



Nya etiska dilemman inom Artificiell Intelligens

En analys av en statlig myndighets jämställdhets och jämlikhetsarbete inom AI.

Författare: Hedvig Abrahamsson

Institution: Etnologi, Religion och Genusvetenskap

Examensarbete 30 HP

Genusvetenskap

Mastersprogram Genusvetenskap 120 HP

Vårterminen 2022

Handledare: Sam Holmqvist

New ethical dilemmas in Artificial Intelligence – An analysis of how a Swedish state institution works with DEI in AI.

Abstract

This study explores the ways a Swedish state agency works with equality during the making of their AI-chatbots. The study explores an AI-team's work in dealing with questions of equality, with equality through own experiences and values, the study has also placed its central themes around equality work and language. The study does this with help of what Sara Ahmed calls institutionalized work. The study concludes that what Edmund Husserl calls the natural attitude to working with equality constitutes a risk of leaving out perspectives and experiences. Working with Swedish as the state language, the study discusses how the use of language might represent whiteness and what effects this use of language may have on the interaction with their chatbot. The study has based its material around interviews made with six members of the state agencies AI-team as well as programmed words and sentences in the chatbot.

Nyckelord: vithet, vitifiering, självmodellen , maktstrukturer, AI, chatbot, jämställdhetsarbete samt jämlikhetsarbete, naturlig inställning.

Innehållsförteckning

.....	0
Abstract	1
1.Introduktion	3
1.1. Syfte och frågeställningar	4
1.2 Bakgrund	5
1.2.1 Vad är artificiell intelligens?	5
1.2.2 Artificiell Intelligens och myndigheter	6
1.2.3 Vithet och Artificiell Intelligens	7
2. Forskningsöversikt	8
2.1 Allmän översikt	8
2.2 Medvetenhet och etik	10
2.3 Vithet och Artificiell Intelligens	13
2.4 Språk och kritiska vithetsstudier	14
3.Teori.....	16
3.1 Fenomenologi	16
3.2 Institutionella vanor och vithet	19
3.3 Språkliga maktstrukturer	20
3.4 Självmodells teori om subjektivitet	22
4. Material och metod.....	23
4.1 Material och studieobjekt	23
4.1.1 Bakgrund om studieobjekt	24
4.2 Semistrukturerade intervjuer	25
4.2.1 Förarbete	26
4.2.2 Genomförande och tematisk analys	26
5. Etik	27
6. Analys.	29
6.1 Arbete med jämställdhet och jämlikhet	29
6.2 Språk	40
6.3 Vithet inom Artificiell Intelligens	58
6.4 Neutralitet	63
7. Slutsatser	65
8. Vidare forskning.....	68
9. Referenser:	69
10. Bilagor:	73

1. Introduktion

Föreliggande studie avser att analysera jämställdhets och jämlikhetsarbete i byggandet av artificiell intelligens i form av en chatbot på en svensk statlig myndighet. Artificiell Intelligens (AI) är enligt Steven G. Chapel och A.B. Simmons definierat som en maskin som kan bete sig på sätt som för en människa skulle anses intelligent.¹ Alan Turings artikel ”Mind” publicerad år 1950 i en filosofisk tidskrift, som medförde stor påverkan inom AI-s utvecklingsområden.² Som idag gett oss chatbotar, automatiserade beslutsprocesser och andra typer av AI som underlättar för människor offentligt och privat.

AI lär sig hela tiden av oss, härmar det vi kanske tar för givet om det programmeras in genom omedveten partiskhet. Riskerna med detta är flertaliga. Reproduktionen av ett segregerat samhälle inom teknologin, precis som det vi lever i idag är enbart ett av alla exempel som visar hur stor påverkan AI har. Kommer nuvarande diskurser påverka hur AI erkänner och inkluderar vissa kroppar mer än andra? Historiskt har AI representerats som något vitt, Steve Cave och Kanta Dihal beskriver kopplingen mellan teknologi och vithet hela vägen tillbaka till kolonialismen, genom en ’överlägsen’ teknologisk utveckling, något som i stor utsträckning fortfarande utgör representationen av rasifiering inom AI.³ Vitheten inom AI kopplas ofta till språk, dialekter och kulturella markeringar, dessa kulturella markeringar medför kulturspecifika ideologier och stereotyper. Dessa stereotyper och ideologier överförs sedan till AI-n och precis som med kolonialismen uppstår här ojämna maktrelationer.⁴

Teknologins koppling till vithet förstärks genom vita kroppar som programmerar, vita värderingar och vita språk som anses universella och därmed i flera fall inte uppmärksammas. Cave och Dihal menar att det som anses intelligent fortfarande centreras kring det vita, något som de kopplar tillbaka till användningen av teknologi för att överta och kolonisera mindre teknologiskt utvecklade länder.⁵ I dag så finns det regelverk för hur AI ska

¹ Simmons A.B. and Chappell S. G, "Artificial intelligence-definition and practice," in *IEEE Journal of Oceanic Engineering*, vol. 13, no. 2, pp. 14-42, April 1988, doi: 10.1109/48.551 Sida: 18.

² Haenlein, M. and Kaplan, A. (2019) "A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence", *California Management Review*, 61(4), sida 5–14. doi: [10.1177/0008125619864925](https://doi.org/10.1177/0008125619864925). Sida: 6-7.

³ Cave, S, Dihal, K. (2020) "The Whiteness of AI." *Philos. Technol.* **33**, 685–703 (2020). <https://doi.org/10.1007/s13347-020-00415-6>

⁴ Marino, Mark C. "The Racial Formation of Chatbots." *CLCWeb: Comparative Literature and Culture* 16.5 (2014): <<https://doi.org/10.7771/1481-4374.2560>

⁵ Cave och Dihal (2020), Sida: 687

programmeras på EU-nivå, dock saknas detta globalt. EU-riktlinjerna finns till för att upprätthålla en standard för hur en tillförlitlig AI behandlar moraliska frågor, dessa riktlinjer ska utgöra en utgångspunkt när europeiska institutioner och företag skapar en egen AI.⁶ Kate Crawford beskriver i sin bok *Atlas on AI* hur de etiska riktlinjer som finns globalt har skapats av rika länder och de länder som påverkas negativt av AI-n får inte en möjlighet att höras i dessa riktlinjer.⁷ Etiska riktlinjer utgör något som inte enligt flera forskare tar en stor nog plats i utvecklingen av AI, bristen på riktlinjer riskerar att skapa och upprätthålla maktstrukturer och stereotyper som överförs från samhället. Bristen på riktlinjer gör att frågan om hur organisationer arbetar med AI blir relevant.

