Obligatorisk oppgave modul 1 iterasjon 1

1.1 Konsepter, definisjon og historie med interaksjon med AI og autonome systemer

Kunstig intelligens (AI) er ifølge Bratteteig og Verne (2018) et underfelt av informatikk som har som mål å utvikle datasystemer som etterligner menneskelig intelligens. I starten var AI regelbasert, med algoritmer som etterlignet menneskelig problemløsning. Rundt 1950-tallet bidro Alan Turing med betydelige ideer om maskiners evne til å imitere menneskelig tenkning, og under 2. verdenskrig ble regelbaserte systemer brukt til å knekke koder (Grudin, 2009). Etter krigen ble forskning på datamaskiner finansiert, og begrepet «kunstig intelligens» ble introdusert i 1956 av John McCarthy. Siden da har AI utviklet seg raskt, med maskinlæring som et viktig underfelt. Maskinlæring lar datamaskiner lære og forbedre seg ved å gjenkjenne mønstre i store datasett, noe som har revolusjonert AI fra tidligere faste, regelbaserte systemer til mer fleksible og adaptive modeller (Bratteteig & Verne, 2018).

Jeg vil kort redegjøre for tre definisjoner av AI. Den første er fra Bratteteig og Verne (2018), som beskriver AI som «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» (s. 2). Denne definisjonen kommer fra Participatory Design-feltet og sier at AI er datasystemer som kan etterligne menneskelig intelligens. Den andre er fra Jonathan Grudin (2009), som i artikkelen *«AI and HCI: Two Fields Divided by a Common Focus»* skriver at «AI focused on devising better algorithms, HCI on how to improve the use of existing algorithms. AI originated in mathematics and engineering, HCI in psychology» (s. 48). Grudin fremhever hvordan AI og HCI har utviklet seg i forskjellige retninger, til tross for deres felles mål. Til slutt har vi definisjonen fra Jo Herstad (2024), som sier: «The words Interaction, Artificial Intelligence and Autonomous are key concepts that are used both to describe «humans, like you and me» and «things, like a computer or a service» (s. 4). Herstad understreker hvordan begreper som tidligere beskrev menneskelige egenskaper, nå brukes for å karakterisere teknologi. Jeg definerer AI som datamaskiners evne til å simulere menneskelig logikk og beslutningstaking. AI kan overgå menneskelige evner ved å utføre oppgaver raskere, med færre feil, og gjøre ting mennesker ikke kan. Likevel har AI begrensninger, som mangel på menneskelig kreativitet, følelsesmessig innsikt, og nyansert kontekstforståelse.

Forskjellen mellom menneske autonomi og maskinell autonomi

Ifølge Thrun (2004) handler maskinell autonomi om roboter som kan ta egne beslutninger i forskjellige situasjoner, inkludert sosiale interaksjoner. Mennesker tar beslutninger basert på følelser, etikk, og personlig erfaring, mens maskiner følger algoritmer og dataanalyse. Dette reiser etiske spørsmål om ansvar når AI-systemer, som selvkjørende biler, forårsaker skade eller tap av liv (Holmquist, 2017). Maskiner kan ha varierende grader av autonomi, fra fullstendig AI-kontroll til delt kontroll med brukeren, i kontrast til menneskelig autonomi, som er knyttet til personlige verdier og kontekst.

Likheter mellom begrepet automatisk og begrepet autonom

I artikkelen *«Ironies of Automation»* beskriver Bainbridge (1982) automatisering som en erstatning for menneskelig kontroll, planlegging og problemløsning gjennom automatiske enheter og datamaskiner. Begrepene «automatisk» og «autonom» overlapper ved at begge handler om systemer som opererer uten direkte menneskelig innblanding. Automatiske systemer følger forhåndsprogrammerte regler, mens autonome systemer har en høyere grad av selvstyring og kan tilpasse seg endringer og ta egne beslutninger basert på sanntidsinformasjon.

Kort gjennomgang av artikkelen av *«Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle»* av Schulz, Herstad & Torresen (2018).

