Modul 1

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

Konseptet om at maskiner skal simulere intelligens og kritisk tenkning ble først beskrevet av Alan Turing i 1950. Alan Turing beskrev en test, kalt "Turing test", til å determinere om maskiner er i stand til menneskelig intelligens. Senere beskriver John McCarthy begrepet kunstig intelligens som vitenskapen og konstruksjon for å lage intelligente maskiner. (Home, 2020)

Tre forskjellige definisjoner av AI med beskrivelse og forklaring av disse. I tillegg, en egen definisjon av AI med egen beskrivelse og forklaring.

I artikkelen Artificial intelligence in medicine beskrives kunstig intelligens som: Artificial Intelligence (AI) is general term that implies the use of a computer to model intelligent behavior with minimal human intervention. Artikkelen handler om rollen til kunstig intelligens i den medisinske verdenen, hvordan den har utviklet seg historisk sett og hvordan dens rolle er i dag. I artikkelen om AI i medisin, deler de kunstig intelligens i to grupper; Virtuell (informatikk og maskinlæring) og fysisk (roboter). (VanItallie, T. B. (2017).

Regjeringen beskriver kunstig intelligens basert på EU sin ekspertgruppe slik: "*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. Enkelte KI-systemer kan også tilpasse seg gjennom å analysere og ta hensyn til hvordan tidligere handlinger har påvirket omgivelsene".* De beskriver at kunstig intelligens sanser og tolker data, analyserer og handler ut ifra informasjonen den blir gitt. De deler AI i svake og sterke former for kunstig intelligens og bruker begreper som maskinlæring, maskinsonnering og robotikk. (Kommunal- og moderniseringsdepartementet. (n.d.).)

T.Bratteig og G.Verne (Bratteig & Verne, 2018) beskriver AI som *" a subfield of computer science aimed at specifying and making computer systems that mimic human intelligence or express rational behavior, in the sense that the task would require intelligence if executed by a human".* Artikkelen beskriver et kritisk syn på AI, sett fra en designer perspektiv.

Dersom jeg skulle kommet med min egen definisjon av hva AI er, ville jeg brukt en kombinasjon av de tidligere sitatene. Det de tidligere sitatene har til felles, er blant annet at kunstig intelligens er teknologiske systemer som gjenspeiler menneskelig intelligens . Det nevnes også at AI er et bredt felt og innebærer både virtuelle og fysiske teknologier. Jeg vil derfor definere AI som: **En fysisk eller digitalt teknologisk system som etterligner menneskelig intelligens, ved hjelp av sansing, analysering og selvlæring.**

Likheter og forskjeller mellom menneskelig autonomi og maskin autonomi. Likheter mellom begrepet automatisk og begrepet autonom.

Menneskelig autonomi handler om selvstendighet i den grad et mennesket er i stand til å rasjonere og handle ut fra egen vilje. (Holmøyvik, 2019). Maskiner vil bli mer autonome dersom selvstendigheten til maskinene vil øke. Kunstig intelligens gjenspeiler menneskelig egenskaper, og jo mer intelligent et system er jo mer vil systemet være autonom. Likhetene med maskinenes autonomi og menneskelig autonomi vil være evnen til å handle etter tidligere kunnskap, og at begge kan ta valg basert på tidligere erfaringer. En annen likhet er at begge kan ta avgjørelser basert på hvordan en befinner seg i omgivelsene. Noen forskjeller vil være at menneske har evnen til å handle ut fra følelser og instinkter. Mennesker har evne til rasjonere og ta etiske valg, mens en maskin tar valg basert på tidligere erfaringer og hvordan den er programmert.

Likhetene mellom automatikk og autonomi er at begge utfører en oppgave uten involveringer utenfra. Selv om autonomi handler med om selvstendig læring, er automatiske systemer selvstendige i den grad at systemet utfører oppgaver automatisk. Jeg vil derfor si at det største likhetstrekket mellom autonomi og automatikk er at de utfører oppgaver automatisk.

