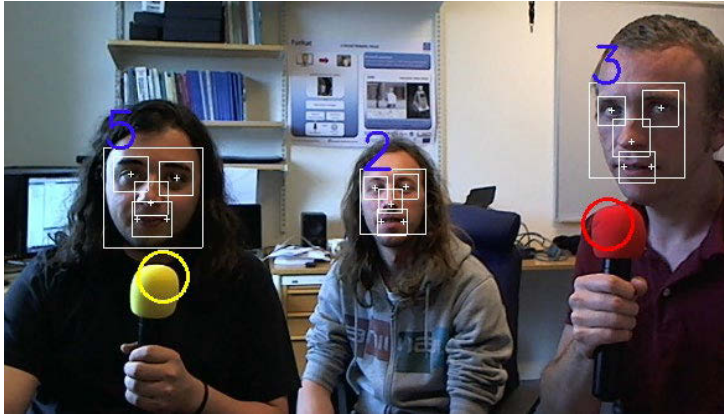
McGurk-effekten

“Ba” eller “Da” eller “Ga”?

Vad ansiktet tillför samtalet

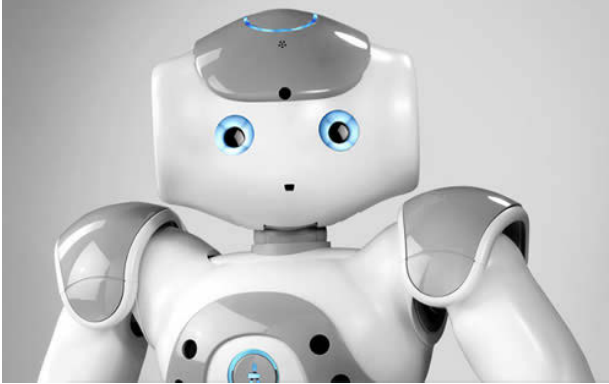


- Identifiering av samtalspartner
- Uppmärksamhet (blick)
- Ansiktsuttryck (känslor, återkoppling)
- Läppläsning



