Individuell oppgave

kristavi

Modul 1

Historien til AI

I 1949 skrev Alan Turing i London Times om hvordan han "[...] ikke kunne se hvorfor datamaskiner ikke skal være del i felt som vanligvis er dekket av den menneskelige intelligens, og etter hvert kunne konkurrere på likt nivå" (Grudin, 2009, s. 49). Noen år senere, i 1956, var det matematiker og logiker John McCarthy som var først ute med å bruke selve begrepet "artificial intelligence" (Grudin, 2009, s. 49). Gjennom tiårene i ettertid har interessen for AI gått i bølger, men utviklingen har vært enorm.

Definisjoner av AI

Her er tre noe ulikt vinklede definisjoner av kunstig intelligens:

"Kunstig intelligente systemer utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål" (Regjeringen, 2022). Dette skriver regjeringen på sine nettsider, i en tekst om nasjonal strategi for kunstig intelligens. Definisjonen er basert på EUs ekspertgruppe sin formulering.

I deres artikkel om sammenhengen mellom AI og deltakende design, skriver Bratteteig og Verne: "AI er et underfelt av informatikk som tar sikte på å spesifisere og lage datasystemer som etterligner menneskelig intelligens eller uttrykker rasjonell atferd, i den forstand at oppgaven ville kreve intelligens hvis den ble utført av et menneske" (Bratteteig og Verne, 2018). Artikkelen er fra 2018.

Til slutt skriver Store Norske Leksikon (forfatter Axel Tidemann) i 2020 at "kunstig intelligens er informasjonsteknologi som justerer sin egen aktivitet og derfor tilsynelatende framstår som intelligent" (Tidemann, 2020).

Basert på disse definisjonene blir mitt forsøk på en definisjon av AI: "Et datasystem som, med basis i tolkning av data, utfører handlinger som ville krevd intelligens dersom det ble utført av et menneske." Det er dette med at AI kan tolke og analysere data på en måte som

får den til å virke intelligent som for meg er det mest fascinerende, samt definere med kunstig intelligens.

Likheter og forskjeller mellom menneskelig autonomi og maskinautonomi

I likhet med menneskelig autonomi, handler maskinenes autonomi om å gjøre handlinger uten påvirkelse utenfra, basert på omgivelser og inntrykk. Grunnlaget for valgene som blir gjort er på mange måter liknende eller like, men også ulike, mellom menneske og maskin. En maskin kan ha et kamera der menneske har øyne, og vil dermed kunne visuelt se den samme fysiske situasjonen, men mennesket kan basere tolkningen sin av situasjonen på faktorer maskinen ikke kan, som for eksempel følelser, empati eller nysgjerrighet.

Likheter mellom konseptet automatisk og konseptet autonomisk

Både autonomi og automatikk handler om å gjennomføre handlinger uten menneskelig assistanse, men der noe automatisk kan gjennomføre en oppgave frem til den opplever et problem som får den til å stoppe opp, kan noe autonomt reagere og endre oppførsel, og dermed fortsette på egenhånd.

Gjennomgang av en artikkel

Jeg har valgt å skrive om artikkelen “Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design”, som er skrevet av Tone Bratteteig og Guri Verne (Bratteteig og Verner, 2018). Den tar opp noen svært interessante problemstillinger rundt design av AI-teknologi ved bruk av deltalkende design, samt om man i det hele tatt trenger PD når AI kan personalisere løsninger av seg selv. Særlig dette med at et AI-basert design forslag alltid vil forandre og utvikle seg av seg selv, og at man aldri vet helt hvordan løsningen vil fungere i fremtiden, er interessant. Jeg synes artikkelen er konsis og godt skrevet, og sier meg ikke uenig i noe av det de presenterer.

Et moderne selskap som jobber med AI

Det norske fintech selskapet Strise bruker AI-begrepet på forsiden sin slik: “Strise is an AI-powered platform that solves complex challenges in your KYC work” (Strise, 2022). De utvikler og leverer en tjeneste der skandinaviske selskaper som trenger å gjøre bakgrunnssjekker på andre selskaper kan finne “all” relevant informasjon på ett sted. De beskriver ikke i detalj hvordan AI er involvert i tjenesten, men den blir presentert som at den er sentral i hvordan informasjon fra ulike kilder hentes, tolkes og så blir presentert i tjenesten. AI blir dermed å se på som en *service* fra Strise.