En kort oppsummering av en artikkel fra pensum.

Jeg valgte artikkelen til Tone Bratteig og Guri Verne: Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design.

I artikkelen diskuteres det hvorvidt kunstig intelligens utfordrer deltakende design (participatory design, PD), og hvordan. I artikkelen utleder de grunnleggende ulikheter mellom PD og AI og hvordan AI utfordrer PD: Det første er at en designer skal kunne forstå teknologien som blir designet for, noe som er vanskelig når AI allerede er en svært krevende teknologi. Det andre er at tidsrommet i en PD prosjekt og AI er svært ulike. Det tredje er hvordan man skal skille når AI er i vanlig bruk og når den trenes opp.

Ved disse utfordringene blir det vanskelig å utføre tradisjonell PD prosjekter med AI, selv om begge har som mål å lage det beste produktet for en bruker. Det argumenteres at PD kan ta del i AI systemer, men at AI sin kompleksitet stiger over PD. De legger til at PD kan bruker som et verktøy for brukere og designere, for å best mulig forhandle rundt utfordringene med AI. (Bratteig & Verne, 2018)

Et moderne selskap som jobber med AI og hvordan dette selskapet presenterer AI som et produkt, som en plattform, en tjeneste, et rammeverk, som en økologi eller som en «ide»?

Jeg kom over det amerikanske startup selskapet SafeBeat RX som bruker kunstig intelligens for å fremvise ECG skanninger (hjerterytme), på en raskere måte. Fremdeles måles disse hjertemålengen får hånd, noe som kan ta lang til, men SafeBeat RX bruker maskinlæring slik at systemet kan gjenkjenne pasientenes hjerterytme, og derfor predikere fremtidig ECG-målinger. Dette selskapet baserer store deler av deres produkt på teknologi innen kunstig intelligens, men kombinerer og muliggjør AI ved hjelp av hjerterytme-fysiker, big-data analytikere og hjerteleger.

En dokumentar, fiksjonsfilm, bok eller spill som handler om bruk og interaksjon med AI og interaksjon mellom mennesker og AI.

Filmen “HER” fra 2013 av Spike Jonze, er en amerikansk film som handler om en ensom tekstforfatter,Theodore Twombly, som forelsker seg i sin virtuelle talebaserte assistent. Filmen er satt i en fremtidig verden der den kunstige intelligensen er svært lik menneskelig intelligens slik at det gjør det vanskelig å skille mellom maskin og menneske. Theodor danner en uvanlig relasjon den virtuelle maskinen, og overbeviser seg selv om at deres relasjon er eksklusivt og spesiell. Utover i filmen går det opp for Theodor at deres relasjon er kunstig, og at den talebaserte assistenten er programmert til å tilfredsstille eierens behov. I dette tilfelle var behovet til Theodore Twombly å redusere ensomheten, og den digitale tjenesten eksiteres til dette formålet. Filmen fremlegger det etisk spørsmålet om menneskelig relasjoner kan likestilles med relasjon med kunstig intelligens.

1.2 Roboter og AI systemer

Ordet “Robot” kommer fra det slaviske “Robota og finnes på flere slaviske språk. Den opprinnelige betydningen er tungt monotont arbeid eller slavearbeid. Ordet Robot fikk en annen betydning da den tsjekkiske dramatikeren Karel Čapek skrev R.U.R. (Rossum’s Universal Robots), som hadde premiere i Praha i 1921 .I stykket R.U.R. tjener arbeids-robotene menneskene ved å gjøre jobben sin, og robotene har derfor ikke behov for følelser eller intellektuelt liv. Etter en stund gjør de opprør og dreper sin menneskelige mester Rossum og ødelegger også alt liv på jorden. I stykket er robotene preget av både overmenneskelig styrke og intelligens. (Wallén, 2008)

To forskjellige definisjoner av "robot". I tillegg, en egen definisjon basert på de to, med beskrivelse og forklaring.