Att ge maskinen ett ansikte

NAO



Jibo

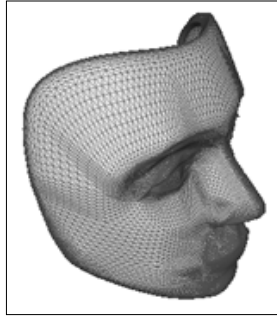


Sophia

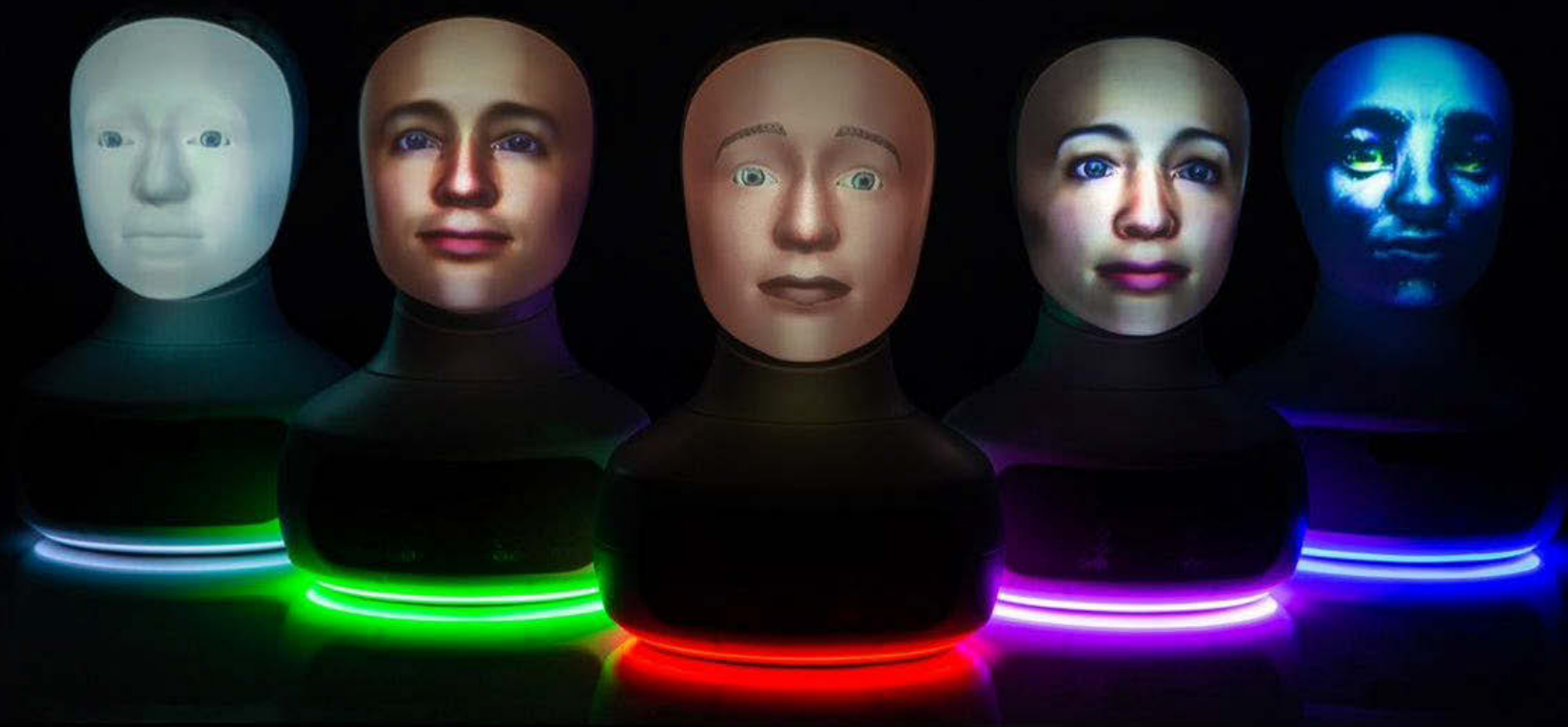


Animerade agenter

Furhat – ett ansikte projicerat på en mask



 Furhat Robotics





Furhat för anställningsintervjuer



BLI FÖRST I VÄRLDEN ATT INTERVJUAS AV EN SOCIAL REKRYTERINGSROBOT

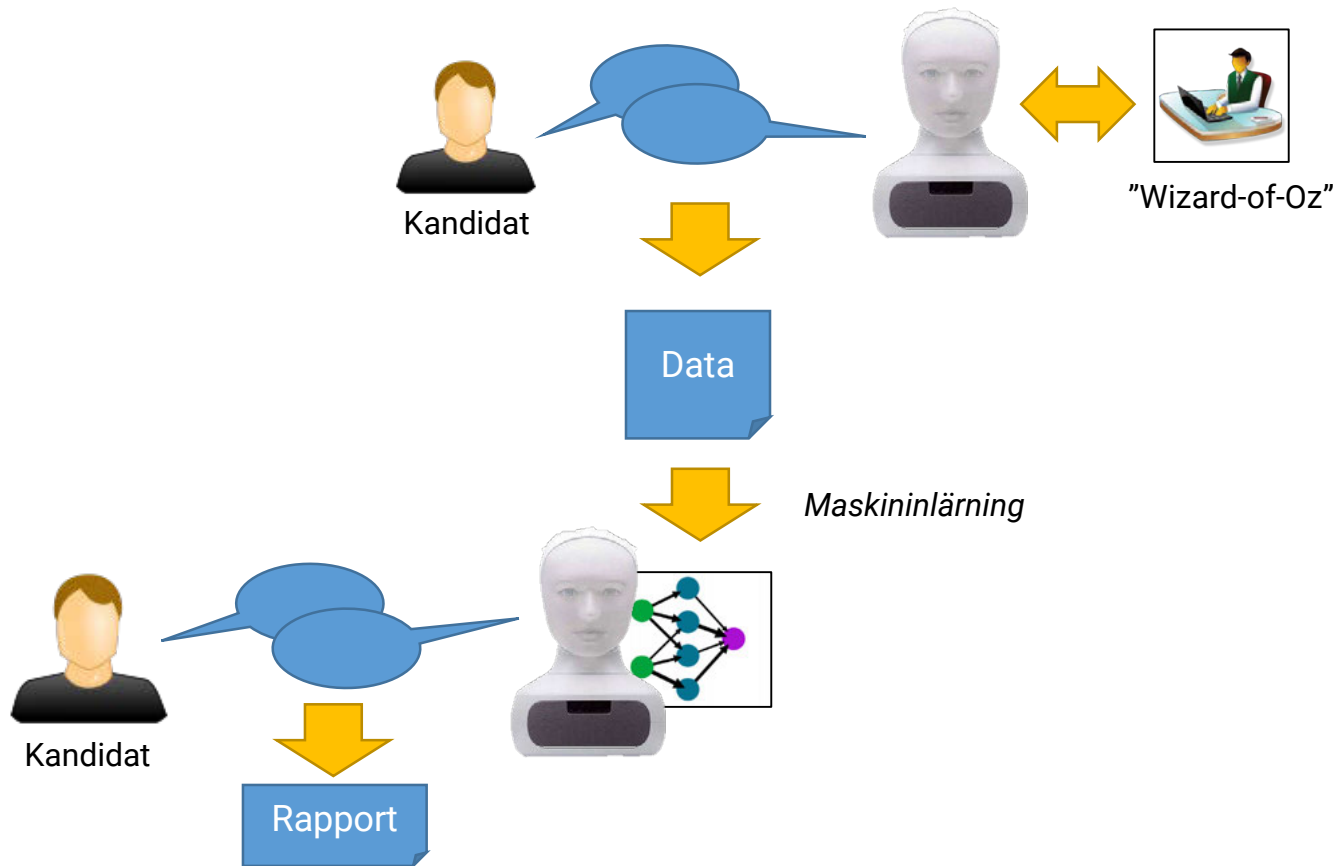
Nu drar vi igång arbetet med att träna upp Tengai – vår nya fördomsfria rekryteringsrobot. Vi behöver nu dig som vill vara med och bidra till en hållbar arbetsmarknad och forskning kring fördomsfri rekrytering! Vad du behöver göra? Träffa Tengai och bli intervjuad. Ansökan hittar du längst ner på den här sidan!

