IN5620
Første modul og første iterasjon:
Individuell innlevering

1.1 Konsepter, definisjoner, og historie rundt interaksjon med AI- og autonome systemer

1.1.1 Hvordan AI ble til

Første gangen begrepet «*artificial intelligence*» (AI) ble brukt var av John McCarthy i 1956 (Cordeschi, 2007). Mye av hans forskning rundt AI har vært med på å definere, og utvikle, den formelle teorien vi har om AI i dag.

Selv om bruken av begrepet AI blir antatt å for første gang å ha blitt brukt da, så har konseptet om kunstig intelligente roboter vært lenge. Forfatteren bak «The Wonderful Wizard of Oz», L. Frank Baum, skrev i 1907 om «mekaniske menn» som tenker, snakker, og gjør alt utenom å være i livet (Buchanan, 2005). En annen viktig person å nevne er Alan Turing. Han utforsket den matematiske muligheten for å skape AI, og i 1950 foreslo han en definisjon som bestemte om en programvare var intelligent eller ikke (Benko & Lányi, 2009).

1.1.2 Definisjoner på AI

Definisjon 1 – John McCarthy, 2007 (først brukt under en AI-konferanse i 1956):
«It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable» (McCarthy, 2007).

Definisjon 2 – Bedriften PwC, 2017 (en del av et forbrukermagasin):
«Defines technologies emerging today that can understand, learn, and then act based on that information» (Series, 2017).

Definisjon 3 – Bratteteig & Verne, 2018 (i relasjon til deltakende design):

«AI is a subfield of computer science aimed at specifying and making computer system that mimic human intelligence or express rational behaviour» (Bratteteig & Verne, 2018).

Egen definisjon:
«AI er maskiner/systemer som handler, tar valg, og justerer seg, basert på kontinuerlig læring og erfaring».

Jeg er veldig enig i definisjonene over der det forklares at AI kan forstå og lære på bakgrunn av erfaring og informasjon den får inn. Dette ville jeg selv få inn i min egen definisjon, og at AI også hele tiden kan justere seg basert på dette.

1.1.3 Menneskelig autonomi og Maskin autonomi – Automatisk og Autonomt

I filosofien blir ordet autonomi brukt om å følge sin egen indre lovmessighet og bestemme over seg selv, uten påvirkning fra andre (Dworkin, 1988). Slik jeg forstår det, kan en likhet mellom menneskelig- og maskin autonomi være at begge tar valg basert på tilgjengelig informasjon man har fått tilgang på tidligere. Forskjellen er bare at informasjonen til maskinen har blitt programmert av en utvikler, mens menneske sin informasjon er basert på livshendelser.

Konseptet automatisk blir ofte brukt om maskiner som gjøre noe uavhengig av menneskelig kontroll. På denne måten kan man kanskje se på automatisk som maskinenes versjon av å være autonome.

1.1.4 Kort sammendrag av artikkel fra pensum

Jeg har valgt å ta for meg artikkelen til Dautenhahn (2018): «Some Brief Thoughts on the Past and Future of Human-Robot Interaction». Forfatteren deler tanker om Menneske-Robot interaksjon (HRI). Som hvilke signifikante forandringer som har skjedd i dette feltet i løpet av de siste 20 årene. Dautenhahn (2018) skriver hvorfor hun er bekymret for HRI, som at det har vært problematisk å vite hvordan HRI-forskning burde bli fasilitert, analysert, og nedskrevet. Det blir også nevnt hvorfor hun er optimistisk for HRI, og at det er avgjørende å holde diskusjonen om hvordan

man gjennomfører HRI-forskning i live. Det er viktig å være åpne for å bruke den metodologien som passer best for sin studie, og å bruke metodologier fra flere felt enn eksperimentell psykologi.

1.1.5 Bedrift som jobber med AI

Microsoft er en stor bedrift som er kjent for sitt arbeid med AI. Dette er heller ikke noe de prøver å skjule, da de har dedikert flere sider på deres nettside om AI. De deler opp AI i ulike formål: «for verden», «for helsa», «for tilgjengelighet», «for kulturell arv», og «for humanitær handling». Med andre ord, de snakker om å bruke AI for å skape en bedre verden. Dette gjør de ved å tilby programmer, midler, og AI-teknologi. De tilbyr AI-tjenester pakket inn i en større økologi/ideologi som de selv har skapt.