Hur arbetar då myndigheter med AI och jämställdhet och jämlikhet idag 2022? I diskussioner med bekanta till uppsatsförfattaren inom IT och programmerings industrin kom det fram att det finns en brist på perspektiv och kunskap som kan göra AI-n användbar för fler och huruvida AI representerar vissa mer än andra.⁸ Studien centreras därmed att se hur en svensk statlig myndighet arbetar med sin AI i jämställdhets och jämlikhetsfrågor. Särskilt intresse för att analysera ett myndighetsarbete med dessa frågor uppkom genom dess betydelse för samhället, en myndighet som är till för medborgare och som ska vara tillgänglig för dessa. Myndigheter framställs som auktoritära med en maktposition, hur de applicerar jämställdhets och jämlikhetsarbete på sin egen AI tjänst framstår därmed som något intressant.

1.1. Syfte och frågeställningar

Syftet med studien är att analysera jämställdhets och jämlikhetsarbete i byggandet av en chatbot på en svensk statlig myndighet, samt hur chatboten rasifieras genom myndighetens AI-team sätt att arbeta i samband med språkanvändning. Syftet är att genom intervjuer analysera vad begrepp som naturligt och det bekanta betyder för hur det statliga AI-teamet programmerar och avgränsar chatboten. Chatboten utgör en utgångspunkt för studien för att se hur myndighetens jämställd och jämlikhetsarbete och språk tar sig uttryck i interaktioner med offentligheten.

⁶ Kate, Crawford (2021). *The atlas of AI power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press Sida: 224–225

⁷ Crawford (2021), Sida: 224

⁸ Cave och Dihal (2020), Sida: 687.

Frågeställningar:

- Hur arbetar myndigheten med jämställdhet och jämlikhet inom AI?
- Hur producerar arbetssättet en potentiell rasifiering av chatboten?
- Hur kan en potentiell rasifiering skapa maktstrukturer och upprätthålla sociala ojämlikheter överförda från samhället till teknologi?

1.2 Bakgrund

Artificiella intelligensens historia kan enligt Michael Haenlein och Andreas Kaplan spåras till 1940 talet och mer specifikt till när science-fiction författaren Isaac Asimov publicerade sin bok *Runaround* år 1942. Historien centreras kring två karaktärer och *Three laws of robotics*, en roman som inspirerade flera forskare inom robotik. Vid ungefär samma tidpunkt började den engelska matematikern Alan Turing arbeta med utvecklingen av verklighetstroga koder inom AI. År 1950 publicerar Turing sin artikel "Computing Machinery and Intelligence" där han beskrev hur en skapar intellektuella maskiner och hur en testar deras intelligens. Turing-testet används än idag.⁹

1.2.1 Vad är artificiell intelligens?

Artificiell intelligens är uppbyggd av något som kallas algoritmer och data, dessa algoritmer liknar Virginia Dignum vid ett recept på äppelpaj. En algoritm blir enligt Dignum ett resultat av ingredienserna i äppelpajen. Varje steg i receptet utgör ett steg i skapandet av AI, här tydliggörs människors roll som grundsten i byggnaden av AI och riskerar att påverka slutresultatet. Resultatet är därför beroende av vilken data som programmeras in, precis som en äppelpaj är beroende av de olika ingredienserna.¹⁰ Jan Weyerer och Paul Langer presenterar i sin artikel "Garbage In, Garbage Out: The Vicious Cycle of AI-Based Discrimination in the Public Sector" hur det som programmeras in i artificiell intelligens får en direkt översättning i hur AI:n kommer att agera. De beskriver detta i form av AI-baserad diskriminering som de

⁹ Haenlein and Kaplan (2019), Sida: 6-7.

¹⁰ Dignum, Virginia. (2020). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer Nature Switzerland AG. Sida: 102.

beskriver genom tre olika steg: 1. AI Input, vad som läggs in i form av mänskliga värderingar, 2. AI learning där det finns tre olika steg, där AI-n läser och processar värderingar. Resultatet av föregående steg är det som sedan integreras och skapar basen för hur databasen fungerar och svarar. Steg nummer tre blir att de mänskliga värderingarna överförs och blir AI-ns värderingar.¹¹

1.2.2 Artificiell Intelligens och myndigheter

Idag arbetar flera svenska myndigheter med AI, till stor del för att effektivisera arbetet. Det kan handla om en chatbot som ska underlätta trycket på en viss del av myndigheten, den som har kontakt med medborgare. En chatbot kan definieras som en mjukvaruapplikation som ofta används för att skapa och genomföra en chattkonversation online, som ett alternativ till kontakt med en mänsklig agent.

År 2021 fick fyra svenska statliga myndigheter i uppdrag att arbeta med AI, för att främja det svenska offentliga förvaltningsförmåga i användandet av AI, för att stärka Sveriges välfärd och konkurrenskraft till andra länder. Uppdragets stora uppgift landade i att ta fram en AI-guide för offentlig förvaltning. Avsikten var att ta fram en modell för hur en verksamhet ska ta sig fram för att använda sig av AI, inkluderat är även hur AI ska struktureras och tillgängliggöras.¹²

AI utgör idag en stor del av den industriella marknaden, Professor Spyros Makridakis skriver om den revolution AI har medfört. En revolution som har potentialen att ersätta, förstärka och komplettera i stort sett alla arbetsuppgifter utförda av människor. En revolution som kan ge positiva effekter för företag som behöver öka produktionsprocesser, men som samtidigt riskerar att höja arbetslösheten.¹³ EU parlamentet beskriver AI-s betydelse som något som hjälper oss inte bara i vardagen genom digitala röstassistenter, men även som något som har en stor betydelse för cybersäkerhet och kampen mot desinformation, genom att kunna detektera falska uttalanden online. AI har också enligt Europaparlamentet en stor betydelse inom hälso- och

¹¹ Weyerer, Jan & Langer, Paul. (2019). "Garbage In, Garbage Out: The Vicious Cycle of AI-Based Discrimination in the Public Sector." 509–511. 10 1145/3 325 112 3 328 220. Sida: 509–511.