Tengai Unbiased är TNG:s och Furhat Robotics nya rekryteringsrobot. Tillsammans utvecklar vi en fördomsfri social AI-robot som ska assistera våra rekryterare i deras arbete. Tengai är en fysisk robot med mänskliga drag, inte en chatbot som ofta får benämningen rekryteringsrobot inom HR-världen.

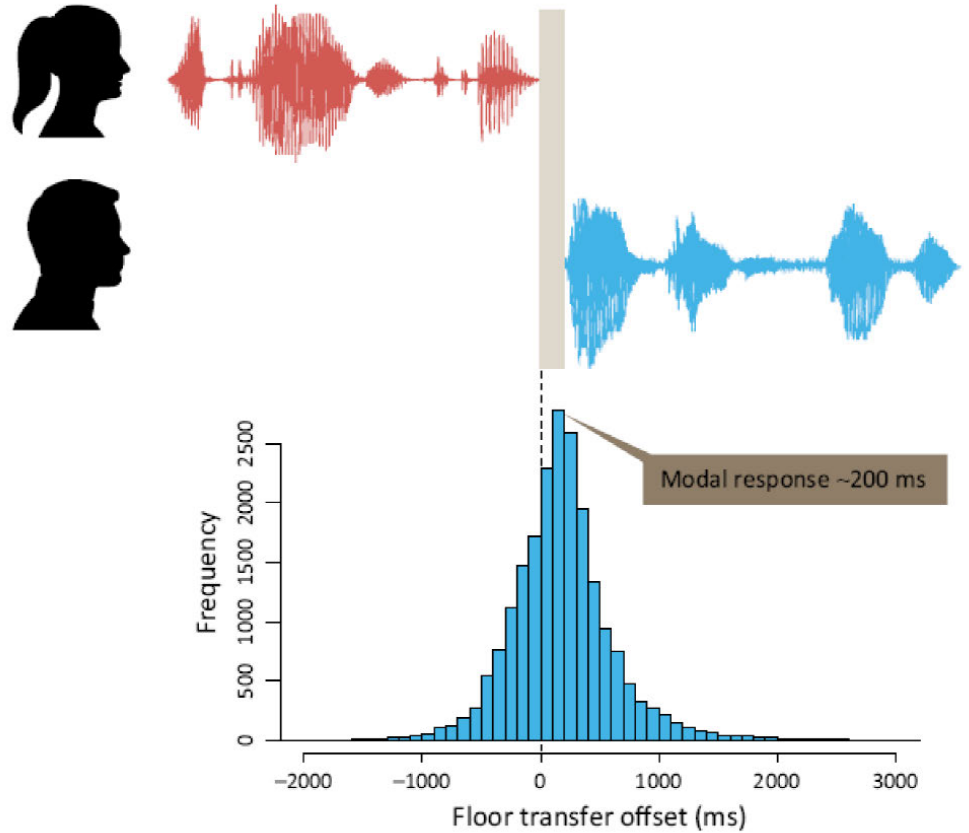
Tanken är att Tengai ska bidra till att göra TNG:s rekryteringsprocess ännu mer rättvis – objektiv oavsett sökandes ålder, kön, ursprung och utseende. Och på sikt också fördomsfritt kunna hantera sökande med funktionsnedsättningar som t ex ADHD och autism (neurodiversity). Sökande som, tyvärr omedvetet, ofta väljs bort av mänskliga rekryterare.

