

# IN5480 – Individuell levering

Sigrid Linn

## Modul 1:

- Etter første modul fikk jeg tilbakemeldinger om å forklare mer under hver definisjon, spesielt under definisjonene til de ulike personene for å understreke forskjellene. Det var også et ønske om å forberede kildehenvisningen.

### 1.1 Concepts, definition and history of AI and interaction with AI

Kunstig intelligens er et seksti år gammelt emne som består av en samling disipliner, ideer og teknikker, som blant annet inkluderer matematisk logikk, statistikk, sannsynlighet og datavitenskap. Disse disiplinene tar sikte på å gjenskape et menneskes kognitive ferdigheter. Utviklingen begynte midt i andre verdenskrig, og har ført til at datamaskiner kan utføre stadig mer komplekse oppgaver, som tidligere bare kunne delegeres til et menneske. (Council of Europe Portal, n.d.) Den britiske logikeren Alan Turing er den første kjente til å utføre betydelig arbeid innen kunstig intelligens utført på midten av 1900-tallet. Deretter steg forskningen innen kunstig intelligens videre gjennom 1960-årene.

Ed Burns beskriver kunstig intelligens som dette: «*Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems*»

(SearchEnterpriseAI, 2021)

Videre forklarer han at kunstig intelligens krever grunnlag for spesialisert maskinvare og programvare for skriving og opplæring av maskinlæringsalgoritmer.

John McCarthy sin forklaring om kunstig intelligens: «*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*» (Stanford, 2007)

Han forklarer hvordan intelligens er den beregningsmessig delen av evnen til å oppnå mål i verden. Mennesker, mange dyr og noen maskiner har alle forskjellige intellektuelle nivåer.

Alan Turing beskriver kunstig intelligens som dette: «*The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings*» (Britannica, 2020)

Uttrykket er mye brukt for å referere til et prosjekt som tar sikte på å lage systemer med menneskelignende kognitive evner, for eksempel evnen til å resonnere, skille mening, generalisere og lære av tidligere erfaring.

Min definisjon av kunstig intelligens:

«*Et datamaskinsystem som er designet for å utføre oppgaver basert på etterligning av menneskelig intelligens*».

Valgt artikkel:

«*Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle*», av Trenton Schulz, Jo Herstad og Jim Torresen. (UiO, 2018)

Denne artikkelen beskriver undersøkelsen for robotbevegelser i hjemmet, og klassifiserer bevegelsen sammenlignet med mennesker og deres bevegelser. Dette gjorde de ved å koble kjent bevegelse utenfor hjemmet med den ukjente bevegelsen av robot inne i hjemmet. Fokusgruppen er eldre mennesker, hvor de diskuterer problemene med roboter hjemme og hvordan robotens utseende og bevegelse påvirker dem. På bakgrunn av dette og andre resultater gitt fra forskningen deres, vil de i fremtidige eksperimenter se hvilke animasjonsteknikker som kan være med å gi roboten en særegen måte å bevege seg på.

Veritone er et moderne selskap som jobber med kunstig intelligens. De presenterer verdens første operativsystem for kunstig intelligens, kalt aiWARE, hvor de beskriver at man kan få innsikt i sanntid fra enorme datamengder ved hjelp av en rekke kognitive ferdigheter. De presenterer dette som en service de vil tilby, og presenterer at de er der for å gjøre verden til et smartere, raskere og morsommere sted. (Veritone, 2021)

«The Social Dilemma», utgitt i 2020, er en dokumentar som diskuterer fremveksten av sosiale medier og hvordan det påvirker samfunnet som en helhet. Den dekker ulike aspekter av sosiale medier, som snakker om hvordan den manipulerer enkeltmennesker hvis det ikke

brukes riktig. Dokumentaren bringer også opp problemene med ungdommens mentale helse, og viser hvordan algoritmer for kunstig intelligens kan fungere i henhold til sosiale medier.

