

Individuell oppgave

1.1 Historien, definisjon og interaksjonen med kunstig intelligens

Kunstig intelligens kom til verden gjennom en rekke vitenskapelig prosjekter. Det startet med den britiske matematikeren Alan Turing, som lagde en maskin for å knekke kryptiske koder under andre verdens krig. Begrepet kunstig intelligens ble først introdusert i 1956 av den amerikanske matematiker og logiker John McCarthy (Grudin, 2009, s. 49).

Definisjoner om kunstig intelligent

Ifølge John McCarthy definerer han kunstig intelligens som "It is related to the similar task of using computers to understand human intelligence, but AI (artificial intelligence) does not have to confine itself to methods that are biologically observable." (McCarthy, 2004, s. 2).

En annen definisjon av Ziyad Saleh fra den britiske universitet i Egypt nevner at kunstig intelligens " is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals" (Saleh, 2019, s. 2).

Sist og ikke minst en definisjon fra den store norske leksikon, skrevet av Axel Tidemann "Kunstig intelligens er informasjonsteknologi som justerer sin egen aktivitet og derfor tilsynelatende framstår som intelligent" (Tidemann, 2020).

Ut ifra disse tre definisjonen, definerer jeg mitt synspunkt på kunstig intelligent. Kunstig intelligens er en datamaskin med ferdigheter i likhet med mennesker og dyr, ved å imitere og lære av handlinger. Jeg tar for meg at kunstig intelligent har tekniske sensorer som bidrar til å gi datamaskinen en form av informasjon. Dette kan da sees som læring, deretter kan datamaskinen benytte informasjonen til å kalkulere eller planlegge en tilbakemelding for oss mennesker.

Artikkel eksemplar om kunstig intelligent

Tone Bratteteig og Guri Verne har skrevet en interessant artikkel om å utforske vanskeligheter med kunstig intelligens innenfor deltagende design. Deltagende design er hvor flere brukere er involvert i å skape et design og utelukker unødvendige tekniske feil fra et brukers perspektiv (Bratteteig & Verne, 2018, s. 2). Videre sies det at kunstig intelligens er en trenings prosess som vil kreve tid, akkurat som å trene en hund. For eksempel å trene opp programvaren til å kunne se forskjellen på en bil og en motorsykkel, ved å benytte flere eksemplar bilder. Bratteteig og

Verne viser til tre utfordringer med kunstig intelligens teknologien innenfor deltagende design. Første utfordring for designere er å kunne vite hvilken effekt designet kan gi for framtiden. Den andre utfordringen er å kunne evaluere resultatet av design ideen innenfor prosjektets tidsramme, fordi kunstig intelligens vil kreve tid. Tredje utfordring er å kunne differensiere mellom bruk av designet og trenings prosessene av designet, så lenge kunstig intelligens er i bruk vil den være under trenings prosessen (Bratteteig & Verne, 2018, s. 4).

Kunstig intelligent produktet Birdsy

Birdsy er et kompani som har laget et videokamera for fugleelsker, og er til salg vinteren 2021. Nettside: <https://store.birdsy.com/>

På nettsiden viser Birdsy bruk av kunstig intelligens for opptak av fugler nær ditt nabolag. De reklamerer for at kunstig intelligent kameraet skal kunne markere opptaksvideoene slik at brukere kan lett bla gjennom opptakene og se de forskjellige fugleartene. Ved å navigere lengere inn på nettsiden, forklarer Birdsy i mer detaljer over hvordan kunstig intelligent gjennomføres i produktet. Kunstig intelligent blir brukt i to steg. Første steg er å gjenkjenne om en fugl eller dyr er innenfor kameravisningen for så å ta et opptak. Det vil også danne et miniatyrbilde ved å zoome inn på dyrene. Andre steget er å få Birdsy serveren til å gjenkjenne hvilke fugleart som er på bildet, og hvis Birdsy ikke klarer å gjenkjenne arten vil ikke fuglen bli identifisert. Hvis Birdsy identifiserer feil fugleart så skal brukeren kunne lett ta vekk eller legge til identiteten på sin brukerkonto. Kunstig intelligens klassifisering av fugleartene vil forbedres etter tid og det er Birdys mål å gjenkjenne nøyaktig de fleste arter i nær fremtid (*Birdsy*, u.å.).

Film eksemplar med kunstig intelligent

"I, robot" er en science-fiction film som kom ut i 2004, regissert av Alex Proyas. Filmen baserer seg i året 2035 med skuespilleren Will Smith som hovedpersonen, detektiv Del Spooner (Proyas, u.å.). "I, robot" tar for seg tre regler for roboter basert av forfatteren Isaac Asimov.