Med din hjälp vill vi nu integrera TNG:s kompetensbaserade och fördomsfria rekryteringsmetodik i Tengai. Vi på TNG kommer lära hen att intervjua fördomsfritt och behöver personer som vill vara med i det här unika forsknings- och





Turtagning i samtal



Hur koordineras turtagning i samtal?



Signal	Ge turen	Hålla turen
Syntax	Avslutad fras	Oavslutad fras
Intonation	Stigande eller fallande	Platt
Intensitet	Lägre	Högre
Duration	Kortare	Längre
Blick	Titta på samtalspartnern	Titta bort
Gester	Avslutad	Ej avslutad

Ju fler ledtrådar, desto starkare signal! (*Duncan, 1972*)

När ska Furhat ta turen?

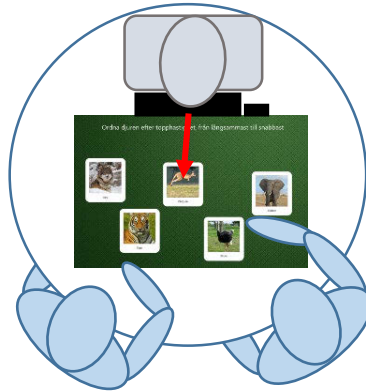


Furhat lämnar över turen

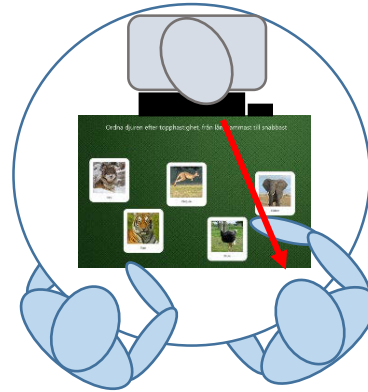
Tittar på båda



Tittar på korten

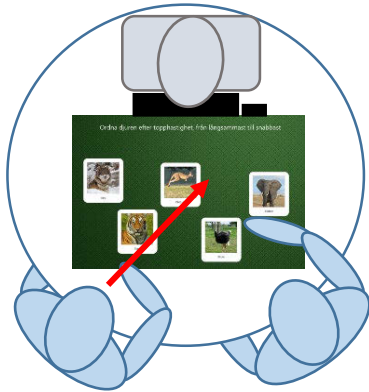


Titta på en användare



Användaren lämnar över turen

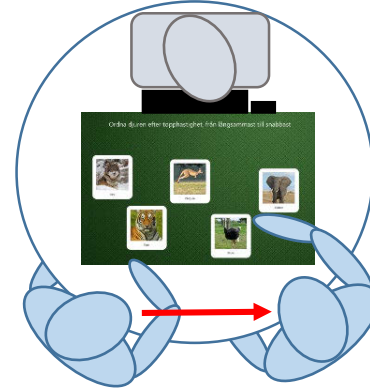
Tittar på korten