Artikkelen tar for seg Human-Robot Interaction HRI, og hvordan man kan benytte animasjonsteknikker for å tilpasse robotbevegelser mer i forhold til mennesker. Forfatterne fokuserer spesielt på prinsippet om «Slow In, Slow Out», som er en teknikk innen animasjon der bevegelsen starter og slutter sakte, noe som gir en mer naturlig og menneskelignende bevegelse. Studien analyserer menneskelig og robotbevegelse i hjemmemiljøer og diskuterer hvordan roboter kan tilpasse sine bevegelser for å bedre samhandle med mennesker.

En bedrift som jobber med AI - Telenor

Telenor presenterer kunstig intelligens (AI) som både en plattform og en tjeneste som driver innovasjon internt og eksternt, samtidig som de framhever AI som en transformerende idé med betydelig potensial for samfunnsutvikling. De beskriver AI som en kjernekomponent i deres

overgang til å bli en datadrevet virksomhet, med fokus på optimalisering av operasjoner, automatisering av prosesser og forbedring av kundeinteraksjoner. På nettsiden finnes det en spesifikk side om AI, samtidig som det finnes flere nyhetsartikler relatert til AI.

Kunstig intelligens i Black Mirror

I episoden «Joan is Awful» fra Black Mirror utforskes et dystopisk scenario hvor strømmetjenester bruker en kraftig AI til å lage hyperrealistiske TV-serier basert på vanlige menneskers liv, uten deres samtykke. Hovedpersonen Joan oppdager at hennes daglige liv blir omgjort til en TV-serie, der en populær skuespillerinne spiller henne. Episoden tar for seg hvordan AI-teknologi kan bli brukt til å overvåke, analysere, og deretter produsere innhold uten noen form for etisk vurdering eller hensyn til individets rettigheter.

1.2 Robots and AI systems

I 1979 definerte Robot Institute of America en robot som en «omprogrammerbar, multifunksjonell manipulator» som flytter materialer, deler, verktøy eller spesialiserte enheter gjennom forskjellige programmerte bevegelser for å utføre varierte oppgaver (Thrun, 2004). Merriam Webster's collegiate dictionary definerte robot som «An automatic device that performs functions normally ascribed to humans or a machine in the form of a human.» (Thrun, 2004, s.11). Denne definisjonen beskriver robot som en enhet som kan utføre oppgaver som vanligvis krever menneskelig innsats, samt en maskin som etterligner menneskelig form eller handlinger. Jeg vil selv definere robot som en fysisk maskin eller plattform som kan programmeres for å gjøre spesifikke oppgaver som simulerer eller utkonkurrerer mennesker i forhold til presisjon, effektivitet og kapasitet.

FN klassifiserer roboter i tre hovedkategorier: industrielle roboter, profesjonelle serviceroboter, og personlige serviceroboter, basert på deres bruksområder. Industrielle roboter, som er den eldste og mest utbredte typen, brukes til oppgaver som montering, sveising, pakking og transport, og har hatt en betydelig vekst siden 1960-tallet.

Relasjonen mellom AI og roboter

Basert på min definisjon av robot, kan ikke en chatbot sees som en form for robot ettersom den ikke har en fysisk kropp eller plattform. For å bli betraktet som en robot, måtte den ha en fysisk komponent, for eksempel en fysisk enhet som kan interagere med den virkelige verden. AI gir

roboter intelligens og beslutningstakingsevne, mens roboter gir en fysisk plattform for å manifestere disse funksjonene. AI kan eksistere uavhengig av fysisk maskinvare, som i tilfelle av chatbots og andre programvarebaserte systemer, som beskrevet av Bratteteig & Verne (2018) som «a subfield of computer science aimed at specifying and making computer systems that mimic human intelligence.». Dette betyr at AI kan utføre oppgaver som krever menneskelig intelligens, men uten behov for fysisk kropp.

Eksempel på en moderne fysisk robot

I artikkelen *«Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle»* av Schulz, Herstad & Torresen (2018) blir TurtleBot3 brukt som eksempel på en moderne robot som anvender animasjonsteknikken «slow in, slow out» for å forbedre bevegelsene sine. Denne teknikken gjør at roboten akselererer og bremser gradvis, noe som resulterer i mer naturlige og forutsigbare bevegelser. Når TurtleBot3 navigerer i et hjemmemiljø, gir denne myke bevegelsen brukeren en bedre opplevelse ved å forhindre plutselige bevegelser og kollisjoner, og gjør interaksjonen med roboten både mer intuitiv og komfortabel.