1. En robot skal ikke skade et menneske, eller tillate mennesker å bli skadet.
2. En robot må lyde til gitte ordre fra mennesker, hvis ikke det går i konflikt med første regel.
3. En robot må beskytte seg selv, så lenge det ikke går i konflikt med første og andre regel.

Kilde: <https://www.historyofinformation.com/detail.php?entryid=4108>

Filmen tar for seg intelligente menneskelige roboter, med diverse oppdrag gitt fra mennesker. Innbyggerne bruker robotene som sin egne butler, for eksempel husarbeids tjenester. Robotene bruker stemme gjenkjenning for å ta inn ordre fra mennesker, og kamera som gjenkjenner objekter. I tillegg er robotene bygget opp slik at de er mye sterkere og stødige enn et menneske. De fleste personene i filmen har stor tro på at robotene er fremtiden, unntatt hovedpersonen Spooner som har lite tillit i roboter. Grunnen til Spooner skeptiske adferd mot roboter, er på grunn av en hendelse hvor en robot kalkulerte overlevelse prosenten mellom Spooner og et annet menneskes liv. Roboten bestemte seg for å redde Spooners liv, men i Spooners mener at roboten skulle ha reddet det andre mennesket.

1.2 Roboter og kunstig intelligent systemer

I Tsjekkia 1920 dukket ordet "robot" for første gang i teaterstykket R.U.R (Rossum's Universal Robots) skrevet av Karel Čapek. Ordet robot på tsjekkisk betyr tvangsarbeid (*Origin of the Word «Robot»*, 2016). Det er vanskelig å si hva som kommer til å bli kalt roboter i fremtiden, fordi robot designet endres raskt med nye opprettede robot kompanier samtidig som andre robot kompanier blir stoppet (Dautenhahn, 2018, s. 2).

Definisjoner om robot

Sebastian Thurn trekker frem definisjonen av ordet "robot" fra robot instituttet i Amerika. "A robot is a reprogrammable, multifunctional manipulator designed to move materials, parts, tools, or specialized devices through various programmed motions for the performance of a variety of tasks" (Thurn, 2004, s. 11). Videre blir det nevnt at denne definisjonen mangler å nevne om autonome ferdigheter som en robot kan ha.

Ifølge den store norske leksikon blir ordet "robot" definert som "Robot, en datastyrt enhet som ved hjelp av sensorer kan motta data fra omgivelsene, bearbeide disse og reagere ved å iverksette handlinger i henhold til forhåndsprogrammerte regler" (Liseter, 2020)

Begge definisjonene nevner at en robot er programmert til å gjøre bestemte oppgaver, men som Thurn har nevnt så sier ikke den første definisjonen noe om autonome ferdigheter. Den store norske leksikon never at en robot er en datastyrt enhet, som tyder på at roboten jobber automatisk ved å styre data-ene for oss mennesker. Min definisjon av en robot er, en datastyrt eller automatisert enhet programmert til å gjøre menneskelige handlinger slik at hverdags livet blir lettere, ved hjelp av bevegelser og sensorer. Jeg fokuserer mer på at roboter gjør hverdags oppgaver litt enklere i livet, men det finnes også ubrukelige roboter. For eksempel en robot

som skrur av en bryter hver gang en person skrur på bryteren. Disse robotene er mer egnet for underholdning.

Forholdet mellom roboter og kunstig intelligent

Roboter og kunstig intelligens er ganske like, ved at begge parter gjør handlinger for oss mennesker og kan etterligne oss mennesker. Det som utelater en robot for å være kunstig intelligens er at roboten ikke er programmert til å lære over tid, men må heller omprogrammeres hvis det skulle være endringer. Kunstig intelligens derimot læres og trenes over tid, noe som gjør det vanskelig å si om en kunstig intelligens er et fullt ferdig produkt. Det kan tenkes som et menneske, hvor kunstig intelligens er hjernen og roboten er kroppen.

Robot omeletten

En automatisk robot kokk som har egenskapen til å tilbedre og servere en omelett. Roboten ble utviklet i Singapore med prototype navnet AUSCA (Dorn, 2019). Roboten oppfører seg om et menneske arm og beveger seg fleksibelt rundt stasjonen sin. Maskinen bidrar til å gi oss perfekte omeletter hver gang, slik at vi mennesker ikke uheldigvis server en brent omelett. Interaksjonen med roboten starter med at vi mennesker selv setter hvilken topping til omeletten vi vil ha. Deretter trykker på omelett knappen, som setter i gang roboten.