Hvordan menneskelig interaksjon med AI blir presentert i et fiksjonelt verk

I filmen Ex Machina fra 2015 møter vi Caleb, som blir valgt ut til å være med i et vitenskapelig eksperiment hvor han skal være med på å evaluere *Ava.* Ava er en svært sofistikert AI-basert robot laget for å se ut som, og oppføre seg, som et menneske. Caleb interagerer med henne slik man ville interagert med et menneske: ved å snakke til henne (inkludert bruk av tonefall, pauser osv.), samt ved å utøve kroppsspråk og andre fysiske bevegelser. Ava responderer på samme måte, men det er tydelig i starten at hun ikke har det menneskelige grunnlaget hennes fysiske "alder" tilsier, og hun spør mange spørsmål.

Historien til ordet "robot"

Ordet *robot* ble brukt for første gang i et Tsjekkisk teaterstykke laget av Karel Čapek i 1920. Ordet kommer fra *robota,* som betyr *tvungen arbeid.* I stykket har mennesket laget menneskelignende roboter, som først er fornøyd med å arbeide, men så gjør opprør og utrydder mennesket (Science Friday, 2011). Fra 20-tallet til den dag idag har begrepet og konseptet *robot* vært i stadig endring, og som Dautenhahn skriver har det vært signifikante endringer i robot-miljøet de siste 20 årene (Dautenhahn, 2018).

Definisjoner av robot

Thrun presenterer flere ulike definisjoner på robot i sin artikkel om HRI (Thrun, 2004). Han trekker blant annet frem the Robot Institute of America, som i 1979 definerer ordet slik: "En omprogrammerbar, multifunksjonell manipulator designet for å flytte materialer, deler, verktøy eller spesialiserte enheter gjennom ulike programmerte bevegelser for utførelse av en rekke oppgaver» (Thrun, 2004, s. 11). Her er det et fokus på at robotene hovedsakelig brukes til å flytte gjenstander, noe de ofte blir brukt til innen industri, og som trolig var det vanligste bruksområdet på denne tiden. Definisjonen åpner heller ikke for at en robot kan være autonom, men at den kun kan gjøre det den programmeres til.

Standarden ISO 8373:2012 beskriver en robot som "en aktivert mekanisme som er programmerbar i to eller flere akser med en grad av autonomi, som beveger seg i omgivelsene for å utføre tiltenkte oppgaver" (ISO, 2012). I motsetning til den første definisjonen sier de altså her at roboter *har* en grad av autonomi, som da åpner for varierende grader av AI-implementasjon. Resten av definisjonen likner mer på den over 30 år eldre definisjonen, ved at en robot er programmerbar, og beveger seg for å utføre oppgaver.

Min definisjon er som følger: "En robot er et programmerbart objekt som beveger seg i, og interagerer med sine omgivelser for å gjennomføre tiltenkte oppgaver, og som kan ha en

grad av autonomi.” Jeg er enig i den meste med den nyeste definisjonen, utenom at en robot nødvendigvis må ha en grad av autonomi, og den eldre definisjonen synes jeg blir for spesifikk for dagens mer varierende bruksområder.

Relasjonen mellom AI og robot

Med basis i definisjonene av både robot og AI, mine egne og de andre som er nevnt, er den første og mest åpenbare forskjellen på en AI og en robot det fysiske. En robot er per definisjonene et fysisk objekt som eksisterer i fysiske omgivelser. En AI eksisterer ved bruk av fysiske gjenstander som datamaskiner, men AI-en i seg selv er ikke noe man kan se (som en gjenstand, man kan se programkoden den består av), eller berøre.

Når det kommer til likheter ser vi at både roboter og AI eksisterer for samme grunn: å hjelpe mennesket med oppgaver. Roboter hjelper oss med fysiske oppgaver, og AI med avanserte (eller enkle) oppgaver som ikke krever fysisk bevegelse. En robot kan også *ha* AI, og dermed bli mer autonom i dens løsning av oppgavene, men separert arbeider de ulikt.