¹² Larsson, S., & Ledendal, J. (2022). "AI i offentlig sektor: Från etiska riktlinjer till lagstiftning." I: K. de Vries, & M. Dahlberg (Red.), De Lege årsbok 2021: Law, AI & Digitalization (s. 467-495). (De Lege årsbok; Vol. 2021). Iustus förlag. Sida: 467.

¹³ Spyros Makridakis, "The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms, Futures", Volume 90, 2017, Sida: 46-60, ISSN 0016-3287, <https://doi.org/10.1016/j.futures.2017.03.006>. Sida 48–49.

sjukvården, genom att kunna identifiera mönster som kan leda till nya upptäckter. AI har även en betydelse för oss under vår fritid, du kanske har en roomba eller ett kylskåp anslutet till ditt Wifi.¹⁴ Inom fältet av AI finns det olika riktlinjer för hur arbetet ska utföras på ett etiskt sätt. Inom EU finns rapporten ”Etiska riktlinjer för tillförlitlig AI” som beskriver hur man bygger en tillförlitlig AI, detta dokument gäller för alla inom EU. Dessa riktlinjer är till för att guida och skapa en grund för hur en bra AI ska programmeras. I rapporten presenteras även möjligheter och kritiska frågor som kan uppkomma och väckas av AI.¹⁵

1.2.3 Vithet och Artificiell Intelligens

I antologin *Om ras och vithet i det samtida Sverige* skriver Tomas Hübinette, Helena Hörnfeldt, Fataneh Farahani och Rene Leon Rosales i det inledande kapitlet om kritiska rasstudier och dess snarlika kusin Kritiska vithetsstudier. Författarna presenterar forskningsfältet som fortfarande är relativt nytt i Sverige, något som gjort att det fortfarande inte fått möjlighet att etablera en legitim plats i det akademiska kretsarna. Antologin har med hjälp av teorin och forskningsfältet till syfte att ge ett alternativt sätt att se på Sverige i dess nuvarande samhällsställning gällande vithet och makt.¹⁶ Vithetens historia inom teknologi beskrivs av Dihal och Caves två forskare på Leverhulme Centre for the Future of Intelligence, University of Cambridge i England. I deras artikel ”The Whiteness of AI” drar de kopplingen mellan vithet och teknologi hela vägen tillbaka till kolonialismen, där de beskriver att de vita kolonistörerna behövde vara teknologiskt överordnade för att erövra land. Utgångspunkten i artikeln centreras i det vita rasiala ramverket som beskrivs som följer av Joe R. Feagin “the White racial frame includes a broad and persisting set of racial stereotypes, prejudices, ideologies, interlinked interpretations and narratives, and visual images”.¹⁷ Cave och Dihal visar på vithetens existens inom AI genom språk, robotarnas utseende och hur AI uppfattas utifrån synen på vithet som

¹⁴ Nyheter, Europaparlamentet. Samhälle. ”Vad är artificiell intelligens och hur används det?” Publicerad 04-09-2020 - 12:38 <https://www.europa.eu/press/2020/09/04/ai-artificial-intelligence-how-it-is-used>

¹⁵ AI HLEG (oberoende expertgrupp skapad av Europeiska kommissionen juni 2018) ”Etiska riktlinjer för tillförlitlig AI”, Europeiska kommissionen, offentliggjordes 8 april 2019.

¹⁶ Hübinette, Tobias, Hörnfeldt, Helena, Farahani, Fataneh & León Rosales, René (red.) (2012). *Om ras och vithet i det samtida Sverige*. Tumba: Mångkulturellt centrum Sida: 14.

¹⁷ Feagin, J. R. (2013). *The white racial frame: centuries of racial framing and counter-framing* (2nd ed.). New York: Routledge. Sida: XI.

det 'rationella'.¹⁸ De demonstrerar vithetens utbredning i AI genom en Google-bildsökning med ordet AI. Här kommer nästan uteslutande vita konstruktioner av AI upp.¹⁹

2. Forskningsöversikt

Forskning inom AI har under de senaste åren utvecklats och mer och mer fokus har lagts på den sociala delen av AI. Perspektiv utgår ifrån social rättvisa, vem som blir representerad och hur samt vilka konsekvenser detta medför. Föreliggande studie skrivs in i fältet genom fokus på arbete med jämställdhet där kön och lika rättigheter mellan dessa sätts i fokus och jämlikhet där alla oavsett etnicitet, språk och hudfärg får samma möjligheter och rättigheter. Dessa begrepp representerar i studien två separata begrepp. Samt språkbruk i AI-n, givet de forskningsfrågor som ställts i undersökningen har relevant tidigare forskning valts ut för att ge en bakgrund inom ämnet.

2.1 Allmän översikt

Under de senaste åren har forskningen kring AI i etiska och sociala sammanhang ökat enligt en studie ledd av Anna Felländer, grundare till företaget AI Sustainability center, ett svenskt företag som arbetar för att utveckla och presentera innovativa metoder för att skapa hållbar AI. Svensk forskning från Vinnova Sverige genomförd av AI Sustainability center har kartlagt användning av AI i offentlig sektor. Projektets syfte var att studera och inventera kunskapen kring etiska, sociala och lagliga utmaningar i anslutning till AI och maskinlärning.²⁰ Projektet motiveras av behovet att av nya innovativa metoder att utveckla och arbeta inom fältet har identifierats. Felländer et al beskriver att flera företag under de senaste åren startat program för att undvika negativa konsekvenser i sitt arbete baserat i artificiell intelligens. Företag som Microsoft, IBM och Google, även Europeiska unionen har genomfört forskning inom artificiell intelligens, i dessa rapporter presenteras och betonas vikten av att adressera etiska utmaningar

¹⁸ Cave och Dihal (2020) Sida: 697.