1.3 Universal design og kunstig intelligent systemer

"Universal design can be defined as the design of products and environments to be usable to the greatest extent possible by people of all ages and abilities. Universal design respects human diversity and promotes inclusion of all people in all activities of life" (Story et al., 1998, s. 11).

Universal design på norsk er kjent som UU (universal utforming), handler om å være tilgjengelig for alle. Uansett om en person har nedsatte ferdigheter, for eksempel hørsel, syn eller en brukket arm. Skal personen kunne benytte et design eller en tjeneste i likhet med de som ikke har nedsatte ferdigheter.

Kunstig intelligent inkluderer mennesker

Kunstig intelligens kan bidra til å få et design til å bli mer universell utformet. Et eksempel på det er videotjenesten, Youtube. En video som blir lastet opp på Youtube-nettsiden kan benytte kunstig intelligens for å generere undertekst på videoen, slik at de med hørsels vanskeligheter ikke går glipp av innholdet på videoen. For å generere undertekst benyttes det kunstig

intelligens med stemme gjenkjenning, for å få talende ord til skriftlig ord. Et annet eksempel hvor kunstig intelligens inkluderer flere parter er selvkjørt biler, som kan bidra til flere med nedsatte funksjonsevne til å kunne kjøre bil.

Kunstig intelligent inkluderer og ekskluderer mennesker

Kunstig intelligens kan også bidra til å inkludere samtidig ekskludere brukere. For eksempel på flyplassen hvor det blir utført passkontroll. Hvis det er en selvbetjent passkontroll som benytter kunstig intelligens med bruk av ansikts gjenkjenning for å gjøre prosessen mer raskere. I likedan kan kunstig intelligens bli brukt til å gjenkjenne ansiktstrekk for å finne en kriminelle personer. Problemet med et slikt eksempel scenario er at kunst intelligens kan melde til feil person for kriminelle handlinger, fordi noen ansikts trekk ser like. Det kan også hende at de reisende har tatt skjønnskirurgi som gjør at ansikts trekk ikke stemmer overens med passbildet.

Har maskiner en forståelse?

Min oppfatning om konseptet "forstå" og "forståelse" er fleksibel læring, hvor vi mennesker har en forståelse av det vi har lært til å kunne benyttes til flere situasjoner. Maskiner derimot forstår kun programmerings koder, instruksjoner som vi mennesker gir til maskinen. En maskin vil ikke kunne kombinere sine funksjonaliteter eller egenskaper av seg selv, med mindre vi mennesker programmerer maskinen slik. Kan maskiner forstå? Nei, maskiner kan kun forstå kommandoer gitt fra oss mennesker.

1.4 Retningslinjer for menneskelig -kunstig intelligens interaksjon

Retningslinje nummer 12 fra Microsofts menneskelige interaksjon med kunstig intelligens sier "Remember recent interactions". Et eksempel for retningslinje 12 er når en bruker titter i nettbutikker, dermed vil nettsiden vise en historikk av besøkte produkter. Ut ifra historikken kan nettbutikken benytte kunstig intelligens til å rekommandere tilpassende produkter for brukeren.

Shneidermans åtte gyldne regler er et sett med HCI(Human computer interaction) design retningslinjer (*Guidelines in HCI*, u.å.). Shneidermans og Microsofts retningslinjer har mye til felles. For eksempel: enkelt å angre handlinger, gi brukeren en form for tilbakemeldinger, og være konsist på informasjonen som vises til brukeren. Retningslinjene har også noen ulikheter, for eksempel hos Microsoft nevnes det om å lære om brukerens atferd. Shneiderman har ikke et slik retningslinjer.

Endringer til modul 1

Definisjonen min om roboter gikk ikke helt inn på automasjon, med tanke på roboter i fabrikker. Jeg har da endret definisjonen min litt.

Modul 2

2.1 Karakteristiske trekk innenfor KI-tilførte systemer

Karakteristiske trekk for KI-tilførte systemer er "systems that have features harnessing AI capabilities that are directly exposed to the end user" (Amershi et al., 2019).

KI-tilførte har fire karakteristiske trekk, læring, forbedring, black-box og drives av store datasett. Black-box er når brukeren tilfører en form av data inn til KI systemet. Deretter gir KI systemet en type tilbakemelding eller respons tilbake til brukeren. Hvordan systemet kommer frem til svaret kan virke mystisk for brukeren. KI systemer drives av store datasett ved å innsamle en del data basert av brukerens interaksjoner. Systemer kan reagere forskjellig eller oppføre seg annerledes over tid, som da er basert av samme tekst input fra en bruker eller neste bruker (Amershi et al., 2019). KI systemer kan også virke uforutsigbar, som vil skape utfordringer for brukergrensesnitt / interaksjons designer. Studier viser at HCI (Human Computer Interaction) designere strever med å se for seg designet og prototypen av KI systemer (Yang et al., 2020, s. 1).