Det Store Norske Leksikon beskriver en robot som: “*en maskin som kan utføre handlinger automatisk, og som ved hjelp av elektroniske «sanser» (sensorer) og «oppskrifter» på hva de skal gjøre (algoritmer) kan bearbeide inntrykk fra omgivelsene, som igjen påvirker hvordan disse oppgavene utføres. En robot består derfor både av programvare (software) som instruerer den i oppgavene, og maskinvare (hardware) som gir den en «kropp» å bevege”*. Her beskrives

en robot som en fysisk maskin, satt sammen av både maskinvare og programvare. Definisjonen gir kriterier om at en robot skal kunne sanse og handle ut fra omgivelser. (Søraa, 2022)

Trenton Schulz beskriver en robot som " *a physical object that interacts with the physical environment, either on its own or via a person, to accomplish a task".* Han beskriver at en robot som er i stand til å handle ut i fra omgivelsene roboten er i. Han beskriver derfor at en robot er et objekt som interagerer med den fysiske verdenen og ikke den virtuelle.

Dersom jeg skal definere hva en robot er, vil jeg basere meg på de tidligere definisjonene. Jeg er enig i at en robot skal kunne være i stand til å utføre oppgaver påvirker den fysiske verdenene. Begrepet "robot" har forandret seg over til, og en maskin som ville blitt kategorisert som en robot før (f.eks. vaskemaskin), ville ikke blitt kategorisert som en robot i dag. Jeg vil derfor definere robot som: **en maskin, som ved hjelp av sansing, prosessering og handling, løser en oppgave som påvirker de fysiske omgivelsene den er i.**

Er "en robot" forskjellig fra "en AI"? På hvilke måter er de forskjellige og like?

Relasjonen mellom AI og robot, ifølge meg, er at man kan finne noen likhetstrekk, men også noen forskjeller mellom de to. Slik jeg beskrev AI og robot, så har begge et formål å utføre en oppgave som innebærer noe ytterligere enn å løse en enkel statisk løsning. Både AI og roboter til dagens standard har blitt avanserte nok til å utføre dynamiske og autonome oppgaver. En annen likhet er at begge interagerer med mennesker og til en viss grad påvirker menneskers hverdag, på en positiv men muligens negativ måte.

De største ulikhetene er at jeg legger vekt på at roboter utfører oppgaver som påvirker de fysiske omgivelsene, der AI ikke nødvendigvis trenger det. AI simulerer intelligens, en verdi som ikke er synlig men en egenskap. Jeg mener at å se på roboter og AI som to ulike teknologier blir litt feil, i og med at de komplimenterer de hverandre godt. En robot er en mer autonom dersom den har en form for AI, og er derfor en "bedre" robot.

Eksempel på en moderne fysisk robot.

S7 MaxV Ultra fra Roborock er populær og privinnende robotstøvsuger. Robotstøvsugeren bruker 3D-lyn-skannings teknologi for å orientere seg i omgivelsene den er i, slik at den på best mulig måte kan bevege seg i et fysisk rom. Roboten er også programmert med maskinlærings algoritmer, slik S7 MaxV Ultra, over tid, vil finne raskere og smartere måter å bevege seg rundt i rommet på. S7 MaxV Ultra er derfor ment for å ha minimal interaksjon med mennesker annet en oppsett, og problemer som kommer i veien for støvsugingen. (roborock, 2022)

1.3 Universell utforming og AI-systemer

Definisjon av universell utforming

Det Store Norske Leksikon beskriver universell utforming som: "*Universell utforming er det å planlegge produkter, omgivelser, programmer og tjenester slik at de kan brukes av så mange mennesker som mulig på en likeverdig måte".* De beskriver universell utforming en metode i design som er inkluderende og ikke diskriminerende. Jeg vil legge til at universell utforming også handler om å inkludere alle, uansett fysiske hemninger, slik at man designer inkluderende løsninger. (Pence, 2021)

Potensialet til AI med hensyn til menneskelig oppfatning, menneskelig bevegelse og menneskelig kognisjon/følelser.