## 1.2 Robots and AI systems

«Robot» stammer fra 1920-tallet og betyr «tvangsarbeid» på tsjekkisk. Ordet «Robot» ble først brukt og kjent i det vellykkede skuespillet kalt Rossumus Universal Robots til den tsjekkiske dramatiker, romanforfatteren og journalisten Karel Capek (NPR, 2011). Skuespillet handler om en fabrikk som skaper kunstige mennesker, og i dag brukes flertallet av roboter til å operere i industrien hvor de utfører de samme repetitive aktivitetene om og om igjen.

Cambridge Dictionary definerer robot som:

*«A machine controlled by a computer that is used to perform jobs automatically»*  
(Cambridge, n.d.)

En definisjon som poengterer at det skal være en fysisk maskin som skal kunne utføre jobber automatisk. En veldig rett frem definisjon som forklarer kort og presist, uten andre innvendinger.

Erico Guizzo definerer robot som:

*«Robots are a diverse bunch. Some walk around on their two, four, six or more legs, while others can take to the skies. Some robots help physicians to do surgery inside your body; others toil away in dirty factories»* (Robots, 2020)

En veldig åpen definisjon hvor den inkluderer mange muligheter for hva en robot kan være, og hva den kan være i stand til å utføre.

Sammenlignet med kunstig intelligens, er robot en maskin som har en fysisk form, og som ikke nødvendigvis trenger intelligens for å utføre spesifikke aktiviteter. Kunstig intelligens derimot trenger ikke å være fysisk til stede. Vi finner kunstig intelligens i datamaskiner, telefoner, og også i roboter, men det er en robotform som er ubevegelig og uten kropp. Et eksempel kan være «Chat-bots» som opererer innenfor programmer. Med en Chat-bot kan du stille spørsmål til boten og få et spontant og relevant svar umiddelbart, og den er i stand til å

behandle en stor mengde slike spørsmål uten å trekke en eneste muskel. En robot trenger ikke å etterligne menneskelig intelligens, mens kunstig intelligens må. De er sammenlignbare i den forstand at de både utfører oppgaver og har autonomi. Likevel vil jeg si at kunstig intelligens har større autonomi enn hva en robot generelt har.

NASA sin «Superhero» robot, ble opprinnelig designet og bygget for bruk på den internasjonale romstasjonen. Roboten, som har fått navnet «Valkyrie», kan gå, plukke opp gjenstander, og bruke verktøy på egenhånd. NASA tar sikte på å bruke den til å hjelpe mennesker i faresoner og katastroferammede land en dag (Nasa, 2015). Roboten består av flere innebygde kameraer, opptaks- og sonarutstyr.

### **1.3 Universal Design and AI systems**

En definisjon av universell design er prosessen med å designe og komponere et rom slik at det kan nås, forstås og brukes i størst mulig grad av mennesker i forskjellige alder, størrelser, evner og funksjonshemninger. Et miljø, bygning, et produkt eller en tjeneste bør lages for å passe kravene til alle som ønsker å bruke det. (Universal Design, n.d.) Dette forklarer at god design er å inkludere alle, uavhengig av funksjonshemming. Målet er at alle brukere skal respekteres og ingen skal bli ekskludert.

Kunstig intelligens har et potensiale med å kunne hjelpe mennesker som er avhengige av andre mennesker til daglige oppgaver i hverdagen. Ved hjelp av et «smart-hjem» med hjelpeverktøy som ulike talekommandoer, kan man lettere gjennomføre eller få hjelp til aktiviteter som ellers er vanskelige å gjennomføre uten denne teknologien. Eksempler på dette kan være «Ring indoor-cam», som er et kamera som lar deg se over hjemmet ditt når som helst med direkte overføring. Det kan også være en utendørs sensor som blant annet kan sørge for at lysene i hjemmet ditt skrur seg på automatisk når du kommer hjem.