¹⁹ Ibid

²⁰ Larsson.S, Anneroth, M. Felländer, A, Felländer-Tsai, L. Heintz, F, & Cedering Ångström, R (n.d.) (2019) "Sustainable AI [Review of Sustainable AI.]" Sida: VIII-IX

inom självständiga tekniska system och AI.²¹ EU:s forskning presenterar tre förslag för att tackla dessa utmaningar. Europeiska kommissionen föreslår även förberedelser för de socioekonomiska förändringar AI kommer medföra samt att skapa ett tydligt regelverk för en ny modern transformativ teknologi.²²

Flera inflytelserika organisationer och institutioner som ADM och IEEE har även inlett forskningsprojekt för att betona vikten av forskning inom AI, etik och hållbarhet.²³ Forskning inom området anses vara ny och extremt viktig för en hållbar framtid. IEEE har tagit initiativet vidare och upprättat ett certifikationsprogram för autonoma system, system som kan styra sig själva oberoende av utomstående entiteter. Fokuset läggs på transparens, ansvar och att minska algoritmisk partiskhet det vill säga partiskhet i koden som skapar AI. Utanför programmet skapat av IEEE finns det konceptuella metoder för att arbeta mot vad Felländer et al kallar ”Good AI”, ett exempel på detta är ”Montreal Declaration for a responsible development of artificial intelligence”.²⁴

Stefan Larsson och Jonas Ledendal publicerade år 2022 en undersökning där de beskriver hur användningen av AI inom den offentliga sektorn i Sverige ser ut. Genom sin undersökning visar Larsson och Ledendal hur svenska myndigheter arbetar med AI, samt vilka utmaningar och möjligheter som uppkommer i arbetet. Larsson och Ledendal inleder med en presentation av den relativt mjuka inställningen till AI, genom att exemplifiera den så kallad AI-Agendan för Sverige²⁵ som publicerades 2020. Agendan innehåller inga tydliga strategier för hur arbete med AI ska bedrivas. Här efterlyses i stället AI-rådgivare och -utvecklare som kan stå för en ledningsposition i förändringsarbetet med AI. I agendan presenteras även behovet av framtida multidisciplinär forskning som ska driva AI forskning framåt.²⁶ Larsson och Ledendal beskriver vidare DIGGS rekommendation att utöka användningen av AI i den offentliga sektorn samt hur det kan nyttjas för att möta flera samhällsproblem, såsom en ökad och mer jämlik tillgång till

²¹“Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines,” April 25th, 2018. Brussels. Inga sidnummer tillgängliga.

²² Ibid.

²³Larsson, Anneroth, Felländer, Tsai, Heintz & Cedering Ångström (2019) Sida: 10–11.

²⁴[The Declaration - Montreal Responsible AI \(montrealdeclaration-responsibleai.com\)](https://montrealdeclaration-responsibleai.com) Hämtad den 17/7-2022

²⁵ Larsson och Ledendal (2022), Sida:474.

²⁶ Ibid, Sida:474.

sjukvård oavsett ekonomiska tillgångar.²⁷ DIGG förslår även att kompetensen inom utvecklingen av AI behöver utvecklas och uppdateras.

2.2 Medvetenhet och etik

Larsson och Ledendal presenterar även hur den kritiska medvetenheten inom AI har vuxit på senare år, de tar som exempel ansiktsigenkänning i USA, där personer med mörk hy och med kvinnliga ansikten inte uppfattas av kameran på samma sätt som en vit man, något som har lyft debatten och medvetenheten om diskriminering inom AI globalt, Vilket har lett till att fler vetenskapliga perspektiv betonas.²⁸ Uppmärksamheten kring AI och diskriminering har gjort att utvecklingen kring principiella ställningstaganden i offentlig sektor och civilsamhället har möjliggjort och utvecklat etiska och rättighetsbaserade riktlinjer, framtagna för skapandet av AI.²⁹ Dessa har senare visat sig tolkas olika och ges olika prioriteringar av olika aktörer, något som kan utgöra något farligt där aktörer inte anpassar sig efter riktlinjer och visar på de relativt mjuka riktlinjerna.

Riktlinjerna kring AI är i stora företag till största del etiska riktlinjer, etiska riktlinjer som Larsson diskuterar vidare med hjälp etikforskaren Thilo Hagendorff. Hagendorff menar att etiska riktlinjer saknar mekanismer och metoder för att implementera samt upprätthålla och skapa en efterlevnad av dessa.³⁰ Larsson beskriver vidare hur Hagendorff menar att avsaknaden av mekanismer och metoder för implementering hos de etiska riktlinjerna är varför företag och organisationer föredrar att använda sig av dessa, för att undgå rättslig granskning. Dessa etiska riktlinjer skapar en motsättning till rättsliga regelverk, något som medför en individbaserad kontroll av AI som i sig medför risker och undviker utomstående granskning. Användandet av etiska riktlinjer inom AI kan därmed enligt Hagendorff tolkas som ett sätt att undvika rättslig reglering.³¹

²⁷Ibid, Sida:467–468.

²⁸ Larsson och Ledendal (2022), Sida: 479–480.

²⁹ Crawford (2021), Sida: 224.

³⁰ Larsson, S. (2020) AI i EU: Etiska riktlinjer som styrmedel, I: Bakardjjeva Engelbrekt, A., Michalski, A. & Oxelheim, L. (eds.) *EU och teknologiskiftet: Europaperspektiv 2020. Årsbok för europaforskning inom ekonomi, juridik och statskunskap.* . Santérus förlag, p. 89 120 p. (Europaperspektiv). Sida: 105–106.

³¹Ibid, Sida: 106.

