

IN5480 - Individual assignment

Modul 1 & 2, Høsten 2021

Innhold

Modul 1	2
1.1 Historien til AI	2
Tre definisjoner av AI	2
Sammendrag av artikkel	3
En bedrifts presentasjon av AI	3
Fictional Film	3
1.2 Roboter og AI systemer	4
Historien til begrepet «robot»	4
Definisjon av begrepet «robot»	4
Relasjon mellom AI og robot	4
AV1-roboten	4
1.3 Universell utforming og AI systemer	4
Universell utforming	4
Potensiale til AI	5
AI som inkluderende og ekskluderende	5
WCAG 2.1	5
1.4 Retningslinjer for Human-AI interaksjon	5
Retningslinjer fra Microsoft	5
Modul 2	7
2.1 Karakteristikk ved et AI-system	7
Youtube og AI	7
2.2 Human-AI interaksjonsdesign	8
Bender et al. – Large Language Model	9
2.3 Chatbots / samtalebaserte brukergrensesnitt	9
Vedlegg 1 – Tilbakemelding	10
Vedlegg 2 – Tilbakemelding	10
Litteraturliste	11

Modul 1

1.1 Historien til AI

Historien til AI utspiller seg fra etter andre verdenskrig og frem til i dag. I artikkelen «AI and HCI: Two Fields Divided by a Common Focus» skriver forfatteren at begrepet AI første gang ble brukt av matematikeren John McCarthy (Grudin, 2009). Begrepet ble brukt i sammenheng med en workshop i 1956.

I den samme artikkelen ser Grudin på relasjonen mellom AI og HCI og hvordan de springer ut fra to ulike fagområder. AI kommer fra matematikken, og HCI kommer fra psykologien (Grudin, 2009). Etter andre verdenskrig fortsatte datamaskinene å utvikle seg, og Alan Turing beskrev dette med følgende sitat “I do not see why [the computer] should not enter any one of the fields normally covered by the human intellect, and eventually compete on equal terms.” (Grudin, 2009). Dette kan være med på å vise hvordan relasjonen mellom den matematiske AI og psykologiske menneskepåvirkningen ikke kunne bli likestilt en gang i fremtiden. Den samme artikkelen viser også hvordan fagfeltene har beveget seg fra hverandre gjennom årenes løp.

Tre definisjoner av AI

Første definisjon – “AI is a subfield of computer science aimed at specifying and making computer systems that mimic human intelligence or express rational behaviour, in the sense that the task would require intelligence if executed by a human” (Bratteteig & Verne, 2018.).

Andre definisjon – “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, 1998).

Tredje definisjon – Med sitatet “AI set out to devise an artificial rival to human intelligence.” Skrev Gruding en kort beskrivelse av hva AI kan være, en rival til den menneskelige intelligensen (Gruding, 2009).

Disse tre definisjonene viser at AI skal etterligne menneskelige egenskaper, så ut i fra det vil jeg definere AI som en elektronisk maskin som skal etterligne menneskelige egenskaper og vår intelligens på best mulig måte.

Sammendrag av artikkel

Den artikkelen jeg ønsker å trekke fram er skrevet av Jonathan Grudin i 2009. Oppbyggingen av artikkelen er gjort slik at forfatteren tar oss igjennom verdenshistorien fra tiden etter andre verdenskrig til nåtiden. Overskriftene forteller oss tydelig hvilket tiår vi befinner oss i, og underoverskriften gir oss et inntrykk av hva vi nå skal begi oss ut på.

Å bli presentert for flere kjente navn og hendelser, og deretter å høre hva de bidro med til utviklingen, gjør at hele artikkelen føles som en historie. Dette skaper en god flyt i artikkelen og jeg sitter igjen med inntrykket at det var lettere å følge med og trekke tråder i historien.