Til tross for alle fordelene kunstig intelligens gir oss, blir mange av oss bevisst på farene ved skjevhet og diskriminering som disse teknologiene innebærer. Ansiktsgjenkjenning ble vist å være mindre effektivt ved gjenkjenning av ansikter til personer med mørkere hud i et MITstudie, som sammenlignet hvordan teknologien fungerer på mennesker av forskjellige etnisiteter og kjønn. (Pursuit, 2018)

I Human AI-Interaction og i WCAG 2.1, brukes begrepet «forståelse», noe som betyr å kunne gi mening om levert informasjon. Når det gjelder kunstig intelligens og maskiner, tenker jeg at de er i logisk stand til å forstå. På en annen side kan ordet forstå også referere til et mer menneskelig følelsesmessig perspektiv, noe en maskin mangler og ikke kan forstå eller uttrykke tilbake på samme måte som et menneske.

## **1.4 Guideline for Human-AI interaction**

Retningslinjen jeg har valgt er «*Make clear what the system can do*».

Du kan hjelpe brukerne med å ikke bare forstå hva kunstig intelligens er i stand til, men også hva systemet ikke er i stand til. Det vil si at vi først viser kunstig intelligens-systemets begrensninger, slik at brukerne er klar over hva som er, og hva som ikke er oppnåelig når man arbeider med kunstig intelligens. Et eksempel på dette er, at før man kommer med anbefalinger, må man advare brukeren om at kunstig intelligens ikke kan forutsi alderen din automatisk uten at det skrives inn manuelt.

Don Norman har et kjent sett med designprinsipper som består av: Synlighet, begrensninger, tilbakemelding, kartlegging, konsistens og samhandling.

Sammenlignet med Microsofts AI-standarder, som er mer enkle og organisert i flere faser, er Norman sine prinsipper mer abstrakte. De er ganske like i måten de håndterer spørsmål, som tilbakemelding og synlighet, samt det faktum at hovedfokuset er på design for brukervennlige produkter.

## **Modul 2:**

### **2.1 Characteristics of AI-infused systems**

#### **2.1.1 Viktige karakteristikk ved systemer med kunstig intelligens**

Vi lærer om fire viktige aspekter ved kunstig intelligens i dette kurset: Læring, forbedring, black-box, og det faktum at det er drevet av enorme datasett. Kunstig intelligens er lærende,

og tar opp brukerens egenskaper og generell oppførsel mens den er i bruk. De bruker dette til å endre innholdet i brukerens preferanser, i håp om å øke brukerens tid på plattformen.

Forbedring innebærer at det vokser som et resultat av det den har lært. Fordi kunstig intelligens alltid lærer, blir de mange komponentene inkonsekvente når de endres over tid. Dette er i følge Amareshi et al. (2019, s. 2), fordi de til tider kan ha problemer med å forstå oss fordi vi alle oppfører oss annerledes.

Alle de uforståelige skjermene på «interiøret» innen kunstig intelligens er pekt på av BlackBox. Brukeren som administrerer, kan ikke se hvordan interaksjonen oppstår, hvilke input som forårsaker en endring, eller hva kunstig intelligens identifiserer som viktig og hva som er irrelevant. I følge Yang et al. (2020, s. 2), sliter ikke bare den vanlige sluttbrukeren, men selv de som utvikler kunstig intelligens og løsningene, med å forstå hvordan kunstig intelligens fungerer og deres funksjon.

Den siste funksjonen er at kunstig intelligens drives av store datasett, som handler om skalaene der kunstig intelligens må vokse, lære og gjøre en forskjell for både organisasjonen og for brukeren.

### **2.1.2 System som benytter kunstig intelligens – Instagram**

Instagram er en medieplattform som behandler innhold som bilder, videoer og likes, for å presentere anbefalinger til sine 1,4 milliarder brukere (Business of Apps, 2021). Brukere reagerer på innleggene ved å gi «likes», legge igjen kommentarer og bruke «hashtags». All denne handlingen skaper en stor mengde data. Det kan gi stor kommersiell kunnskap og innsikt i menneskelig atferd når de først er analysert, både av mennesker og i økende grad av kunstig intelligens.