Kate Crawford, en forskare inom AI, tidigare medgrundare och forskningsansvarig för AI Now Institute, presenterar sin syn på AI i boken *Atlas of AI*.³² Boken presenterar en ny insyn i hur AI diskuteras och framställs nästan som ett magiskt fenomen men som i själva verket döljer en mörk sida av diskriminering och förstärkning av strukturella ojämlikheter. Crawford tar upp frågan huruvida vi borde demokratisera AI, på detta svarar Crawford att detta är kantat av hinder. Hinder som det faktum att AI representerar makt och centraliserad kontroll, hon liknar detta vid att demokratisera vapenproduktionen för att nå fred. Crawford diskuterar vidare som svar på frågan om en demokratiserad AI på skrivningen av 'AI-ethics principles', dessa principer har tagits fram av ekonomiskt utvecklade länder, där Afrika, Syd och Centralamerika och Centralasien har en mindre representation. De som skadas mest av AI systemen blir därmed inte inkluderade i skapandet av dessa principer. Som Larsson menar blir dessa etiska riktlinjer ett sätt att motsätta sig rättsliga regleringar och därmed kontrollera riktlinjerna kring arbetet i individuella sammanhang. Crawford refererar vidare till Shannon Mattern som påpekar bristen på etiska principer i genomförandet och skapandet av AI. Mattern påpekar även hur AI fältet inte har någon officiell övervakning, strukturer eller normer, det finns inte heller någon gemensam överenskommelse kring definitioner och mål inom fältet.³³ Fältet struktureras i stället av enskilda självreglerande organisationer och institutioner.

Crawfords forskning tillsammans med Larsson och Felländer et al inom fältet är av intresse för föreliggande studie, givet synen på AI som något mer än ett mystiskt och magiskt fenomen.³⁴ Framtoningen av det underliggande sociala, politiska och maktperspektiven blir relevant att tillämpa på ett statligt myndighetsarbete med AI som interagerar med medborgare. År 2018 antog EU uppdraget att skapa en lagstiftning kring AI och dess användande inom Europa, där de uppmuntrade alla medlemsstater att skapa en egen strategi kring AI baserat på lagstiftning.³⁵ Initiativet togs med anledning av det som Mattern beskriver nämligen bristen på en tydlig och gemensam lagstiftning kring AI i Europa.

Felländer et al presenterar fyra kategorier de anser vara extra viktiga att undersöka i användningen och utvecklingen av AI. Dessa fyra kategorier består av "ansvar: erkännande och fördelning", "missbruk och illvillig användning", "transparens och förklarbarhet" samt "bias och partiskhet". Partiskhet inom AI menar Felländer et al är något som i stor utsträckning är en del av AI.³⁶

³² Crawford (2021), Sida: 222–225.

³³ Ibid

³⁴ Ibid

³⁵ Ibid

³⁶ Larsson, Anneroth, Felländer, Tsai, Heintz & Cedering Ångström (2019) Sida: XV, 10–11.

Felländer et al fortsätter att beskriva partiskhet och mer specifikt oavsiktlig partiskhet, som de menar utgör flertal av fallen inom partisk AI. En oavsiktlig partiskhet innebär att social konstruerade fördomar och tankar påverkar vad som programmeras in i AI ofta omedvetet, vilket i sig ger AI en oavsiktlig partiskhet mot vissa personer. Utmaningen menar Felländer et al ligger mellan de inneboende sociala strukturerna och historiska ojämlikheter. Dessa strukturer manifesteras sedan i systemen menar Felländer et al. Om informationen som matas in i systemen är partisk, kommer informationen att återskapas i systemet, något de menar är problem inom till exempel tillämpningar för jobbsökningar. Deras forskning presenterar även andra sätt som partiskhet skapas inom AI, detta genom värdebaserade preferenser hos skapare av system och även användare av system. Med detta menar de att egna erfarenheter påverkar hur vi tolkar och uppfattar resultat i systemen.³⁷ Felländer et al beskriver kort varför ansvarstagande är viktigt att titta på inom AI, genom att jämföra med självkörande bilar. Om en av dessa krockar, vem blir då ansvarig för krocken? För att undvika sådana situationer menar de att en algoritm måste köras flera gånger för att få möjlighet att identifiera och ta ansvar för misstag.³⁸

Under ”missbruk och illvillig användning” diskuterar Felländer et al huruvida enskilda programmerare ska hållas ansvariga för kränkande och illvilliga användningar av AI, något som flera forskare menar att en liten del av ansvaret måste ligga på de som skapat AI-n. Felländer et al menar att en starkare kultur av ansvarstagande över hur verktyg kan användas måste etableras, med betoning på utbildning, etiska standarder och normer.³⁹ Vidare poängterar författarna hur AI är byggt av autonoma system, system som lär sig av det som programmeras in och utvecklar kunskapen med fortsatt interaktion eller programmering. Även här kommer problemet med ansvarstagande in i bilden, menar författarna. Inneboende social partiskhet blir synliggjord i dessa system och blir i sig normativa och riskerar att upprätthålla en normerande bild av samhället. Felländer et al menar i slutändan att programmerarna ska till en viss del hållas ansvariga för de konsekvenser deras programmering skapat, de menar även att dessa problem är något som programmerare måste bli medvetna om.⁴⁰ Hur partiskhet ska undvikas i programmering är i dagsläget omdebatterat i flera olika fält, metoder för att identifiera partiskhet inom algoritmer existerar redan i dagsläget. Dessa metoder blir däremot ofta blockerade på grund av amerikanska CFAA [Computer Fraud Abuse Act] som möjliggör för

³⁷Ibid, Sida: 13.

³⁸Ibid, Sida: 16.

³⁹Ibid, Sida: 16.