1.1.6 Media som ser på interaksjon mellom AI og mennesker

Black Mirror er en TV-serie fra Netflix, der hver episode handler om en alternativ verden der dystopisk teknologi og AI er integrert i hverdagen til menneske. I en av episodene følger man en kvinne som mister sin kjæreste. Hun velger å benytte seg av en AI-tjeneste som lager en digital versjon av hennes avdøde kjæreste basert på hans sosiale medier, og det går så langt at hun bestiller en fysisk robotversjon av kjæresten. Denne robotversjonen ser akkurat ut som den ekte, det eneste som skiller dem er at robotversjonen gjør det den blir bedt om – uansett. Dette gjør kvinnen frustrert, da hun begynner å forstå at robotversjonen ikke kan erstatte den ekte – selv om den ser og føles helt lik ut. Den er bare ikke autonom nok til å erstatte den ekte mannen.

1.2 Robot og AI-systemer

1.2.1 Ordet «Robot» sin opprinnelse

Ordet «robot» ble først introdusert av den tsjekkiske forfatteren Karel Capek i hans forestilling R.U.R. («Rossum´s Universal Robots») (Wallén, 2008).

1.2.2 Definisjoner på robot

Definisjon 1 – Britannica, 2022:

«Any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner» (Moravec, 2022)

Denne definisjonen forklarer hva som skiller robot fra menneske, men at dens hovedfunksjon er å erstatte menneskelig arbeid.

Definisjon 2 – Oxford Reference:

«A machine resembling a human being and able to replicate certain human movements and functions automatically» (Reference, 2022).

Definisjonen her fokuserer på at en robot kan kopiere mange menneskelige funksjoner automatisk.

Egen definisjon:

«En maskin som kan utføre både enkle og kompliserte oppgaver automatisk. Må ikke nødvendigvis se ut eller oppføre seg som et menneske, men er ofte med på å minske menneskelig arbeid»

Jeg ville ikke ta med at en robot erstatter menneskelig arbeid, men heller at den kan være med på å minske den. Automatisering er ord som jeg også ville ha med i min egen definisjon da jeg mener det er en sentral del av en robot.

1.2.3 Relasjonen mellom AI og roboter

AI er kunstig intelligens som lærer av erfaring, og hele tiden justerer sine egne valg på bakgrunn av læringen. En robot lærer ikke på samme måte som AI, da en robot er programmert til å utføre forhåndsbestemte, spesifikke oppgaver.

1.2.4 Robotgressklipperen

En type robot man ser oftere hos en typisk husstand med hage, eller bedrifter med uteområder, er robotgressklipperen. Å klippe gresset er for mange en kjedelig oppgave, og det er nok derfor robotgressklippere har blitt så populære. Den

beveger seg i avgrensede områder som er kartlagt på forhånd av brukeren, og de har blitt såpass smarte at de krever minimalt med menneskelig interaksjon. «Self-docking» gjør at maskinene selv vet når de må lades, og beveger seg da mot ladestasjonen sin. Mange har også regnsensorer som gjør at de automatisk går i dvale når det begynner å regne. Robotgressklipperen er et godt eksempel på hvordan roboter kan minske arbeid for mennesker som har behov eller ønske for det.

1.3 Universelt design og AI-systemer

1.3.4 Hva er universelt design (UU)?

«The design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design» (Persson, åhman, Yngling, & Gulliksen, 2015).

Denne definisjonen passer på å ta hensyn til at noens behov kan skape barrierer for andre, og det derfor er viktig å designe universelt design i *størst mulig grad*. Det er viktig å tenke på UU som et krav man stiller til produkter og tjenester - ikke et ønske. UU er med på inkludering, slik at alle kan bruke og ha nytte av tjenesten.