AI kan brukes som et hjelpemiddel slik at det blir enklere å inkludere flere til å ta del i aktiviteter og systemer. AI kan etterligne menneskelige egenskaper slik at mennesker med hemninger kan inkluderes uten hjelp av et annet menneske. Et eksempel kan være talebasete tjenester som leser opp det som foregår på digitale tjenester, eller beskriver omgivelser i et fysisk rom Dette inkluderer mennesker med nedsatt syn. Et annet eksempel kan være eldre eller andre med fysiske utfordringer som kan få hjelp til å utføre hushjelp, som støvsuging og gressklipping, ved hjelp av automatiske roboter som tar i bruk AI for å utføre disse oppgavene på en selvstendig måte.

Potensialet til AI ved inkludering og ekskludering av mennesker.

En implementasjon av AI kan ha som mål å inkludere, eller å gjøre et platform tryggere for flere mennesker, men kan ha som konsekvens å ekskludere enkelte brukere. Et eksempel er algoritmer på sosiale nettsteder som har til formål å sensurere hatkommentarer og diskriminerende utsagn. Dette kan føre til at enkeltmennesker som har meninger som er i strid med den digitale plattformen vil bli sensurert og ekskludert. Det debatteres om dette er et hinder for grunnleggende ytringsfrihet, eller om disse enkeltindividene er med på å skape et skadelig miljø på plattformen. Flere konspirerer også om dersom det allerede skjer, vil det lede til en orkestrert agenda og kan lede til digital autoritarisme (Wright, 2018).

Hvordan AI er inkluderende er nevnt i forrige oppgave.

Forstår maskinene?

Det å forstå og det å lære, mener jeg, er konsepter som går mye inn i hverandre. Når mennesket lærer over tid utredes det til forståelse, og på denne måten vil man danne seg evnen til å resonnere og reflektere rundt det man har forstått av læringen. Maskiner derimot, har ikke evnen til å reflektere rundt erfaringer den lærer seg underveis. Jeg mener derfor at maskiner har evnen til å lære, hvor maskinlærings teknologien har kommet langt, med at maskiner kan løse oppgaver basert på tidligere erfaringer, men at de ikke kan resonnere og forstå hvorfor oppgavene blir løst.

1.4 Retningslinjer for Human-AI interaction

Retningslinjer for human-AI interaction fra Microsoft og sett med HCI-design retningslinjer fra nettet.

Microsoft guidelines for human-AI interaction nr.11: Make clear why the system did what it did:

Dette punktet handler om at et AI system skal gi en tilbakemelding på hvorfor den gjorde det den gjorde. Det finnes flere måter å gi en tilbakemelding på, enten det er talebasert eller en bevegelse som tydeliggjør begrunnelse for aktiviteten. På denne måten minimeres eventuelle forvirringer, og forårsaker at interaksjonen med systemet bedres.

Jeg fant Shneiderman's 8 Golden Rules of Interface Design:

Shneiderman legger vekt på at systemet skal være enkelt og visuelt. Det forklares at feilmeldinger skal forklares og at systemets struktur er logisk bygget opp. Ved dette vil systemet inkludere flere og interaksjonen vil flyte bedre. Her legger både Microsoft og Shneiderman vekt på at systemet skal gi tilbakemeldinger og visuelle uklarheter. Likevel legger Microsoft mer vekt på systemets tekniske AI løsninger skal være tydelige og at bruker skal kunne blir bedre på å bruke systemet over tid. (Wong, 2020)

(Modul 1 oppdatert etter tilbakemelding etter iterasjon 1. Måtte forandre på hvordan jeg skulle sitere, og fikse opp i skrivefeil.)