⁴⁰Ibid, Sida: 17–18

myndigheter och företag att exkluderas från granskning av produkter för partiskhet.⁴¹ CFAA är en amerikansk lag emot data intrång som tillåter myndigheter och organisationer att använda sig av 'skyddade datorer' för att minska tillgången till viss information.⁴²

2.3 Vithet och Artificiell Intelligens

Enligt Yarden Katz, professor i kognitionsvetenskap, har AI alltid producerat rasifierade, klassificerade och könade modeller av jaget. Katz menar att vitheten i AI även har en mer subtil framtoning, här pekar han på vithetens ideologi och hur den återskapas inom AI. Vidare pekar Katz på AI och dess egenskaper som han liknar till vitheten. AI menar Katz är en produkt av den vita världseliten och som en konsekvens av detta blev AI avbildad till att likna den vita ideologin, genom sina oklara och dynamiska egenskaper, styrda av kapitalism och imperialistiska mål.⁴³ Katz benämner AI som resultat av historien som en "technologies of whiteness", något som han förklarar med hjälp av två argument: AI är en teknologi som arbetar för vithetens mål och blir därmed ett verktyg för att upprätthålla en vit 'överlägsenhet', Katz menar vidare att speglingen av den vita ideologin i AI blir något dynamiskt och diffust. Katz menar att denna diffusa egenskap gör att AI kan tillämpas och anpassas till flera rörelser och fenomen precis som vitheten.

Katz forskning är applicerbar i föreliggande studie genom den historiska synen på vithet inom AI tillsammans med speglingen av en vit ideologi i teknologin, detta för att kunna diskutera hur chatboten hos myndigheten potentiellt speglar en vit ideologi, liknande det som Felländer et al diskuterar kring egna erfarenheter samt hur dessa riskerar att påverka systemens produktion.⁴⁴ Cave och Dihal beskriver precis som Katz vitheten i AI som en dominant ideologi, de menar att avbilden av vithet inte bara presenterar sig genom utseende på robotar utan också genom det talade språket. De menar att en vit representation genom språk, utseende och andra kopplingar

⁴¹Ibid, Sida: 19.

⁴²IT-ord, ord och uttryck inom IT Branschen. 15 april 2019 <https://it-ord.idg.se/ord/computer-fraud-and-abuse-act/> hämtad den 18/8-2022.

⁴³ Yarden, Katz (2020). *Artificial Whiteness: Politics and Ideology in Artificial Intelligence [Elektronisk resurs]*. Columbia University Press. Sida: 154.

⁴⁴ Ibid, Sida: 153-156.

till vitheten riskerar att upprätthålla och förstärka redan existerande stereotyper och sociala rasifierings hierarkier.⁴⁵

Cave och Dihal drar denna slutsats genom att titta på det som kallas det vita ramverket, där en konstruktion av en viss social status får stå för det som representerar synen på vad AI är förbundet med. AI representerar det som anses vara det intelligenta och önskvärdt utseendemässigt något som Cave och Dihal menar är något som allmänheten uppfattar som något vitt. Hur detta framställs i AI-n bidrar till ett upprätthållande av dessa uppfattningar om intelligens och vithet. Partiskhet diskuteras som något som upprätthåller social orättvisa och partiskhet i konstruktionen av AI påverkas av vilka som blir anställda för att skapa AI-n. Genom kontrollen av rekrytering kan därmed företag skapa en egen bild av hur AI-n ska vara, något som även riskerar att leda till homogena arbetsgrupper som i sin tur riskerar att upprätthålla sociala hierarkier. Cave och Dihal kommer även fram till att vitifiering av AI har skapat en ny nivå inom den sociala hierarkin, här får en vit AI en plats i den mänskliga sociala hierarkin som överordnas personer som rasifieras. Detta betyder att rasifierade personer hamnar i en position underordnad maskinen.⁴⁶

2.4 Språk och kritiska vithetsstudier

Språk och kritiska vithetsstudier tillämpas i studien för att möjliggöra diskussion kring chatbotens språk och hur språket kan skapa en identitet eller representera en rasifierad grupp. Sara Trechter och Mary Bucholtz, forskare på California State University och Texas A&M University beskriver i artikeln ”White Noise: Bringing Language into Whiteness Studies” vithet i språk som något med avsaknad av färg och kultur.⁴⁷ I stället för närvaro handlar vithet i språk om dess frånvaro. De beskriver vitheten som något som kan kasta sig ideologiskt som den enda synliga och hörbara rasifieringen av grupper detta kallar de för ”whiting out” andra subjekt.⁴⁸ Trechter och Bucholtz menar att vitheten i språket behåller sin placering genom en historisk norm eller en ”single conceptual category” som alla förstår. De menar även att språk och ideologi hänger ihop, de väljer att identifiera ideologi genom Foucault. Trechter och Bucholtz

⁴⁵ Cave och Dihal (2020), Sida: 690.

⁴⁶ Ibid, Sida: 687.

⁴⁷ Trechter, Sara, and Mary Bucholtz. ”Introduction: White Noise: Bringing Language into Whiteness Studies.” *Journal of Linguistic Anthropology*, vol. 11, no. 1, 2001, pp. 3–21, <http://www.jstor.org/stable/43103950>. Sida 4.

⁴⁸ Ibid, Sida: 5.

menar att ideologier av en rasifierad grupp överförs på språk, därmed skapas en dikotomi mellan det markerade (icke-vita) och det icke-markerade (vita) språken.

Trechter och Bucholtz presenterar ett underliggande problem som centralt att: vithet inte bara är en identitet men mer fundamentalt en ideologi.⁴⁹ Författarna diskuterar språk och identitet, de menar att identitet är en kulturell och social process. Undersökningen presenterar att identitet och ideologi spelar en stor roll i skapandet av vithet i språk. Språk inom AI kan även visa på partiskhet, något som författarna Su Lin Blodgett, Solon Barocas, Hal Daumé III och Hanna Wallach diskuterar i artikeln ”Language (Technology) is Power: A Critical Survey of “Bias” in NLP” språk och sociala hierarkier. I artikeln beskriver de hur sociolingvister och antropologer under senare tid har analyserat inställningar kring språk och språk ideologier och personers metalingvistiska inställningar till språk. Vilka språkvariationer anses bli identifierade som standard och omarkerade?⁵⁰ Även vilka språk som anses korrekta och lämpliga att använda i offentliga bruks sammanhang är frågor som Blodgett et al diskuterar. De menar precis som Trechter och Bucholtz att språkliga ideologier utgör en av grunderna för att försvara och upprätthålla sociala hierarkier, eftersom språkliga ideologier ofta har en betydelse för hur vi uppfattar talaren. De pekar på vikten av att förstå rollen språk har i neurolingvistiskt programmering (NLP) för upprätthållandet av sociala hierarkier, något som de beskriver som avgörande för framtida arbeten med partiskhet i NLP.⁵¹ Mark Marino författare och akademiker inom digitala studier beskriver i sin artikel “The racial formation of Chatbots” hur skapandet av i det här fallet en chatbot är att skapa en rasifierad produkt. Marino menar att språk skapar och representerar kulturer i textbaserade chatbotar något han tydliggör genom att referera till Robert Hodge och Gunther Kress som menar att språk är bärare av medvetenheten i samhället. Genom detta bjuder Marino in till att alla chatbotar utgör en möjlig rasifiering.⁵²