En bedrifts presentasjon av AI

Jeg googlet litt rundt og fant ut at det kunne være litt vanskelig å finne et selskap som uttalte direkte at de bruker AI. For eksempel DNB har en chatbot, og etter litt leting finner jeg at de skriver at «AI-trenerne jobber kontinuerlig med å forbedre roboten.». Ut i fra den teksten jeg fant som var skrevet i 2018, (<https://www.dnb.no/dnbnyheter/no/din-okonomi/chatboten-kan-snakke-med-3000-samtidig>), er det kun den setningen som omtaler AI. Det får meg til å tenke over hvorfor de ikke skriver mer om hvordan AI brukes for å lage denne roboten. Jeg får følelsen at de presenterer AI litt som en selvfølge, men også litt mystisk. Selve roboten er der for å hjelpe deg, men hva med AI som ligger bak? Er det kanskje mer å relatere seg til når vi hører ordet robot?

Fictional Film

I Netflix-serien “Black mirror” er det flere episoder som tar for seg ubehagelige tema rundt teknologi og menneskeheten, og en av episodene som har fått mye oppmerksomhet er «Be right back». Den tar for seg relasjonen til en kvinne og hennes avdøde kjæreste, og hvordan AI gjør det mulig for henne å føle at hun kommuniserer med han. Konseptet er at AI lærer hvordan kjæresten oppførte seg ved å se på hans digitale spor, og det viser seg at hun kan være med på å påvirke hvordan han oppfører seg. Ulempen her er at han aldri oppfører seg helt som hun husker, da han ikke viser fram de negative sidene med seg selv. AI utfører alle ordre, og tar ikke egne valg, som gjør kvinnen frustrert, og AI blir aldri det samme som kjæresten.

1.2 Roboter og AI systemer

Historien til begrepet «robot»

I følge artikkel skrevet av Liseter for SNL ble ordet robot lansert av den tjekkiske kunstneren Josef Čapek i og senere brukt av hans bror i et teaterstykke (Liseter, u.å.). Ordet ble forbundet med pliktarbeid, og det kan videre kobles til definisjonene som kommer nedenfor om at en robot skal utføre oppgaver på kommando.

Definisjon av begrepet «robot»

“Actuated mechanism programmable in two or more axes with a degree of autonomy, moving within its environment, to perform intended tasks” er hentet fra “the International Organization for Standardization” via forelesningen “Module 1 - Robots and movement”. Videre finner vi Trenton Schulz sin egen definisjon: *“A robot ... refers to a physical object that interacts with the physical environment, either on its own or via a person, to accomplish a task.”*. Disse to definisjonene viser at tanken bak en robot er at den skal ta kommandoer for å utføre en oppgave, gitt av oss mennesker. Det skal være et fysisk objekt som beveger seg i et fysisk rom.

Relasjon mellom AI og robot

Når jeg først tenker på AI forestiller jeg meg noe som skal ta egne avgjørelser. Sett opp mot en robot er det tydelig at de to definisjonene krever at en robot får instruksjoner av mennesket for å løse en oppgave. På den andre siden tenker jeg at AI er der for å gjøre ting lettere for mennesket, og det samme skal roboten, og sitatene trekker fram at den skal gjøre en fysisk oppgave. AI kan jeg se på som noe litt skjult, og en robot blir mer fysisk.

AV1-roboten

Roboten til No Isolation “AV1” er en fysisk robot som plasseres i klasserommet, der eleven som ikke er til stede skulle ha sittet, og så kan eleven likevel kommunisere og lytte til det som foregår i timen (No Isolation, u.å). Interaksjonen skjer via en smarttelefon eller tablet, der eleven kommuniserer via det, og så kommer det til uttrykk gjennom roboten. Roboten er skapt for at elever som er langtidssyke skal slippe å gå glipp av skolehverdagen.

1.3 Universell utforming og AI systemer

Universell utforming

Inger Marie Lid for SNL definerer universal design eller universell utforming med følgende sitat: “Universell utforming vil si å planlegge produkter, omgivelser, programmer og tjenester

slik at de kan brukes av så mange mennesker som mulig på en likeverdig måte» (Lid, u.å.). Denne beskrivelsen tar for seg ulike elementer vi som mennesker møter i hverdagen, og påpeker hvor viktig det er å tenke på massen som skal møte sidde elementene. Sitatet legger frem at mennesker skal kunne bruke det på en «likeverdig måte», og det tenker jeg er et viktig fokus for å sørge for inkludering i samfunnet som kan omtales som målet med universell utforming.