Hvordan en moderne kommersiell robot beveger seg og interagerer

Jeg har valgt å se på robotgressklippere, mer nøyaktig Husqvarnas modell Automower 405X. Disse har som navnet tilsier én oppgave: å klippe gress. De vanligste robotgressklipperne idag går på batteri som lades opp via en ladestasjon den selv parkerer ved. Klipperens område defineres ved bruk av en fysisk grense-vaier som legges ut ved installasjon. For å finne tilbake til ladestasjonen legges en annen vaier som klipperen kan følge (Husqvarna, 2022).



Gressklipperen er laget for å kreve så lite interaksjon fra et menneske som mulig, men dersom den setter seg fast eller må vaskes gjøres dette naturligvis med fysisk interaksjon fra en person. Videre har man et grafisk grensesnitt i form av en app, i Husqvarna sitt tilfelle *Automower Connect*, som gir eieren tilgang til å starte, stoppe eller parkere klipperen, eller spore via GPS dersom den skulle bli stjålet eller kjøre seg vill (Husqvarna, 2022).

Definisjon av Universelt Design

Regjeringens nettside beskriver universell utforming som “utforming av produkter, omgivelser, program og tjenester på en slik måte at de kan brukes av alle mennesker, i så stor utstrekning som mulig, uten behov for tilpasning og en spesiell utforming” (Regjeringen,

2021). Det handler altså om å ikke ekskludere personer med diverse nedsettelser og utfordringer fra å kunne ta i bruk, i lik grad som alle andre, det man designer.

AIs potensiale innen menneskelig oppfatning, bevegelse og kognisjon

AI-basert teknologi har potensiale til å forbedre oppfatning, bevegelse og kognisjon for personer som av ulike årsaker har utfordringer med dette. Dette kan eksempelvis være i form av armproteser som ved bruk av AI gjør en stadig bedre jobb med å lese nervesignalene fra hjernen (The Medical Futurist, 2020). Et annet eksempel er Seeing AI fra Microsoft, en app som gjennom et talegrensesnitt beskriver det det den ser gjennom et smarttelefonkamera (Microsoft, 2022).

AIs potensiale for å inkludere og ekskludere

AI er basert på eksisterende data og statistisk analyse av denne, og vil dermed øve seg mest på å forstå og gjenkjenne det som forekommer oftest i dataen. Dette kan ha en ekskluderende effekt for personer som i ulike sammenhenger havner utenfor normalkurven. Et eksempel på dette kan være kjønn. En AI-basert ansiktsgjenkjenning kan komme frem til at en person er mann eller kvinne, selv om denne personen identifiserer seg som noe annet.

Derimot kan AI også ha en inkluderende effekt i diverse sammenhenger. Dersom den blir trent med gode datasett og gitt riktige restriksjoner, kan en AI eksempelvis ignorere faktorer som mennesker kan ha en tendens å ha fordommer mot. Dette kan være inkluderende i for eksempel jobbsøkersammenheng, hvor den kan ignorere opprinnelsen av navn eller hudfarge.

Hva er “forståelse,” og forstår maskiner?

Slik jeg ser det handler forståelse om å skjønne et konsept på et dypere nivå enn å lære seg at det eksisterer, eller hvordan det er sammensatt. Det handler mer om å skjønne *hvorfor* noe er som det er, i den kontekst det eksisterer i. En datamaskin kan se at mennesker skriver “love you” i slutten av meldinger til hverandre, men har ikke mulighet til å faktisk forstå hvorfor. Den kan lære seg at det er noe mennesker sender til de de kommuniserer mest med, men konsepter som kjærlighet vil den ikke kunne faktisk forstå. Maskiner er basert på matematikk og logikk, og om det vil være mulig å digitalt gjenskape noe så komplisert og menneskelig som forståelse blir spennende å se.