1.3 Universal Design and AI systems

Ifølge Herstad (2024) handler universell utforming om «questions concerning inclusion». Dette innebærer å skape løsninger som sikrer at alle mennesker, uavhengig av deres individuelle behov og evner, kan delta på like vilkår. Målet er å utvikle systemer og produkter som er tilgjengelige og brukbare for et mangfold av brukere, og dermed unngå å ekskludere noen. Det dreier seg om vilkår og muligheter for å inkludere alle. Målet er at nye systemer ikke skal skape barrierer for ulike mennesker. I en video fra The Walrus Talks Inclusion understreker Jutta Treviranus (2019) viktigheten av å anerkjenne menneskelig unikhet. Dette prinsippet står i kontrast til masseproduksjon, massemarkedsføring, massekommunikasjon og enkle offentlige retningslinjer. Ofte designes systemer for «flertallet», der begrensede ressurser brukes for å nå mange. 80/20-prinsippet, hvor 20 % av ressursene gir 80 % effekt, er ikke inkluderende, da opptil 20 % av befolkningen kan bli utelatt. Derfor er det viktig i universell utforming å kontinuerlig spørre: «Blir noen utelatt?».

Hvordan AI kan forbedre persepsjon, bevegelser og kognisjon/følelser

For eksempel kan AI forbedre sanseinntrykk ved å hjelpe synshemmede med å tolke visuelle data, noe som kan gi dem bedre navigasjonsmuligheter. Innen bevegelse kan kunstig intelligens og maskinlæring tilpasse proteser etter individuelle bevegelsesmønstre, og dermed skape mer naturlige og flytende bevegelser. Når det gjelder kognisjon, kan AI støtte personifisert læring ved å tilpasse undervisningsopplegg til den enkeltes læringsstil og tempo, noe som kan forbedre læringsopplevelsen og effektiviteten.

Potensialet til AI for både å inkludere og ekskludere mennesker

Selv om AI har stort potensial for å inkludere nye brukergrupper, er det også risikoer for ekskludering. For eksempel kan AI-teknologier oppleves som en trussel mot arbeidsplasser, da automasjon og maskinlæring kan overta oppgaver som tidligere ble utført av mennesker. Dette kan spesielt ramme grupper som er avhengige av enklere arbeidsplasser, som personer med funksjonsnedsettelser som har tilpasset seg bestemte jobber.

WCAG 2.1 «understand» og «understanding»

I WCAG 2.1 (2023) står det: «Understandable - Information and the operation of user interface must be understandable». Dette betyr at brukere skal være i stand til å forstå både informasjonen som presenteres og hvordan brukergrensesnittet fungerer. Utfordringer knyttet til kunstig intelligens i denne sammenhengen kan oppstå fordi menneskelig kognisjon og handlinger skjer i en kompleks kontekst som ikke alltid kan fullt ut representeres eller forstås av algoritmer og maskiner (Bratteteig & Verne, 2018, s. 2). AI-systemer baserer seg på data og mønstre, men mangler den dype kontekstuelle forståelsen som mennesker har. Dette kan føre til situasjoner der AI gir svar eller utfører handlinger som ikke helt matcher brukerens faktiske behov eller situasjon, da maskiner ikke kan forstå den subtile konteksten som er viktig for menneskelig forståelse.

Retningslinjer for Human-AI Interaction av Microsoft

«02 Make clear how well the system can do what it can do» handler om å sikre at brukeren tydelig forstår hvilke oppgaver AI-systemet kan utføre (Microsoft, 2019). Eksempler på dette kan være å vise hva en navigasjonsapp kan søke etter, forklare hvilke data en aktivitetsmonitor

registrerer, eller informere om hvordan et sosialt nettverk anbefaler personer å følge. Brudd på retningslinjen skjer når systemets funksjoner ikke er tydelig kommunisert, som når en nyhetsfeed eller bildeorganisator ikke opplyser om hva som vises eller hvilke kategorier som er tilgjengelige (Amershi et al., 2019).