Potensiale til AI

Talestyrte tjenester som Alexa og Siri har potensiale til å inkludere flere mennesker i bruken av teknologi. For eksempel en synshemmet bruker vil kunne skrive melding ved å si den høyt, og deretter få meldingen opplest. Et annet eksempel er de som bruker skjermleser på datamaskin, for å få lest opp teksten fra for eksempel en nettside.

AI som inkluderende og ekskluderende

Som vist til tidligere er AV1 roboten til No Isolation laget for å inkludere elever som ellers hadde gått glipp av skolegangen.

Når det kommer til chatboter så fungerer de slik at man gir den enkle kommandoer dersom man forstår konseptet av en chatbot. Ulempen er at noen brukere vil skrive slik de gjør til mennesker, og chatboten vil kanskje ikke kommunisere dette godt nok til brukere, og det vil ekskludere noen til å ikke få benyttet seg av verktøyet.

WCAG 2.1

Å forstå et konsept kan knyttes til at du forstår hvordan noe skal brukes, men så kan du ha en forståelse om hvordan det er tenkt å brukes og hvordan det er forståelig å brukes.

De som utvikler AI har i oppgave å lære opp teknologien til å forstå. Med dette som utgangspunkt kan man si at AI er avhengig av mennesket for å kunne lære og i senere tid, kunne begynne å forstå. Om AI kan forstå på eget initiativ er en viktig tanke, men jeg ser for meg at når et menneske må inn for å lære maskinen å forstå, så vil den aldri helt forstå fra et selvstendig perspektiv.

1.4 Retningslinjer for Human-AI interaksjon

Retningslinjer fra Microsoft

Jeg velger å se nærmere på den 12. guidelinen, som tar for seg «Remember recent interactions» (Microsoft, u.å.), der Microsoft velger å trekke frem at AI skal ha et korttidsminne som gjør at data kan hentes frem igjen. Jeg tenker det er relevant å se på hva AI

skal gjøre, for dersom det skal etterligne vår intelligens, er minnet vårt noe som gjør oss unike. Dersom et program skal kunne etterligne kunnskap og minner vi har av hverandre, er det avgjørende at det fokuseres på ved AI. For eksempel en chatbot som er åpen for alle skal kunne svare på enkle spørsmål, men en chatbot med innlogging kan ha muligheter for å finne frem til lagret informasjon om brukeren, for eksempel å tiltale personen ved navn.

Etter å ha sett på Nielsen Norman Group sin oversikt over 10 guidelines for design, vil jeg si at hovedforskjellen kan ligge i hvordan guidelines er beskrevet. Tar vi for oss Microsoft vil jeg karakterisere de som svært konkrete, i den form at det er noe man må tenke på. I Nielsen Norman Group sine formuleringer, viser de til flere eksempler på hvordan heuristikkene kan brukes (Nielsen Norman Group, u.å.). For eksempel viser Microsoft at man skal tilpasse seg brukerens normer, men ikke så mye på hvilken måte. Men Nielsen Norman trekker kanskje frem et mer sosialt perspektiv der det legges fokus på å bruke ord og uttrykk som tilpasses den spesifikke brukeren.