Menneske-AI interaksjonsretningslinjer fra Microsoft

Jeg har valgt nr. 12: “Husk nylige interaksjoner.” Denne retningslinjen beskriver hvordan AI-et burde ha nylige interaksjoner og dets resultat lett tilgjengelig for brukeren, slik at det enkelt kan bli gjort referanser til dette. Et eksempel på dette kan være en høyttalerbasert smartassistent, hvor bruker kan si “Repeter det du sa” og assistenten vil presentere resultatet av en handling på nytt.

Likheter mellom HCI retningslinjer og Menneske-AI retningslinjer

Jeg fant Schneidermanns “åtte gylne regler som vil hjelpe deg med å designe bedre grensesnitt” (Euphemia Wong, 2020). Noen av disse likner veldig på Menneske-AI retningslinjene. Begge nevner hvordan man bør håndtere error og brukerfeil, at bruker skal ha kontroll over interaksjonen, og at løsningene skal passe med det brukeren er vant til og kjenner igjen. Utover dette har Menneske-AI et naturlig større fokus på AI-relevante retningslinjer, som at AI skal presentere hvordan det har endret oppførsel og hvorfor.

Modul 2

Karakteristikker ved AI-tilføyde systemer

Amershi et al. (2019) beskriver flere karakteristikker ved disse typene systemer. Det blir beskrevet at det er bundet en usikkerhet til dem, altså at de ofte gjør feil som i verste fall kan oppleves som forvirrende og i noen tilfeller støtende, og at de er inkonsistente, som vil si at de er svært sensitive til kontekst og små endringer i input. Videre skriver de at systemene ofte har “behind the scenes” personalisering; at de på egen hånd personaliserer og i noen tilfeller da gjemmer informasjon som egentlig var interessant for brukeren, uten at de får det med seg (Amershi et al., 2019, s. 1).

Kocielnik et al. (2019), i deres artikkel om å sette forventninger til AI-tilføyde systemer, skriver at disse nesten alltid operer med mindre enn perfekt nøyaktighet, noe som gjør at brukere har lett for å bli misfornøyd da de ofte forventer at det skal være konsistent og perfekt i bruk. Det blir også beskrevet at det kan være utfordrende med transparens rundt den kunstige intelligensen og hva den har basert valgene sine på, samt at de blir påvirket av brukernes handlinger, noe som må kommuniseres til brukeren (Kocielnik et al., 2019, s. 2-3).

Yang et al. (2020) skriver i sin artikkel om Menneske-AI interaksjon at systemene kan gjøre uforutsigbare feil som kan påvirke brukeropplevelsen negativt, og i noen tilfeller ha etiske utfordringer og i noen tilfeller konsekvenser på samfunnsnivå (Yang et al., 2020, s. 1).

Et AI-tilføyd system jeg bruker

Et AI-tilføyd system som jeg bruker på daglig basis er Spotify. Musikkstrømmetjenesten bruker AI til å anbefale sanger, artister, album, podcast og spillelister på flere ulike måter, og man kan finne mange av de nevnte karakteristikkene i tjenesten. Det er blant annet stor grad av "behind-the-scenes" personalisering. Tjenesten bruker flere ulike datapunkter fra hver bruker til å produsere disse anbefalingene, og eksempelvis opplever jeg personlig ofte at albumslipp jeg egentlig er svært interessert i å få med meg ikke blir anbefalt til meg, og dermed går under radaren. Videre er det også svært lite transparens i hva datagrunnlaget for anbefalingene er, sett bort ifra en oppsummering av hvilke sanger og artister jeg har hørt mest på som kommer på slutten av hvert år. Den gir også brukerne noe, men relativt lite, kontroll over å påvirke anbefalingene i form av at man kan gi en anbefalt sang tommel ned om man ikke likte den. Spillelister, album og podcasts kan man ikke gi tilbakemelding på. Dette gjør at man ofte får lite relevante anbefalinger som kan bli stående på forsiden i lang tid og dermed hindrer bedre anbefalinger fra å vises.

To artikler om menneske-AI interaksjonsdesign

Amershi et al. (2019) presenterer 18 retningslinjer for Menneske-AI interaksjon man bør ta hensyn til når man designer AI-tilføyde systemer. Retningslinjene er organisert etter når i bruken de vil være relevante, og blir presentert med eksempelapplikasjoner fra ulike systemer. Artikkelen går i dybden på hvordan de samlet, evaluerte og itererte på retningslinjene i fire faser: konsolidering av nåværende retningslinjer, intern evaluering, brukerstudie, og ekspertevaluering.