Modul 2

2.1 Kjennetegn på AI-tilføyde systemer

AI-tilføyde systemer er systemer som har funksjoner som utnytter AI-egenskaper, og som er direkte eksponert for sluttbrukeren (Amershi et al.2019). Slike AI-tilføyde systemer har spesielle egenskaper. Nevnt fra første forelesning i modul 2 er typiske egenskaper evnen til å lære, forbedre, black box ("usynlige" prosesser) og drives at store datasett. (Følstad)

Ved disse egenskapene stiller Amershi et al. seg kritisk til AI-tilføyde systemer og nevner at "*automated inferences are typically performed under uncertainty, often producing false positives and false negatives, AI-infuse systems may demonstrate unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous" (Amershi et al.2019)*. AI komponentenes evne til å lære over tid vil føre til inkonsekvent adferd. Dette gjør at AI-tilføyde systemer vil reagere ulik fra en bruker til en annen eller reagerer ulikt på samme type input over tid. Dette er med å forvirre brukerne og kan føre til mistillit til AI-teknologi. "*AI functionalities are probabilistic and almost always operate at less than perfect accuracy. How- ever, most users do not expect their applications to behave inconsistently and imperfectly*"*(Kocielnik et al., 2019)*.

Som egenskap hos AI-tilføyde systemer forklarer Yang et al: *"Some AI systems continue to learn from new data after deployment. In the ideal case, the system will "grow," integrating new insights from new data and adapting flexibly to more varieties of users and use contexts" (Yang et al., 2020).* Det forklares at når systemet fortsetter å lære vokser systemet og datasettet fortsetter å utvikle seg basert på brukerkontekst. Allikevel vil utviklende data føre til at systemytelsen kan gå i feil retning og er en del av AIs evne usikkerhet.

Eksempel på et AI-tilføyd system

Et eksempel på en AI-tilføyd system er Grammarly. Grammarly er en digital skrive assistent som vurderer blant annet stavefeil, grammatikk og tegnsetting i engelskspråklig tekst. Basert på kunstig intelligens finner Grammarly erstatninger og feil tilpasset en spesifikk bruker. Med moderne teknologi innen maskinlæring og deep learning lærer maskinen seg hvordan bruker skriver og kan komme med skreddersydde tilpasninger. Grammarly lærer over tid, og samler store mengder data jo mer man bruker systemet. (*Grammarly About Us*, n.d.)

Grammarly lærer seg hvordan type tekst en bruker skriver, allikevel vil ikke alle vurderingene Grammarly kommer med være til forventningene. Dersom man skriver en tekst ved fremmede begreper, vil Grammarly gi falske negative output, og er med på å ødelegge skrive flyten. Jeg erfarer spesielt når man skriver en tekst med begreper fra et fag på skolen, eller en rapport med flere begreper fra informatikken, gir Grammarly forslag som man ikke er interessert i. Til slutt ender man med å ikke ta i bruk systemet i spesielle tilfeller, fordi man forventer at det ikke vil fungere slik man vil.

2.2 Human-AI interaksjonsdesign

Sentrale take-aways fra Amershi et al. (2019) og Kocielnik et al. (2019)

Med stor fremskritt innen AI, presenterer Amershi et al.2019 en problemområde mellom menneske-datamaskin-interaksjon. AI-tilføyde systemer er ofte inkonsekvente og uforutsigbare og kan lede til negative brukeropplevelser. Som svar på dette foreslår de 18 generelle designretningslinjer som et verktøy på videre utvikling av AI-tilføyde systemer og som retningslinjer for videre forskning rundt human-AI interaksjonsdesign. Målet med disse retningslinjene er å utvikle bedre og flere brukerrettet AI-tilføyde systemer slik at den økende inkluderingen av AI vil legge til rette for menneske-AI-interaksjonen.