Eksempel på et KI-tilført system

Nettsiden Youtube har KI-tilførte systemer. Brukeren kan få rekommanderte videoer basert av KI systemets beregninger. Systemet drives av å innsamle data basert av brukeren interaksjoner. Eksempel på dette er når brukeren har sett på en video, og deretter sjekker ut sine rekommanderte videoer. Youtube vil da rekommandere videoer basert av de videoene brukeren har klikket seg inn på, disse kan være basert på tema tagger eller fra samme videoskaper. I noen tilfeller kan brukeren bli rekommandert videoer som har ingen relasjoner fra tidligere sette videoer, dette gir brukeren en black-box effekt.

2.2 Menneskelige-KI interaksjons design

Forskningspapiret av Amershi et al. (2019) introduserer 18 retningslinjer for menneskelige-KI interaksjoner. Disse retningslinjene skal bidra til å hjelpe brukergrensesnitt designere med utfordringer innenfor et KIs brukergrensesnitt. Noen av retningslinjene gjelder ikke for de mindre grafiske grensesnittene, for eksempel stemme baserte, virtuell assistent eller aktivitets tracker (Amershi et al., 2019).

Forskningspapiret av Kocielnik et al. (2019) går inn på slutt brukerens forventninger for et KI-tilført system. KI-systemet som dette ble testet på var Scheduling Assistant, som skulle automatisk oppdage om en e-post melding inneholdt en møteforespørsel. Kocielnik et al.

(2019) trekker frem at bruker forventninger kan blir formet innen 3 prinsipper. 1) Informasjon fra eksterne kilder (spesifiser egenskapene til systemet), 2) argumenter og forståelse (Formuler forventninger basert av forståelse av systemet), og 3) første oppfatnings opplevelse (assosierte handlinger til systemet, gjennom interaksjon og opplevelse) (Kocielnik et al., 2019).

2 retningslinjer fra Amerashi et al. (2019) rettet mot eksempelet av et KI-tilført system

Retningslinje 2: Vis frem hvor godt systemet kan gjøre det den kan gjøre. Hjelp brukeren til å forstå hvordan KI systemet kan lage feil.

Youtube har i noen små tilfeller rekommandert brukere videoer som har lite eller ingen relasjoner fra brukerens video historikk. Oftest er dette i håp om å fange brukeren til nye interesser, og dermed få brukeren til å benytte nettsiden lenger. Det er ingen informasjon som viser til hvorfor KI systemet rekommanderer visse videoer, noe som kunne ha vært interessant å vite. Dette kan være fordi systemet tar det som en selvfølge at brukeren oftest forstår hvorfor visse videoer blir rekommandert.

Retningslinje 8: Støtte effektive oppsigelser. Gjør det enkelt for å avvise eller ignorere uønskede KI-systemtjenester.

I noen tilfeller kan brukeren få rekommandert videoer de ikke liker, eller spesifikke Youtube kanaler de ikke vil se lenger. Systemet har da en funksjonalitet til å gi brukeren muligheten for å ikke vise mer av disse typer videoer. Dette gjøres ved at brukeren klikker seg inn på de tre prikker-ikonet under videoen og velger "ikke interessert". KI systemet vil da ikke lenger vise de uinteressante videoene for brukeren. Hvis brukeren har valgt "ikke interessert" ved uhell, så kan de søke opp videoen. Dermed se på videoen og få relaterte videoer rekommandert igjen.

Sammendrag av Bender et al. (2021)

Artikkelen "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" av Bender et al. (2021) handler om KI maskin læring innenfor store språk modeller. Artikkelen tar for seg maskin læring og miljø risikoer, men også mulige løsninger. Forskning og trening av maskin læring for store språk modeller kan kreve en rekke ressurser som økonomiske kostnader og Co2 utslipp. I likedan blir teknologien oftest designet til en margin av lokalsamfunnet, oftest til rike land. Dermed bør forskere prioriteres av effektiv resurs håndtering for å forhåpentligvis redusere drastiske klima endringer. Ved bruk av internett for datainnsamling om mangfoldets perspektiv i verden, vil vi kunne se at datainnsamlingen kan skape utelukninger og stygge stereotyper. Dette er fordi internett representerer ikke hele mangfoldet, men dominerende meninger fra de som benytter nettet. Dermed blir det viktig å dokumentere datainnsamlingen under språk modellens trenings prosess.