Kocielnik et al. (2019) diskuterer hvordan man kan sette realistiske forventninger for sluttbrukere når det gjelder bruken av AI-tilføyde systemer. For å gjøre dette bruker de et eksempel i form av en automatisk møteplanlegger som bruker AI til å tolke innholdet i e-poster. De presenterer tre teknikker som er med på å sette realistiske forventninger til resultatet: en nøyaktighetsindikator som beskriver hvor bra teknologien vanligvis fungerer, bruk av eksempelbaserte beskrivelser og til slutt ytelseskontroll; at brukeren selv velger hvordan ytelse de ønsker fra systemet. Videre demonstrerer de hvor effektive disse teknikkene er på å oppnå nettopp brukertilfredshet og aksept for AI-tilføyde systemer. De viser også hvordan et slikt system som er optimalisert for å unngå falske negativer oppleves som mer nøyaktig og oppnår høyere aksept enn systemer som er optimalisert for å unngå falske positiver.

Spotify og retningslinjer fra Amershi et al. (2019)

Jeg har valgt å skrive om følgende retningslinjer:

G11: Make clear why the system did what it did
G9: Support efficient correction

“Make clear why the system did what it did” (Amershi et al., 2019) handler om at systemer forklarer hvorfor resultatet av en AI-handling ble som det ble. Det er begrensede muligheter for å se dette i Spotify. Anbefalinger kommer blant annet på forsiden under kategorien “Mer av det du liker”, som i noen grad forklarer grunnlaget for anbefalingene. Andre anbefalinger blir formulert som “Andre som hørte på X hører også på…”, som gir et bedre grunnlag for å forstå hvorfor en artist, sang, album eller podcast blir anbefalt til deg. Et annet eksempel er i en spilleliste hvor man kan trykke på “Enhance”, noe som legger til sanger som den tror vil passe inn i spillelisten. Utover å hinte til at den baserer seg på den spillelisten man er i, får man ingen andre beskrivelser av hva grunnlag vurderingen av de nye sangene er basert på. For å forbedre dette kunne det blitt presentert en forklaring, for eksempel tilgjengelig via en knapp, i de relevante sidene. Når man trykker “Enhance” burde det være mulig å se hvilke sjangre, tempo osv. den tror vil passe inn i listen, samt liknende funksjonalitet ved andre typer anbefalinger.

“Support efficient correction” (Amershi et al., 2019) handler om at brukere enkelt kan korrigere feil den kunstige intelligensen har gjort. I Spotifys tilfelle kan dette være når en anbefaling ikke føles relevant eller virker som den er gitt på feilaktig grunnlag. Tjenesten har, som nevnt tidligere, noe funksjonalitet for å fjerne dårlige anbefalinger, men dette er altså kun ved noen få typer anbefalinger. Som bruker hadde det vært nyttig å kunne redigere Spotifys forståelse av hva musikk og podcasts man liker og misliker. Eksempelvis kunne dette vært en meny hvor man kan gi tydelig tilbakemelding på om de beslutnigene den har tatt om brukerens smak er rett eller galt, og dermed være med på å gjøre anbefalingene mer relevante. I mindre skala burde det vært tommel-opp og tommel-ned vurderinger tilgjengelig flere steder enn det er i dagens løsning.

Problematiske aspekter med store språkmodeller

Bender et al. (2021) skriver i sin artikkel om problematiske aspekter tilknyttet det faktum at dagens store *Natural Language Models* stadig vokser. De kommer med flere risikofaktorer som oppstår, og beskriver hvordan man kan se på veien videre med disse i bakhodet.