Även Hübinette, Hörnfeldt, Rosales och Farahani beskriver en koppling mellan språk och vita kroppar. De skriver i antologin ”Om ras och vithet” hur det svenska språket enklare tillämpas på vita kroppar. I föreliggande studie kommer språk att utgöra en central del av undersökningen, precis som tidigare forskning presenterar avser studien att undersöka hur språket representerar, vem språket representerar samt hur detta görs.⁵³ Kopplingen mellan språk och vithet diskuteras

⁴⁹Ibid, Sida:6.

⁵⁰Su, Blodgett &, Solon Barocas &, Hal III &, Hanna Wallach. (2020). “Language (Technology) is Power: A Critical Survey of “Bias” in NLP. 5454-5476. 10.18653/v1/2020.acl-main. Sida: 5458.

⁵¹Ibid, Sida: 5458.

⁵² Marino (2014).

⁵³ Hübinette, Tobias, Hörnfeldt, Helena, Farahani, Fataneh & León Rosales, René (red.) (2012). Sida: 53–54.

vidare av Katarina Mattsson. Mattsson beskriver fem kategorier som associerar en person som svensk. Första kategorin menar Mattsson handlar om födelse inom Sverige, är du utlandsfödd eller invandrare hamnar du i en annan kategori. Kategori två beskriver Mattsson som medborgarskap, här finns det två underkategorier: svenskar som föds in i medborgarskap och invandrare som får sitt medborgarskap och som då ofta benämns som ”naturaliserade svenskar”. Kategori tre menar Mattsson består av blodsband, om båda dina föräldrar är födda i Sverige. Kategori fyra består av språk och kultur, där du som född i Sverige anses vara ”kulturell svensk”, något som sedan kopplas till en rad olika egenskaper till följd av att ha svenska som modersmål. Kategori fem menar Mattsson utgår ifrån utseende, där personer delas upp utefter hur svenska de anses se ut. Ser du inte svensk blir du indelad i kategorin som utgörs av personer med ”annat utseende” som Mattsson beskriver ofta talas om i form av ”synliga minoriteter”.⁵⁴

3. Teori

3.1 Fenomenologi

Sara Ahmed är en brittisk kultur och samhällsteoretiker. Hennes fenomenologi och orientering används i föreliggande studie för att analysera hur uppfattningar och inställningar till arbetet med jämställdhet och jämlikhet skapas. Hur påverkas informanten av sin omgivning genom hens omgivning? Ahmeds syn på medvetenhet används i studien genom att undersöka vilka objekt och tankar som får ta plats i jämställdhets och jämlikhetsarbete. Studien använder sig även av Ahmeds reflektioner om det bakomliggande, något som appliceras i stor utsträckning.

Ahmeds fenomenologiska perspektiv utgår ifrån bland annat Edmund Husserl, Husserl anser att fenomenologi är en lära om riktning mot objekt, om deras upplevda uppenbarhet som ett objekt givet till det medvetna, samt hur riktningar blir till.⁵⁵ Sara Ahmeds förhållande till medvetande skapas genom att definiera medvetenhet som något avsiktligt mot en persons orientering mot objekt. Ahmed utvecklar detta genom att beskriva hur vår riktning mot objekt

⁵⁴ Mattsson, Katarina. ”Diskrimineringens andra ansikte – svenskhet och ”det vita västerländska” Bortom Vi och Dom - Teoretiska reflektioner om makt, integration och strukturell diskriminering” *SOU 2005:41* Publicerad 01 juni 2005 Uppdaterad 02 april 2015 Sida: 148–149.

⁵⁵ Sara Ahmed (2006). *Queer phenomenology. [Elektronisk resurs]: orientations, objects, others*. Durham: Duke University Press. Sida: 35.

och världen är knutna till vår medvetenhet, med detta menar Ahmed att vårt medvetande och vår orientering mot objekt påverkar vår världsbild genom egen plats i samhället. Medvetandet påverkas i sin tur genom personens plats i samhället, hur personen uppfattas av andra och detta påverkar i sin tur hur personen uppfattar världen. Genom till exempel hudfärg eller kulturell tillhörighet. blir en orientering subjektiv och relativ till personen som besitter den, något som vidare i stort sett omöjliggör att en persons medvetande kan utgå ifrån andra personers orienteringar. Uppfattningen av världen påverkas därmed genom personers orientering mot objekt, vad vi orienterar oss emot beror på vad som finns inom räckhåll för vår kropp. Människor har från början en tydlig riktning mot och inställning till objekt som skapas genom egna upplevelser av och riktningar mot samhället.⁵⁶ Ahmed beskriver orienteringen som personers utgångspunkt i världen, där personer riktar sig själva i en specifik riktning gentemot ett objekt. Här beskriver hon även att denna punkt påverkas av hur andra upplever oss, eftersom detta påverkar hur en person upplever objekt.