2.3 Chatbots / dialog brukergrensesnitt

For å designe en chatbot bør fokuset være dialog orientert design. Sentrale utfordringer til designet er læring, krav på brukerinnsats, mangel på tilbakemelding og at chatboten ikke møter brukerens forventninger. Under læringsprosessen er det restriksjoner av språk bruk, fordi chatboten ikke kan ta for seg komplekse setninger med en gang. Det vil også kreves en del brukerinnsats over tid for å trene opp en chatbot. En chatbot har ingen minne eller kunnskap, men istedenfor etterligner en samtale (Luger & Sellen, 2016, s. 5287). Luger et al. (2016) påpeker også på dialog utfordringer som kan skape mindre suksess, for eksempel at chatboten gir uforståelig tilbakemeldinger, og vanskeligheter for å navigere rundt dialoger.

Retningslinjen 1 og 2 fra Amerashi et al. (2019) kan løse chatbot utfordringer

Retningslinje 1: Gjør klar hva systemet kan gjøre. Hjelp brukeren til å forstå hva KI systemet er i stand til å gjøre.

Hvis en chatbot er i stand til å vise brukeren hva de kan gjøre, så vil chatboten kunne møte brukerens forventninger. Dette kan være for eksempel at startdialogen til chatboten gir klar beskjed over hva hensikten chatboten kan hjelpe med til brukeren. Slik at dialogen ikke er åpen for hvilke som helst samtaler, enn det chatboten er i stand til.

Retningslinje 2: Vis frem hvor godt systemet kan gjøre det den kan gjøre. Hjelp brukeren til å forstå hvordan KI systemet kan lage feil.

Ved å gi chatboten egenskapen til å gi feilmeldinger basert av f.eks. ord som ikke er registrert i systemet, vil gi brukeren en bedre brukeropplevelse. Enn om chatboten repeterer uforståelige tilbakemeldinger eller gjentakende tilbakemeldinger. Dermed kan brukeren prøve seg på andre ord i håp å få en bedre dialog med chatboten. Hvis chatboten kontinuerlig sender feilmeldinger til brukeren, da bør kanskje chatboten gå tilbake til trenings prosessen.

Kilder

Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for Human-AI Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300233>

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>

Birdsy. (u.å.). Hentet 10. september 2021, fra <https://store.birdsy.com/>

Bratteteig, T., & Verne, G. (2018). Does AI make PD obsolete? Exploring challenges from artificial intelligence to participatory design. *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial - Volume 2*, 1–5. <https://doi.org/10.1145/3210604.3210646>

Dautenhahn, K. (2018). Some Brief Thoughts on the Past and Future of Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction*, 7(1), 4:1-4:3. <https://doi.org/10.1145/3209769>

Dorn, L. (2019, mai 22). An Autonomous Self-Service Robotic Chef That Prepares Customized Omelettes on Demand. *Laughing Squid*. <https://laughingsquid.com/omelette-making-robot/>

Grudin, J. (2009). AI and HCI: Two Fields Divided by a Common Focus. *AI Magazine*, 30(4), 48. <https://doi.org/10.1609/aimag.v30i4.2271>

Guidelines in HCI. (u.å.). Hentet 10. september 2021, fra https://www.tutorialspoint.com/human_computer_interface/guidelines_in_hci.htm

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3290605.3300641>

Liseter, I. M. (2020). Robot. I *Store norske leksikon*. <http://snl.no/robot>

Luger, E., & Sellen, A. (2016). «Like Having a Really Bad PA»: The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5286–5297. <https://doi.org/10.1145/2858036.2858288>

McCarthy, J. (2004). *What is artificial intelligence?* 14.

Origin of the Word «Robot». (2016, juli 15). Robotics Business Review.
https://www.roboticsbusinessreview.com/rbr/origin_of_the_word_robot/

Proyas, A. (u.å.). *I, Robot*. Hentet 10. september 2021, fra
https://www.rottentomatoes.com/m/i_robot

Saleh, Z. (2019). *Artificial Intelligence Definition, Ethics and Standards*.

Story, M. F., Mueller, J. L., & Mace, R. L. (1998). *The Universal Design File: Designing for People of All Ages and Abilities. Revised Edition*. Center for Universal Design, NC State University, Box 8613, Raleigh, NC 27695-8613 (\$24). <https://eric.ed.gov/?id=ED460554>

Thrun, S. (2004). Toward a Framework for Human-Robot Interaction. *Human-Computer Interaction*, 19(1–2), 9–24.

Tidemann, A. (2020). Kunstig intelligens. I *Store norske leksikon*.
http://snl.no/kunstig_intelligens

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3313831.3376301>