Et av problemene som blir nevnt er den økende ressursbruken de store datasette krever. Dette har direkte finansielle konsekvenser, og de økende økonomiske kostnadene gjør at det blir en større barriere for tilgang for ny forskning i området. De beskriver også hvordan større ressursbruk har en negativ påvirkning på miljøet, som påvirker marginaliserte

samfunn som i mange tilfeller ikke får noen betydelige goder ut av resultatet av forskningen. Videre diskuterer de hvordan den utrolige størrelsen på datasettene gjør at man mister oversikt over dataen som blir brukt. Dette kan blant annet resultere i at negative ideologier, stereotyping og nedverdigende assosiasjoner får plass i datasettene, som kan resultere i at den syntetiske teksten som blir produsert inneholder dette. Generelt ber de forskere være forsiktige med teknologi som er til for å etterligne mennesker, da dette medfører enorm risiko. Mennesker finner mening der det ikke finnes, og det tekstuelle innholdet som blir produsert som et resultat av disse modellene kan av mange bli tolket som om det er et ekte menneske som har ment og produsert det, og at det ikke bare er et resultat av kunstig intelligens.

Utfordringer ved design av chatboter og muntlige brukergrensesnitt

Følstad og Brandtzæg (2017) skriver om den nye verden av HCI, med fokus på interaksjon med systemer gjennom muntlig språk. De skriver at den største forskjellen er det at det er nettopp en *samtale* som er designobjektet. Designprosessen vil kreve en ny retning i at det ikke lenger vil handle om å i hovedsak designe grafiske grensesnitt, men mer *service design,* som vil si at fokuset må skifte mer mot brukerens behov. Generelt vil grafiske elementer bli å ta mindre og mindre plass i interaksjonen, og det blir brukerens egendefinerte input som styrer den. En annen nøkkelutfordring som blir trukket frem er hvordan de fleste slike grensesnitt i dag er basert på en “one-size-fits-all” tilnærming, der alle brukere vil få samme svar fra samme spørring, uansett behov, preferanser og digital kompetanse. Å kunne tilby skreddersydde formuleringer av svar basert på slike forskjeller vil kunne forbedre brukeropplevelsen drastisk for eksempelvis barn og eldre. Et annet viktig punkt som krever forbedring fra dagens løsninger, og som er utfordrende å løse, er måten et system kan forstå den *muntlige prosessen* bedre, slik at den, blant annet, kan gi akseptabel respons når en samtaletråd bryter ned, eller går i uønsket retning.

Retningslinje G1 og G2, og dagens muntlige grensesnitt

G1 “Make clear what the system can do” (Amershi et al., 2019) handler om å gjøre det tydelig for brukeren hva systemet kan gjøre. Det er stor variasjon i hvor gode dagens løsninger er på å følge denne retningslinjen. Mange av dagens chatbotløsninger åpner samtalen med et avsnitt om hva chatboten er til for å gjøre, og henviser ofte til noen eksempelspørringer. Andre gir svært lite utdypende informasjon om dette i det hele tatt. En godt formulert slik introduksjon, samlet med en mulighet for å åpne en side med mer utfyllende informasjon om chatboten er en god start. En tjeneste som løser dette på en effektiv måte er Google Home. Stadig vekk legger systemet til “Om du ønsker, kan du spør

meg om hva jeg kan gjøre” etter svaret på en spørring. Dermed er informasjonen alltid tilgjengelig, og man bli påminnet om hvordan man får tak i den med jevnlige mellomrom.

G2 “Make clear how well the system can do what it can do” (Amershi et al., 2019) handler i likhet med G1 om å sette realistiske forventninger til systemet hos brukeren, slik at de ikke blir frustrert og eventuelt bryter interaksjonen. Chatboten eller det muntlige grensesnittet må informere om hva man kan forvente fra det - både i form av talegjenkjenning, og i form av spørsmålene man kan stille, samt hva slags kvalitet på svarene man kan forvente å få i retur. Dersom det ikke blir oppgitt noe informasjon om dette, kan brukere forvente at det vil forstå alt man spør om, uansett hvor komplisert og dårlig formulert det skulle være.