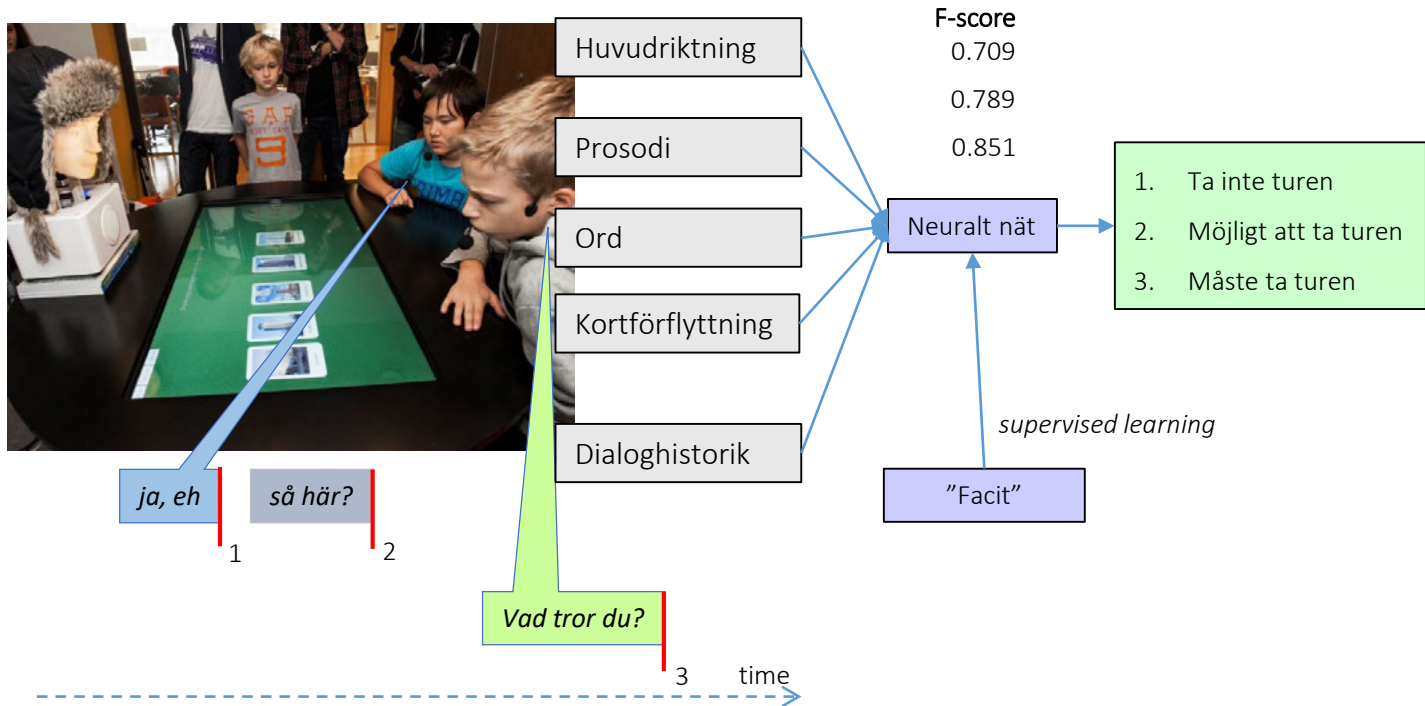
Tittar på Furhat



Tittar på den andre talaren



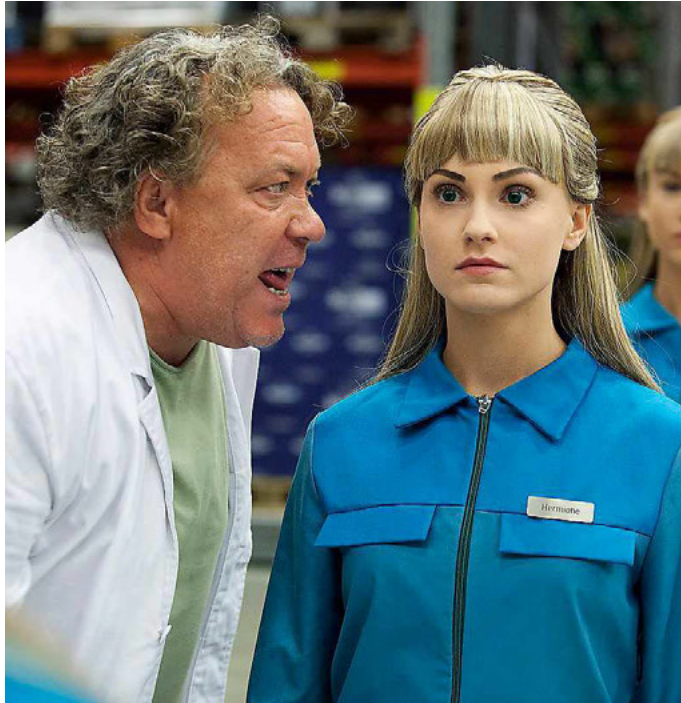
Maskininlärning för turtagning



Vanliga frågor

- Varför ska vi prata med robotar istället för med varandra?
- Kommer robotarna att ta våra jobb?
- Kommer robotarna bli självmedvetna och ta över världen?
- Hur ska vi kunna lita på robotarnas beslut?
- Ska vi bete oss mot robotarna som om de vore människor?

Hur ska vi bete oss mot robotar?



"Animals are not self-conscious and are there merely as a means to an end. The end is man."

"If any acts of animals are analogous to human acts and spring from the same principles, we have duties towards the animals because thus we cultivate the corresponding duties towards human beings."

Immanuel Kant

Amazon Alexa to reward kids who say: 'Please'

© 25 April 2018

f b t e Share



Tack!

Gabriel Skantze
skantze@kth.se