For å sikre at brukerne finner verdi i plattformen, er det viktig for Instagram å vise dem hva de vil like. Det blir gradvis vanskeligere å finne informasjon som hver bruker vil finne relevant etter hvert som innholdet stiger. Instagram benyttet seg med å presentere innlegg som de tror brukerne ville glede seg over og dele, og derfor ble maskinlæringsalgoritmer satt i gang for å hjelpe til med å forstå over tid hva som er mest verdifullt og relevant for hver bruker for å lage en tilpasset «feed».

For å bekjempe spam bruker Instagram kunstig intelligens. Spamfilteret kan oppdage falsk kommunikasjon på ni forskjellige språk, inkludert engelsk, kinesisk, russisk, arabisk med fler. Meldinger slettes umiddelbart når de blir identifisert, og den anvender DeepText, et Facebook kunstig intelligens-system som kan forstå konteksten i en melding nesten like godt som mennesker (Business of Apps, 2021).

## **2.2 Human-AI Interaction Design**

### **2.2.1 Hovedtrekk – Amershi et al. (2019) og Kocielnik et al. (2019)**

Hovedtrekkene fra Amershi et al. (2019) er at papiret presenterer 18 generell design retningslinjer for menneskelig AI-interaksjon. Spørsmålene som ble nevnt angår brukerens forventninger og erfaring med et system. Mangelen på kunnskap om hvordan kunstig intelligens fungerer og hva det oppnår, ifølge forfatterne, har en innvirkning på forholdet og bruken av kunstig intelligens i samfunnet. De påpeker blant annet at kunstig intelligens blir vanskeligere å forstå, og at de inkonsekvente interaksjonene dette innebærer er vanskelig å forstå. De benytter seg av «Scheduling Assistant», en kunstig intelligens tjeneste laget av Microsoft Outlooks, for å se på hvilken effekt ulike teknikker har på brukerens forventninger. Her undersøker de virkningen av ulike feil, samt kobler dette til tidligere forskning på hvordan brukere reagerer på falsk positivt og falskt negativt.

Artikkelen forklarer hvorfor retningslinjene er nødvendige, og de argumenterer for at økt mønstergjenkjenning og automatisert slutninger skaper uforutsett atferd, usikkerhet og mulighet for krenkelse av etablerte retningslinjer for brukervennlighet som konsistens, forebygging og forståelse av feil. (Amershi, 2019, s. 3) Artikkelen eksemplifiserer kompleksiteten til kunstig intelligens, og hvor vanskelig det er å designe systemer som kan ta det tekniske, etiske, moralske og menneskelige aspekter tatt i betraktning.

### **2.2.2 Kunstig intelligens sammenlignet med designretningslinjer**

*G5 – Match relevant social norms*

Når det gjelder retningslinjer tror jeg det er umulig å gi et enkelt ja eller nei svar.

Retningslinje G5 er et ja/nei-spørsmål fordi innholdet du får på Instagram, nærmest vil

matche dine relevante sosiale normer, ved å få anbefalinger basert på dine interesser, demografi, og blant annet det du tidligere har likt. Samtidig må huske på at Instagram er en western plattform, noe som betyr at vestlige sosiale og kulturelle normer, kan utgjøre størstedelen av innholdet, til tross for at de har brukere fra hele verden.

### *G8 – Support efficient dismissal*

Denne retningslinjen kan bli forbedret hvis Instagram ga en lett tilgjengelig mekanisme for å avvise eller se bort fra deres AI-systemtjenester. En måte de kan oppnå det på er å gi en liste over funksjoner som brukeren kan «slå av» i innstillingene sine. De kan også merke eller informere brukere når en bestemt funksjon brukes. Brukeren vil kunne lære mer om AI-standarden og kanskje få muligheten til å slå den av.