Kilder

Modul 1

Bratteteig, T., Verne, G. (2018) Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design, UiO. Tilgjengelig fra: https://dl-acm-org.ezproxy.uio.no/doi/pdf/10.1145/3210604.3210646 (Hentet 13. september 2022). DOI: 10.1145/3210604.3210646

Dautenhahn, K. (2018) Some Brief Thoughts on the Past and Future of Human-Robot Interaction, University of Hertfordshire Tilgjengelig fra: https://dl-acm-org.ezproxy.uio.no/doi/pdf/10.1145/3209769 (Hentet 13. september 2022). DOI: 10.1145/3209769

Euphemia Wong (2020) *Shneiderman’s Eight Golden Rules Will Help You Design Better Interfaces* Tilgjengelig fra: https://www.interaction-design.org/literature/article/shneiderman-s-eight-golden-rules-will-help-you-design-better-interfaces (Hentet 13. september 2022).

Grudin, J. (2009) AI and HCI: Two fields divided by a common focus, i The AI magazine. Tilgjengelig fra: https://go-gale-com.ezproxy.uio.no/ps/i.do?p=AONE&u=oslo&id=GALE%7CA216267417&v=2.1&it=r (Hentet 13. september 2022). DOI: 10.1609/aimag.v30i4.2271

Husqvarna (2022) *Automower 405x* Tilgjengelig fra: https://www.husqvarna.com/no/robotgressklippere/automower-405x/ (Hentet 13. september 2022).

ISO (2012) *ISO 8373:2012(en) Robots and robotic devices - Vocabulary* Tilgjengelig fra: https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-2:v1:en (Hentet 13. september 2022)

Microsoft (2022) *Seeing AI* Tilgjengelig fra: https://www.microsoft.com/en-us/ai/seeing-ai (Hentet 13. september 2022).
Regjeringen (2022) *Nasjonal strategi for kunstig intelligens* Tilgjengelig fra: https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/?ch=3 (Hentet 13. september 2022).

Regjeringen (2022) *Veileder i universell utforming* Tilgjengelig fra: https://www.regjeringen.no/no/dokumenter/veileder-i-universell-utforming/id2850026/?ch=3 (Hentet 13. september 2022).

Science Friday (2011) *The Origin Of The Word 'Robot'* Tilgjengelig fra: https://www.sciencefriday.com/segments/the-origin-of-the-word-robot (Hentet 13. september 2022).

Strise (2022) *The future of KYC* https://www.strise.ai/ (Hentet 13. september 2022).

Thrun, S. (2004) Toward a Framework for Human-Robot Interaction, Human–Computer Interaction, 19:1-2, 9-24, DOI: 10.1080/07370024.2004.9667338 Tilgjengelig fra: https://www-tandfonline-com.ezproxy.uio.no/doi/pdf/10.1080/07370024.2004.9667338?needAccess=true (Hentet: 13. september 2022).

Tidemann, A. (2020) *Kunstig intelligens* https://snl.no/kunstig_intelligens (Hentet 13. september 2022).

The Medical Futurist (2020) *The Future Of Prosthetics Depends On A.I.* Tilgjengelig fra: https://www.regjeringen.no/no/dokumenter/veileder-i-universell-utforming/id2850026/?ch=3 (Hentet 13. september 2022).

Modul 2

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Teevan, J. (2019). Guidelines for human-AI interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 3). ACM. Tilgjengelig fra: https://dl-acm-org.ezproxy.uio.no/doi/abs/10.1145/3290605.3300233

Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623). ACM. Tilgjengelig fra: https://dl-acm-org.ezproxy.uio.no/doi/abs/10.1145/3442188.3445922

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. interactions, 24(4), 38-42. Tilgjengelig fra: https://dl.acm.org/citation.cfm?id=3085558

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. Tilgjengelig fra: https://dl-acm-org.ezproxy.uio.no/doi/abs/10.1145/3290605.3300641

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In Proceedings of the 2020 CHI conference on human factors in computing systems (Paper no. 164). Tilgjengelig fra: https://dl.acm.org/doi/abs/10.1145/3313831.3376301

Vedlegg

Feedback og endringer på første modul

Jeg fikk tilbakemelding på at jeg kunne vært noe mer utfyllende i svarene mine, og har dermed lagt til tekst på flere av svarene i modul 1.

Feedback og endringer på andre modul

Jeg fikk tilbakemelding på at jeg manglet en riktig kildehenvisning, samt noe annen formatering rundt retningslinjene som blir nevnt. Dette er nå fikset.