Modul 2

2.1 Karakteristikker ved et AI-system

Sitatet “systems that have features harnessing AI capabilities that are directly exposed to the end user.” viser hvordan Amershi et al. karakteriserer systemer som er i direkte kontakt med brukeren (Amershi et al. 2019). Med dette så utdyper de at egenskapene til AI inkluderes i brukergrensesnitt slik som for eksempel oversettelse av tekst og ansiktsgjenkjenning. Selv om dette er områder der brukeren kan føle de får mye hjelp fra systemet, kommer det ikke uten utfordringer. Amershi et al. trekker tydelig frem utfordringer ved AI-baserte systemer kan gi brukeren «falske positive og falske negative» (Amershi et al., 2019). Dette kan komme av at AI som regel fungerer under automatikk som kan være vanskelig å styre og predikere, spesielt i en setting der AI opererer i bakgrunnen. Forfatterne karakteriserer AI-baserte funksjoner som kan foregå som en synlig funksjon for brukeren eller en funksjon som skjer i bakgrunnen. Synligheten kan vi se i artikkelen til Kocielnik et al. fra 2019, der de trekker fram et AI-basert system som detekterer møteinformasjon i en e-post som brukeren skriver, der brukeren får beskjed om at den har funnet informasjon som kan indikere et møte.

Amershi et al. karakteriserer AI-baserte systemer som systemer som lærer, og at det knyttes en del utfordringer til dette. Med dette så trekker de fram at ulike brukere og situasjoner kan gjøre at den tiltenkte spesialiseringen man ønsker med AI, kan skape en uønsket usikkerhet i hva som faktisk ligger bak brukeren (Amershi et al., 2019).

Fra første forelesning i modul 2 gir Følstad oss fire mer konkrete karakteristikker ved AI systemer. De lyder som følger at AI-systemer kan karakteriseres som systemer som «lærer, forbedrer, black box og baserer seg på store datasett.».

Youtube og AI

En tjeneste jeg benytter meg mye av er Youtube, der det er en personlig startside etter det du som regel pleier å se på. Jeg kan ofte kjenne meg igjen av innholdet som finnes på startsidene og som kommer opp som forslag underveis i bruken. Når det kommer til black box konseptet så kan man tenke på det i sammenheng med det personlige, for man som bruker bare ser resultatet og ikke egentlig hva som ligger bak personliggjøringen. Andre karakteristikker kan kobles til læringen som både Følstad og Amershi skriver om, og det kan komme til syne ved at noen videoer dukker opp ved jevne mellomrom som jeg kan tenke meg at henger med årstiden eller tidspunkt.

Som sagt så kan man se at det skjer mye med Youtube ut ifra hva man har søkt på før og hva som er populært. Amershi skriver at AI ofte opererer i bakgrunnen, og ikke alt er like synlig, og det kommer godt til syne her, fordi bruker ikke ser tydelig hvordan dette gjøres. Det kan være fascinerende og skummelt å se hvordan Youtube «kjenner deg».

2.2 Human-AI interaksjonsdesign

Amershi et al. tar for seg diskusjonen av ulike guidelines for design av AI-systemer. Disse retningslinjene er basert på studier av menneske og AI, og skal være til hjelp for å sørge for bedre menneske-chatbot interaksjon. Videre tar artikkelen for seg utfordring med AI-baserte systemer, og hvordan mennesker kan oppleve kontakten. Noe av motivasjonen for artikkelen er sammenhengen mellom AI-baserte systemer og HCI, og hvordan disse to retningene kan skape utfordringer for hverandre.

Kocielnik et al. tar for seg et fokus på brukerens forventninger til AI-baserte systemet. I forhold til Amershi, som legger frem retningslinjer for design, så retter Kocielnik et al. seg mer mot å styre forventningene til brukeren. Dette kan ifølge forfatterne gjøres ved hjelp av tre teknikker som de omtaler som: nøyaktighetsindikator, forklaringer og kontroll av ytelse (Kocielnik et al., 2019).

G12: Remember recent interactions.

Jeg velger å trekke denne retningslinjen opp mot hvordan Youtube «husker» hva du har sett på sist du brukte systemet. Det kan hende Youtube tar hensyn til denne retningslinjen for å skape en bedre opplevelse for brukerne, spesielt hvis vi ser på videoer som ikke er sett ferdig. Minnet om hva brukerne sist interagererte med kan være med på å skape en bedre opplevelse for brukerne, for de trenger ikke selv å huske hva de sist så på.

G13: Lear from user behaviour.