Selv om det har blitt gjort grundig studie rundt maskinlæring og AI-algoritmer, har ikke undersøkelser rundt brukerens forventninger blitt like grundig studert. I artikkelen til Kocielnik et al. 2019 tar de derfor for seg brukerens forventninger til AI-tilføyde systemer ved hjelp av AI basert planleggingsassistent. Videre foreslår de tre teknikker for å sette brukerens forventninger: 1) Nøyaktighetsindikator - som eksplisitt angir nøyaktigheten til systemet, 2) Eksempelbasert forklaring - som søker å øke brukerforståelsen, og 3) Ytelseskontroll - som lar brukeren direkte justere ytelsen til systemet. Ved eksperimentelle undersøkelser finner de at teknikkene kan brukes for å øke brukerforventninger. Artikkelen konkluderer med at fokus på høy precision må balanseres med like høy recall , slik at brukernes oppfatning av systemets aksept og nøyaktighet ikke blir redusert.

Design guidelines

Design Guideline G9:

“Support efficient correction: Make it easy to edit, refine, or recover when the AI system is wrong.“

Dette følger ikke Grammarly helt. Systemet vil komme med flere forslag til hvordan man kan forbedre teksten sin. Dersom man velger å redigere det man har skrevet, legger Grammarly inn dette i teksten. Problemet oppstår dersom man angrer på Grammarly sitt forslag, den eneste måten å gå tilbake på, er å manuelt redigere teksten. Om man har valgt flere vurderinger fra Grammarly og velger å trekke det tilbake alt sammen, blir dette en større jobb. Med enklere oversikt over hva som har redigert og en “tilbake” funksjon ville gjort det tunge arbeidet lettere, og bruker ville hatt mer positiv erfaring med bruken av systemet.

Design Guideline G2:

“Learn from user behavior. Personalize the user’s experience by learning from their actions over time.”

Dette punktet følger Grammarly. Systemets selvlæring er tungt prioritet og Grammarly er flink på å komme med forslag der en bruker ofte gjør feil. Systemet blir bedre over lengre tid, men dette gjelder dersom bruker også er flink på å bruke systemet. Om Grammarly har nok data vil vurderingene fra systemet være relevante og mer spesifikke.

Problematiske sider ved generering av store språkmodeller.

Artikkelen til Bender et al (2021) kommer med flere problematiske aspekter ved AI-tilføyde systemer som baserer å generere store språkmodeller. Et problem artikkelen presenterer handler om hvordan generering av store språkmodeller krever mye energi, og konserverer i negativ påvirkning av klima og miljø. I artikkelen fremviser de en urovekkende sammenligning som beskriver hvordan et menneske bruker i snitt 5T CO2e per år, mens trenings prosedyren til en stor Transformer brukte 284T CO2 per år. Mye av denne energien er ikke hentet fra karbon-nøytrale kilder eller fornybare kilder. Selv energien som er hentet fra fornybare kilder er både kostbart og kunne blitt brukt av andre brukere.

Bender et al (2021) argumenterer også rundt et annet problem rundt miljø. "*And, while some language technology is genuinely designed to benefit marginalized communities, most language technology is built to serve the needs of those who already have the most privilege in society. Consider, for example, who is likely to both have the financial resources to purchase a Google Home, Amazon Alexa or an Apple device with Siri installed and comfortably speak a variety of a language which they are prepared to handle" Bender et al (2021)*. De forklarer hvordan land som blir miljø-diskriminert ved global oppvarming er land som allerede er verdens mest marginaliserte samfunn. Av dette belønnes de som har mest, mens de utsatte samfunnet får de negative konsekvensene.

Et annet problem som blir fremhevet er hvordan trening med store mengder data fra internett leder til at de som ytrer nedsettende meninger rundt kjønn, rase, etnisitet og funksjonshemming er de som blir representert. " *In the case of US and UK English, this means that white supremacist and misogynistic, ageist, etc. views are overrepresented in the training data, not only exceeding their prevalence in the general population but also setting up models trained on these datasets to further amplify biases and harms" Bender et al (2021)*. Konsekvensen av dette fører til at datatreningen rettes mot slike meninger, øker diskriminering og gir skjevhet i samfunnet.