### **2.2.3 Bender et al. s argumenter rettet mot kunstig intelligens sin påvirkning på tekstinnhold og bruk av store språkmodeller**

I følge Bender et al., (2021), har dype læringssystemer problemer med miljømessige og økonomiske spørsmål, samt tekstinnhold. Dette må bli adressert, spesielt fordi det er anslått at opplæring av bare en av disse store språkmodellene vil forbruke samme mengde energi som et Amerikansk charterfly (Bender et al., 2021, s. 612). Først skisserer de noen av farene, for eksempel det at å ha en stor språkmodell ikke inkluderer variasjon. Dette skyldes forskjeller i internettilgang, kjønn og alder representasjon. En liten gruppe mennesker deler også sine ideer og skaper plattformer eller arenaer for å fremme sitt verdensbilde, eller kanskje skape et ekkokammer (s. 163).

Det er vanskelig å forstå og filtrere ut hva som ligger i dataene, hvordan man kategoriserer det, og analyserer det når det er for mye av det. Videre kan ikke for store datasett dokumenteres tilstrekkelig (s. 1), og Bender et al. (2021), forklarer også hvordan mennesker tolker tekst. Når dette er sammenkoblet med språkmodellens evne til å lære mønstre, er virkelige skader en mulighet (s. 168).

Til dette kommer han med følgende løsning:

Flytte fokus fra å lage modeller, til å lære om hvordan maskiner oppnår mål, og på hvilken måte de tar del i sosio-teknisk systemer. Med dette anbefaler han at forskere og utbyggere av språkmodeller blir mer proaktive og endrer tankegangen.



Andre viktige endringer er nøye planlegging av datainnsamling, ha et nøye utvalg av datasett som er egnet for oppgaven, og vurdere økonomiske og miljøkostnader på forhånd, samt gi grundig dokumentasjon.

## **2.3 Chatbots/Konversasjonsgrensesnitt**

### **2.3.1 utfordringer knyttet til utforming av chatbots**

En av hovedutfordringene med å designe chatbot er å sørge for at chatboten vil svare på en riktig måte i henhold til input. Det er viktig for interaksjonsdesignere å konstant prøve å forberede kunnskapene deres i samtalebrukergrensesnittet i chatbots. Følstad og Brandtzaeg (2017), ønsker at designeren skal gå bort fra en spesifikk oppgave som en «forklarende oppgave», og sette søkelys på å forstå hva brukeren trenger i stedet for å forklare innhold og funksjoner som er tilgjengelige for brukeren.

Følstad og Brandtzaeg (2017), utforsker den skiftende overgangen til chatbots og konsekvensene dette vil ha. De understreker at det alltid vil være krevende at dialogen er det vi trenger å designe, og at designet er knyttet til en lang serviceprosess, og samarbeidet mellom chatbots. Forventninger bidrar til å påvirke brukeropplevelsen, slik Kocielnik et al. (2019), og Amareshi et al. (2019) nevner. Kunstig intelligens som vi møter i chatboter er vanskelig å forutse hvordan de vil handle og reagere tilbake til brukeren. Når forventninger brytes, reduseres brukerens tillitt på chatboten, og kan føre til at brukeren velger en annen måte å kommunisere med.

Viktige årsaker som er nødvendige av menneskesentrert kunstig intelligens, er at kunstig intelligens må tilfredsstille et ekte menneskelig behov, entydige vurderinger av menneskelige evner, og etablere tillit til mottakeren. Dette er også behov som er nødvendige under utvikling av chatbots. Yang et al. (2020), diskuterer hvor krevende det er å demonstrere for brukeren hva kunstig intelligens kan oppnå, og viser hvordan en balanse av kontroll mellom kunstig intelligens og mennesker kan oppnås i fremtiden, noe som vil bidra til å forbedre folks evne til å tilfredsstille folks behov og etablere tillit. Chatbots er unike i det aspektet at det er samtalekunsten som bli avgjørende, og byr derfor på komplekse utfordringer.