1.3.2 Potensialet til AI i forhold til mennesket

AI har et stort potensial med å være en hjelpende hånd for mennesket. Et eksempel er språk-til-tekst/tekst-til-språk teknologien. Amazon sin *smart speaker*, Alexa, gjør det enkelt for brukeren å snakke og be om ulike tjenester ved bruk av tale. Dette er spesielt nyttig for de med motorikk- eller synsvansker. Mens stemme-systemet Voiceitt kan lære brukerens talemåte over tid, og oversette brukerens ord til klar og normalisert tale i både lydopptak og meldinger. Et godt verktøy for folk med språkvansker.

1.3.3 Potensialet til AI for inkludering/ekskludering av mennesker

Eksemplene i forrige avsnitt hvor hvordan AI kan inkludere, men det har også vist seg at AI kan ekskludere mennesker. AI bias oppstår, og dette gjelder spesielt for kvinner, mennesker som ikke er hvite, og mennesker med funksjonsnedsettelse (Whittaker et al., 2019). For eksempel har man sett at AI-systemer ikke «hører»

stemmer med høyere tone (feminine toner), diagnostiske systemer som fungerer dårlig på mennesker med mørk hud, og ansettelsesalgoritmer som prioriterer menns CV. Man ser hvordan utviklerne sitt bias skinner gjennom produktet.

1.3.4 Forstå og forståelse - «Forstår maskiner?»

Slik jeg oppfatter konseptet forståelse, så er det en mer dyptgående mental prosess av det å forstå noe. Å ha forståelse for noe gjør at man kan se noe ut ifra ulike kontekster, og drøfte rundt det. Dette er en menneskelig evne, som ikke nødvendigvis er noe roboter får til. En robot kan forstå en oppgave den er programmert til å gjøre, men har ikke en personlig forståelse for hvorfor den må gjøre oppgaven.

1.4 Guideline for Menneske-AI-interaksjon

1.4.1 Guideline 16: Convey the consequences of user actions

«Immediately update or convey how user actions will impact future behaviors of the AI system».

Dette betyr at AI skal være responsiv til brukerinput, og kunne veilede brukeren videre på basis av det.

1.4.2 HCI design guideline og Menneske-AI-interaksjon guideline

Nummer tre av Shneiderman (2004) sine åtte gylne regler, står det at man skal gi informativ tilbakemelding. Dette ser vi også er et viktig punkt i G16 til Microsoft. Microsoft sin G09 handler om å støtte effektiv korreksjon, og i den sjette gylne regelen blir det også nevnt at det er viktig å tillate enkel reversering av handling.

Schneiderman sine guidelines passer spesielt bra for evaluering av GUI, og er mer overordnet enn Microsoft sine. Allikevel ser man at det er likheter ved dem, og det viser at en god brukeropplevelse kan oppnås på samme måte i flere felt.

1.5 Referanser

Benko, A., & Lányi, C. S. (2009). History of artificial intelligence. In *Encyclopedia of Information Science and Technology, Second Edition* (pp. 1759-1762): IGI Global.

Bratteteig, T., & Verne, G. (2018). *Does AI make PD obsolete? exploring challenges from artificial intelligence to participatory design.* Paper presented at the Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial-Volume 2.

Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *Ai Magazine, 26*(4), 53-53.

Cordeschi, R. (2007). AI turns fifty: revisiting its origins. *Applied Artificial Intelligence, 21*(4-5), 259-279.

Dautenhahn, K. (2018). Some brief thoughts on the past and future of human-robot interaction. In (Vol. 7, pp. 1-3): ACM New York, NY, USA.

Dworkin, G. (1988). *The theory and practice of autonomy*: Cambridge University Press.

McCarthy, J. (2007). What is artificial intelligence?

Moravec, P. H. (2022). robot. Encylopedia Britannica. Retrieved from https://www.britannica.com/technology/robot-technology

Persson, H., åhman, H., Yngling, A. A., & Gulliksen, J. (2015). Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. *Universal access in the information society, 14*(4), 505-526. doi:10.1007/s10209-014-0358-z

Reference, O. (2022). robot. In: Oxford Reference.

Series, C. I. (2017). Bot. Me: A revolutionary partnership.

Shneiderman, B. (2004). Designing for fun: how can we design user interfaces to be more fun? *interactions, 11*(5), 48-50.

Wallén, J. (2008). *The history of the industrial robot*: Linköping University Electronic Press.

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., . . . Salas, M. (2019). Disability, bias, and AI. *AI Now Institute*.