Et annet problem som blir argumentert i artikkelen er hvordan man kan være sikker på hvilken data som blir brukt ved læring av systemet. Modellen har ingen forståelse av meninger, noe som kan være risikabelt dersom dataene er feil, eller at det blir brukt manipulerende data. Til slutt stilles spørsmålet: *" If a large LM, endowed with hundreds of billions of parameters and trained on a very large dataset, can manipulate linguistic form well enough to cheat its way through tests meant to require language understanding, have we learned anything of value about how to build machine language understanding or have we been led down the garden path?"Bender et al (2021)*

En løsning på problemet Bender et al (2021) presenterer er at man er nødt til å godta at det er en risiko ved AI-tilføyde språkmodeller, og man må vurdere om fordelene oppveier risikoen. Forfatterne oppfordrer til å nøye vurdere data slik at utviklingen ikke går i en negativ retning og være forberedt på at teknologi påvirker mennesker. Det forklares at tidlig dokumentering og dataselektering i designprosess vil reduserer de problematiske aspektene ved AI-tilføyde systemer som baserer å generere store språkmodeller

2.3 Chatbots/ konverserende brukergrensesnitt

Chatbots

Chatbots er en populær type AI-tilføyd system. I artikkelen "Chatbots and the New World of HCI" av Følstad & Brandtzaeg (2017) beskriver de chatbots økende popularitet og dens potensielle design utfordringer. Et eksempel på en utfordring er hvordan fokus på grafisk brukergrensesnitt har blitt redusert og det vektlegges mer mot det språklige brukergrensesnittet. Det grafiske grensesnittet har det blitt gode undersøkelser rundt, men dersom de store teknologiske selskapene retter seg mot språklige grensesnitt, kan det føre til store endringer HCI feltet. Man har i flere tiår blitt vant med å benytte designmetoder for visuelle oppsett, dersom man bruker noen av de samme designmetodene for konverserende oppsett vil det lede til hulrom og implikasjoner.

En annen implikasjon ved chatbots er at man må heller tenke på design som en tolkningsoppgave framfor en forklarende oppgave. Det vil si at man må bevege seg vekk fra å designe forklaringer rundt hvilke innhold og funksjoner bruker må ta for å oppnå målet sitt, men heller gå i retningen mot å forstå hva bruker ønsker og hvordan man kan oppnå brukerens mål. *"The need to move from user interface design to service design" Følstad et al (2017).*

Et annet problem som blir diskutert er hvordan digitale teknologier, som chatboter, er med på å skape forskjeller og skjevheter på tvers av kjønn, alder og samfunnsstatus. Det antas at chatboter ikke diskriminerer, men dersom chatboten ikke har et godt utformet naturlig språkgrensesnitt som ikke støtter digitale teknologier kan dette føre til utestenging av marginaliserte grupper.

Guideline G1 og G2

Guideline G1 handler om at det skal være tydelig hva et AI system er i stand til å gjøre. Dette kan man implementere ved å gi en åpningsmelding når man åpner opp chatenboten slik at brukeren kan forvente hva systemet er i stand til å gjøre. En mulig implikasjon ved dette er som nevnt i Følstad & Brandtzaeg (2017) er at denne type løsning er basert på visuell grensesnitt, framfor en løsning som baserer seg på språk og kommunikasjon. En optimal løsning er å tydeliggjøre hva chatbotsen kan gjøre gjennom samtale, og ikke at brukeren blir forklart hvordan systemet funker.

Guideline G2 handler om å tydeliggjøre for bruker hvor godt systemet får til å gjøre det den gjør. En implikasjon ved en løsning på dette kan være likt det som ble beskrevet i G1 om grafisk og språklig grensesnitt. En annen mulig implikasjon kan være at bruker ikke forventer å få en beskrivelse av hvor godt systemet får til oppgavene den får. Dette kan på noen områder gi en negativ opplevelse, dersom bruker kun vil bruke chatbot som et raskt virkemiddel for få svar på det en har lyst til å få svar på.