### **2.3.2 Retningslinjer G1 og G2**

Retningslinje G1: *Make clear what the system can do*

Hvis en bruker er klar over formålet en chatbots skal gi, kan innspillet de gir være mer tilpasset de svarene som chatboten kan gi. Dette innebærer å oppgi hva systemet kan og ikke kan gjøre, gi tilbakemelding når chatboten ikke forstår, og definere hva som er riktig innspill.

Retningslinje G2: *Make clear how well the system can do what it can do*

Denne retningslinjen kan ha potensiale til å ha en betydelig innvirkning på hvordan mennesker og chatbots samhandler. Chatbots er ofte ment å etterligne mennesker og menneskelig kommunikasjon så nære som mulig, noe som kan få brukerne til å tro at teknologien vil forstå og svare på samme måte som et menneske. Forventingen brukeren har kan bli kontrollert hvis begrensningen til en bestemt chatbot ble gjort mer eksplisitte.

## Modul 2:

- Jeg fikk tilbakemelding hvor det var et ønske om å lese mer om utfordringer knyttet til utforming av chatbots.

## Litteraturliste

Anyoha, R. (2017). *“The History of Artificial Intelligence”*. Tilgjengelig fra: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

Burns, Ed (2021). *“What is artificial intelligence?”* Tilgjengelig fra: <https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence>

Cambridge Dictionary. Tilgjengelig fra: <https://dictionary.cambridge.org/dictionary/english/robot>

Copeland, B.J. (n.d.) *“Artificial Intelligence”*. Tilgjengelig fra: <https://www.britannica.com/technology/artificial-intelligence>

Geek, Uber (2014). *“10 Modern Robots that will shape our future.”* Tilgjengelig fra: <https://wonderfulengineering.com/10-modern-robots-that-will-shape-our-future/>

Guizzo, E. (2020). *“What is a robot?”*. Tilgjengelig fra: <https://robots.ieee.org/learn/what-is-a-robot/>

Iqbal, M., (2021). *Instagram Revenue and Usage Statistics*. Business of Apps. Tilgjengelig fra: <https://www.businessofapps.com/data/instagram-statistics/>

Kisliuk, E. (2017). «R5». Tilgjengelig fra: <https://www.nasa.gov/feature/r5/>

McCarthy, J. (2007). “*What is Artificial Intelligence?*”. Tilgjengelig fra: <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>

Paterson, Maker. 2018. “*Why does artificial intelligence discriminate?*” Tilgjengelig fra: <https://pursuit.unimelb.edu.au/articles/why-does-artificial-intelligence-discriminate>

Schulz, T. Herstad, J. Torresen, J. (2018). “*Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle*”. Tilgjengelig fra: [https://www.uio.no/studier/emner/matnat/ifi/IN5480/h19/undervisningsmateriale/schulz-et-al\\_2018\\_classifying-human-and-robot-movement-at-home-and-implementing-robotmovement.pdf](https://www.uio.no/studier/emner/matnat/ifi/IN5480/h19/undervisningsmateriale/schulz-et-al_2018_classifying-human-and-robot-movement-at-home-and-implementing-robotmovement.pdf)

Ukjent forfatter, (n.d.) “*History of Artificial Intelligence*”. Tilgjengelig fra: <https://www.coe.int/en/web/artificial-intelligence/history-of-ai>

Ukjent forfatter. (n.d.) Universal Design. “*What is Universal Design?*”. Tilgjengelig fra: <http://universaldesign.ie/What-is-Universal-Design/>

Ukjent forfatter. NPR. «*Science Diction: The Origin of the word robot*”. Tilgjengelig fra: <https://www.npr.org/2011/04/22/135634400/science-diction-the-origin-of-the-wordrobot?t=1631128522351>

Veritone. (2021). Hjemmeside. Tilgjengelig fra: <https://www.veritone.com/>