Relatert til den over så får man opp kanaler man har brukt mye tid på, uten å være abonnent. Dette viser at Youtube følger med på hva man bruker mye tid på, som regel liker og ofte vender tilbake til. På samme måte som Youtube husker nylige interaksjoner, henter algoritmene også informasjon om hva brukere bruker mye tid på, og det kan da tenkes at Youtube tolker dette til at dette er noe brukeren liker.

Fordelen med disse to kan være som jeg har skrevet om før at det føles personlig og at systemet er spesielt for deg som bruker. Muligheten for å predikere innhold og kjenne seg igjen kan kobles til det Amershi et al. (2019) trekker frem i sin artikkel.

Bender et al. – Large Language Model

Asbjørn Følstad reflekterer rundt artikkelen i forelesningen 20.10.2021 Jeg trekker derfor frem konsepter fra hans forelesning, og reflekterer rundt det. Dette gjorde jeg fordi jeg ikke fant artikkelen til forrige modul, men jeg fikk den tilsendt som tilbakemelding.

Et av de mest spennende aspektene var slik Bender et al. ser på det økologiske og økonomiske presset store språkmodeller sørger for. Det krever mye energi og ressurser, og er en stor belastning, slik som Følstad sammenlikner det med Bitcoin, som man vet krever store ressurser. Annet er hvordan disse store språkmodellene kan være med på å spre feilinformasjon i samfunnet, og være med på å påvirke brukeren i en negativ retning.

Et tillegg til Asbjørns forelesning trekker Bender et al. fram viktige aspekter ved konsekvensen av store språkmodeller, og slik dette sitatet viser «the authors note that the majority of cloud compute providers' energy is not sourced from renewable sources and many energy sources in the world are not carbon neutral.», trekker forfatterne frem at det er flere dilemmaer å tenke over ved utviklingen (Bender et al., 2021). Artikkelen trekker også frem risikoer ved systemene, slik som hvordan bias kan påvirke treningsdata «When humans produce language, our utterances reflect our worldviews, including our biases» (Bender et al., 2021).

2.3 Chatbots / samtalebaserte brukergrensesnitt

En ekstra artikkel som vi har lagt til i gruppeoppgaven er skrevet av Brandtzaeg and Følstad i 2017. Denne tar for seg hvorfor mennesker bruker chatbotter, og et av de viktigste poengene er slik som dette sitatet at: «Users' interactions with chatbots often mimic interactions between humans» (Brandtzaeg & Fløstad, 2017). Jeg vil trekke ut fra dette at det man ønsker med en chatbot er å kunne etterligne menneskelig samtale, men at etterligning vil kunne være vanskelig for at man ikke får med det «ekte menneskelige». Som Asbjørn Følstad forteller i forelesningen 20.10.2021 (*Kilde har ikke blitt lagt ut på tidspunktet dette er skrevet. 20.10.2021, 17:18*), om at mennesker tolker mening inn i samtalen, men maskinen mangler denne egenskapen. Kocielnik et al. trekker fram viktigheten av et naturlig språk i samtalen mellom menneske og chatbot (Kocielnik et al., 2019).

Brandtzaeg og Følstad trekker fram at det er viktig å tenke på hvorfor brukere tar i bruk chatbotter. Diskusjonen om ulike motivasjoner peker på at folk bruker det fordi det er effektivt og en enkel måte å ta til seg informasjon. For å sørge for dette mener de at designere må tenke på hvilke situasjoner brukere ønsker et slikt utfall, og sørge for at chatbots faktisk løser problemene bedre og mer effektivt (Brandtzaeg & Følstad, 2017).

De to retningslinjene tar for seg noe av den samme tematikken, som jeg vil omtale som åpenhet om hva og hvordan systemet faktisk fungerer. Slik som jeg har nevnt over er det viktig at en chatbot er designet slik at den sørger for en så naturlig samtale med menneske som mulig. Med tanke på at roboten mangler ekte menneskelige faktorer, det kan være viktig å påpeke dette ovenfor brukeren, med tanke på retningslinjene. G2 skal sørge for at brukeren får en nyanse over hvor god chatboten vil fungere, og det kan være en fordel i dette tilfellet for å gjøre det tydelig for brukeren at chatboten ikke vil fungere på lik linje som et menneske. G1 vil kunne sørge for at brukeren vet at en samtale er mulig, men G2 skaper litt nyanse i hvor godt det faktisk kan gjøres.