(Oppdatert etter tilbakemelding etter iterasjon 2. Måtte skrive hele teksten på enten norsk eller engelsk. Måtte også drøfte og reflektere litt mer. (dette har jeg prøvds på så godt jeg kunne))

Referanseliste

Modul 1

Bratteig, T., & Verne, G. (2018, August). Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design. Retrieved September 13, 2022, from https://www.researchgate.net/publication/327131040_Does_AI_make_PD_obsolete_exploring_challenges_from_artificial_intelligence_to_participatory_design

Holmøyvik, E. (2019, august). *autonom – Store norske leksikon*. Store norske leksikon. Retrieved September 13, 2022, from https://snl.no/autonom

Home. (2020, October). History of artificial intelligence in medicine. Retrieved September 13, 2022, from https://reader.elsevier.com/reader/sd/pii/S0016510720344667?token=6C873BA1BCDD42E35067C40EE8FA738E6D2F49143C1D74008C4F7AFAE99361E489EFDD4F00F5F418DAB324D06FC4F2FC&originRegion=eu-west-1&originCreation=20220913122215

Kommunal- og moderniseringsdepartementet. (n.d.). *Nasjonal strategi for kunstig intelligens - regjeringen.no*. Regjeringen.no. Retrieved September 13, 2022, from

https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf

Pence, C. (2021, December 27). *universell utforming – Store norske leksikon*. Store norske leksikon. Retrieved September 13, 2022, from https://snl.no/universell_utforming

roborock. (2022). *Roborock S7 MaxV-serien – Alt blir enkelt*. Roborock Norway. Retrieved September 13, 2022, from https://no.roborock.com/pages/roborock-s7-maxv

Schulz, T. (2020). *Exploration of Moving Things in the Home*. Exploration of Moving Things in the Home. Retrieved September 13, 2022, from https://www.duo.uio.no/handle/10852/74061

Søraa, R. A. (2022, May 12). *robot – Store norske leksikon*. Store norske leksikon. Retrieved September 13, 2022, from https://snl.no/robot

VanItallie, T. B. (2017, April). *Home*. History of artificial intelligence in medicine. Retrieved September 13, 2022, from https://reader.elsevier.com/reader/sd/pii/S002604951730015X?token=D3C05A1473E2DF8DC9F4141EF9EBC626774FA6704A57C126330F5CFE514994D1C9CB98F40253B72B110A8CA529D6A485&originRegion=eu-west-1&originCreation=20220913130121

Wallén, J. (2008, mai 8). *The history of the industrial robot*. DiVA portal. Retrieved September 13, 2022, from https://www.diva-portal.org/smash/get/diva2:316930/FULLTEXT01.pdf

Wong, E. (2020, July 31). *Shneiderman's Eight Golden Rules Will Help You Design Better Interfaces*. Interaction Design Foundation. Retrieved September 14, 2022, from https://www.interaction-design.org/literature/article/shneiderman-s-eight-golden-rules-will-help-you-design-better-interfaces

Wright, N. (2018). *DIGITAL AUTHORITARIANISM*. How Artificial Intelligence Will Reshape the Global Order. Retrieved September 14, 2022, from https://alinstitute.org/images/Library/ArtificialIntelligenceReshapeGlobalOrder.pdf

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. (https://www.microsoft.com/en-us/research/uploads/prod/2019/01/Guidelines-for-Human-AI-Interaction-camera-ready.pdf)

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)

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. (https://www.microsoft.com/en- us/research/uploads/prod/2019/01/chi19_kocielnik_et_al.pdf)

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). (https://dl.acm.org/doi/abs/10.1145/3313831.3376301)

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

Grammarly About Us. (n.d.). Grammarly. Retrieved November 21, 2022, from https://www.grammarly.com/about