Forskjeller mellom HCI design retningslinjer og Human-AI Interaction retningslinjer

Jeg ønsker å sammenligne retningslinje 02 fra Microsofts «18 Guidelines for Human-AI Interaction» (2019) med designprinsippet om «feedback» som beskrevet av Don Norman i *The Design of Everyday Things* (2013, s. 23). Microsofts retningslinje «02 Make clear how well the system can do what it can do» fokuserer på å gi brukerne en klar forståelse av systemets kapabiliteter og begrensninger på forhånd. Dette innebærer å sette klare forventninger til hva systemet kan oppnå før eller under bruken. På den annen side dreier Don Normans konsept om tilbakemelding seg om å gi brukeren umiddelbar informasjon om resultatet av en spesifikk handling, som bekrefter om handlingen var vellykket eller mislykket. Mens «feedback» gir informasjon etter at en handling er utført, søker Microsofts retningslinje å informere brukeren om systemets evner og begrensninger på forhånd. Generelt fokuserer HCI-designretningslinjer på universelle prinsipper som tilbakemelding og brukervennlighet for effektiv interaksjon mellom mennesker og datamaskiner. I kontrast er retningslinjene for Human-AI Interaction spesifikke for hvordan brukere interagerer med AI-systemer, og tar for seg utfordringer som kommunikasjon av AI-kapabiliteter og håndtering av systemets begrensninger.

Dagbok – Refleksjon over daglig bruk av ulike skriveverktøy

Ved å bruke penn og papir la jeg stor vekt på å få håndskriften til å se pen ut. Den taktile følelsen av papir og blekk ga en tilfredsstillende opplevelse, men effektiviteten var begrenset, særlig når det gjaldt redigering og omskrivning. Med QWERTY-tastatur og skjerm økte hastigheten og effektiviteten. Jeg slapp å bekymre meg for håndskriftens utseende, og redigering og navigering ble mye lettere. Men skjermen kunne av og til virke distraherende. Stemmeskrivning introduserte en ny dimensjon ved raskt å fange ideer. Utfordringer med nøyaktighet og feilkorrigering ble tydelige, og selv om denne metoden frigjorde meg fra fysisk skriving, krevde den kontinuerlig overvåking for å rette feil og effektiv organisering og formatering av teksten. Alt i alt foretrakk jeg å skrive med penn og papir når det kommer til å skrive dagbok.

Litteratur

Amershi, S., Inkpen, K., Teevan, J., & Kikin-Gil, R. (2019). Guidelines for human-AI interaction. *In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). https://www.researchgate.net/publication/332742200_Guidelines_for_Human-AI_Interaction

Bainbridge, L. (1982). Ironies of automation. *IFAC Proceedings Volumes*, 15(6), 129-135. https://doi.org/10.1016/S1474-6670(17)62897-0

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

Grudin, J. (2009). AI and HCI: Two fields divided by a common focus. *The AI Magazine, 30(4), 48-57*. https://go-gale-com.ezproxy.uio.no/ps/i.do?p=AONE&u=oslo&id=GALE|A216267417&v=2.1&it=r

Herstad, J. (2024, August 22). IN5620 Interaction with AI and autonomous systems [Lecture notes]. University of Oslo. https://www.uio.no/studier/emner/matnat/ifi/IN5620/h24/fra-forelesning/in5620-lecture-notes-22-august-2024.pdf

Holmquist, L. E. (2017). Intelligence on tap: Artificial intelligence as a new design material. *Interactions*, 24(4), 28-33. https://doi.org/10.1145/3085571

Microsoft. (2019, February 1). Guidelines for human-AI interaction design. https://www.microsoft.com/en-us/research/blog/guidelines-for-human-ai-interaction-design/

Norman, D. A. (2013). *The design of everyday things*. MIT Press. https://www.uio.no/studier/emner/matnat/ifi/IN1050/h23/timeplan/norman-principles.pdf

Sharp, H., Rogers, Y., & Preece, J. (2023*). Interaction design: Beyond human-computer interaction* (Sixth edition). Wiley.

Thrun, S. (2004). Toward a framework for human-robot interaction. *Human–Computer Interaction*, 19(1-2), 9-24. https://doi.org/10.1080/07370024.2004.9667338

The Walrus. (2019, October 25). *We count!* | *Jutta Treviranus* [Video]. YouTube. https://www.youtube.com/watch?v=OAXmCAqZqRk

World Wide Web Consortium. (2023). Web content accessibility guidelines (WCAG) 2.1. W3C Recommendation. https://www.w3.org/TR/WCAG21/