Jeg tenker det er veldig viktig å gjøre det klart for brukere at selv om en chatbot kan svare på spørsmål man må ha, så finnes det en begrensning i hva den kan svare på. Som nevnt jobber AI ut ifra et stort datasett, og det kan tenkes at begrensningen til dette datasettet, vil gjøre at chatboten har svaret. Derfor vil fokus på G1 være å få brukeren til å forstå området chatboten kan dekke, og G2 vil sørge for at man ser hvor godt den kan svare.

Vedlegg 1 – Tilbakemelding

Jeg fikk tilbakemelding på å ikke inkludere hele oppgaveteksten, og det er noe jeg har endret på til denne gang. Da jeg skrev forrige tekst hadde jeg knapt med tid, så jeg har jobbet meg igjennom teksten for å se om noe kunne forbedres.

Vedlegg 2 – Tilbakemelding

Jeg fikk tilbakemelding om å forbedre avsnittet mitt om Youtube, og det har jeg gjort til denne innleveringen. «Hva legger du i guidelinen og hvordan kobler du det opp mot eksemplet ditt?» var det jeg fikk spesifikk tilbakemelding på, og jeg håper jeg har forbedret måten jeg får fram min forståelse. Ellers har jeg lagt til tanker om artikkelen til Bender et al., da jeg ikke fant den til forrige innlevering.

Litteraturliste

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P., 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., & 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. (<https://dl.acm.org/doi/pdf/10.1145/3442188.3445922>)
- Brandtzaeg, P. B., & Følstad, A. (2017). Why People Use Chatbots. *Internet Science*, 377–392. https://doi.org/10.1007/978-3-319-70284-1_30
- Bratteteig, T., & Verne, G. (2018). Does AI make PD obsolete? Proceedings of the 15th Participatory Design Conference, 2, 1–5. <https://doi.org/10.1145/3210604.3210646>
- Følstad, A. (2021). Forelesningen: INTERACTION WITH AI – MODULE 2. Session 1. Hentet 20.10.2021 fra: <https://www.uio.no/studier/emner/matnat/ifi/IN5480/h21/interacting-with-ai-2021---module-2---session-1---handout.pdf>
- Grudin, J. (2009). AI and HCI: Two Fields Divided by a Common Focus. *The AI Magazine*, 30(4), 48. <https://doi.org/10.1609/aimag.v30i4.2271>
- Kocielnik, R., Amershi, S., & Bennett, P. (2019). Will You Accept an Imperfect AI? Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–14. <https://doi.org/10.1145/3290605.3300641>
- Liao, Q., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing Design Practices for Explainable AI User Experiences. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–15. <https://doi.org/10.1145/3313831.3376590>
- Lid, Inger Marie: universell utforming i Store norske leksikon på snl.no. Hentet 10. september 2021 fra https://snl.no/universell_utforming
- Liseter, Ivar M.: robot i Store norske leksikon på snl.no. Hentet 10. september 2021 fra <https://snl.no/robot>

- McCarthy, J. (1998). What is Artificial Intelligence? Hentet 10. september 2021 fra <http://cogprints.org/412/index.html>
- Microsoft (u.å.). Guidelines for human-AI interaction design. Hentet 10.09.2021 fra <https://www.microsoft.com/en-us/research/blog/guidelines-for-human-ai-interaction-design/>
- Nielsen Norman Group (u.å.). 10 Usability Heuristics for User Interface Design. Hentet 10.09.2021 fra <https://www.nngroup.com/articles/ten-usability-heuristics/>
- No Isolation (u.å.). AV1. Hentet 10.09.2021 fra <https://www.noisolation.com/no/av1/>
- Schulz, T. (2021). Exploring Moving Things in a Home Environment. Lecture “Module 1 Robots and movement”.
- 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>