Ahmed beskriver med hjälp av Husserl hur en uppfattning riskerar att utestänga flera perspektiv, genom ett exempel om Husserls skrivbord. Husserl ser sitt skrivbord som något bekant (familiar), han behöver inte reflektera över att det finns där. Ahmed beskriver även hur hennes egen tolkning av detta skapas av hennes tidigare läsning, inte en förstahandsupplevelse. Här riskerar hon att bara se Husserls skrivbord och kontor utifrån böckernas beskrivning. Bakomliggande arbete för att möjliggöra Husserls skrivbord läggs det inte någon större vikt vid, det bekanta riskerar därmed att utestänga andra tankegångar och reflektioner. Enligt Ahmed kan vi ställa två frågor; Vem sitter med blicken mot skrivbordet? Och har skrivbordet ett ansikte och pekar ansiktet mot vissa kroppar mer än andra? Utifrån Ahmeds två frågor kan följande fråga ställas; Hur kan en persons egen riktning mot objekt påverka vilka som får ta plats i ett rum och vem har möjligheten att påverka hur personer ska identifieras? Vår egen blick mot skrivbordet påverkas av vilka vi är, något som Husserl kallar en naturlig inställning mot ett objekt. Husserl beskriver den naturliga inställningen genom att förklara hur vi upplever objekt, precis som vi tidigare beskrivit med hjälp av skrivbordet. Skrivbordet finns alltid tillgängligt, vi behöver därmed inte tänka på om det finns där eller inte eller vilka konsekvenser som kan uppstå om det inte längre finns där. Egna upplevelser av världen skapar alltså det som vi utgår ifrån i vår naturliga inställning, den naturliga inställningen kring skrivbordet gör att vi inte längre tänker vad skrivbordet gör utan att det är ett bestämt objekt som finns till hands.⁵⁷ Genom

⁵⁶Ibid, Sida: 116.

⁵⁷Ibid, Sida: 32–33.

vår upplevelse och uppfattning av världen skapas vår riktning mot objekt som sedan påverkar hur vår naturliga inställning till objekten formas. Som Ahmed beskriver beror detta på vilka rum en person har tillgång till samt hur denna blir uppfattad av samhället, något som antingen medför större möjligheter eller utestänger personer från rum.⁵⁸

Ahmed utgår i sin fenomenologi från vithet, något som passar in i analysen av myndighetens jämställdhet och jämlikhetsarbete i och med dess fokus på kritiska vithetsperspektiv. I den här typen av fenomenologi utgår Ahmed från vithetens orientering. Kroppars medvetande är här centralt: vem har tillgång till möjligheten att agera i vissa situationer? För att förklara kroppars olika tillgång till handlande använder sig Ahmed av Frantz Fanons teorier, där Fanon pekar på hur kroppars handlande påverkas av världen de lever i. Är handlingsutrymmet tillgängligt? Med detta menar Ahmed att handlingen inte är något som beror på kroppars förmåga utan snarare i vilken mån världen är tillgänglig för kroppen.⁵⁹ Ahmed refererar till Fanons syn på historiska rasscheman, något som Fanon menar bestämmer hur kroppar kan röra sig och orientera sig i samhället. Om världen orienterar sig utifrån vithet kan endast vita kroppar känna sig bekväma och därmed ha fri tillgång till handlingskraft utifrån vithetens orientering. Här menar Ahmed att rasifiering blir något socialt och kroppsligt från historiska arv. Ahmed diskuterar vitheten som punkten 'här', hon menar att vitheten går i arv och därmed reproduceras i samhället, som en god boplats. Hon menar att vitheten förblir sig själv genom bekantskapen som uppkommer och därmed fortsätter gå i arv.⁶⁰

För att identifiera hur vitheten upprätthålls använder sig Ahmed av något hon kallar vanevärldar.⁶¹ Hon menar att vi måste undersöka inte bara hur kroppar blir eller inte blir vita men även hur dessa kroppar påverkar skapelsen av rum. Ahmed tar hjälp av Pierre Bourdieu när hon diskuterar vanor som något som kan bli omedvetet eller en andra natur, att beskriva vithet som en vana blir därför att säga att vitheten är det som kroppar gör. Ahmed menar att vanor inte bara är vanemässiga beteenden utan snarare ett kroppsligt och rumsligt arv. Ahmed pekar på vanors osynlighet i vardagen, något som låter sig uppmärksammas endast utifrån. Kroppen följer med i vanornas handlingar den gör varken motstånd eller uppmärksammar agerandet. Genom Maurice Merleau-Ponty förklarar Ahmed att den invanda kroppen inte gör något motstånd eftersom den ligger bakom handlingen. Detta blir relevant för studien genom att

⁵⁸Ibid, Sida 27–28.

⁵⁹Ibid, Sida: 109.

⁶⁰ Ibid, Sida: 111, 129.

⁶¹ Ahmed, Sara. Goldsmiths. "Institutional Habits", (2013). Inga sidhänvisningar tillgängliga.
<https://feministkilljoys.com/2015/02/02/institutional-habits/?amp>.

analysera hur arbetet med jämställdhet och jämlikhet ser ut, samt om informanterna påverkas av egna erfarenheter och hur dessa ser ut.⁶²

3.2 Institutionella vanor och vithet

Ahmed refererar tillbaka till sitt begrepp vane-världar när hon diskuterar institutionella vanor, dessa vanor menar hon har överförts från ett liv till en institution.⁶³ Eftersom den egna kroppen har skapat dessa vanor genom personens uppfattning av världen så förblir dessa något omedvetet, något den egna kroppen själv inte kan se. Hur dessa vanor senare institutionaliseras menar Ahmed påverkar arbetet med jämställd och jämlikhet och kan i sig motverka arbetet med jämställd och jämlikhet. Ahmed lånar argumenten kring kroppens omedvetenhet kring de egna vanorna från Merleau-Ponty som precis som tidigare nämnt menar att en kropp skapar dessa vanor själv. Något blir alltså institutionaliserat genom att vanor överförs och det blir något som brukar göras automatiskt. Institutionerna skapas genom att vi hela tiden reproducerar dem, vanorna utgör ett arbetssätt som i sin tur gör att arbetet riskerar att missa ett eller flera perspektiv.⁶⁴ Ahmed lånar exemplet om hur lutandet mot ett skrivbord enbart blir något medvetet för händerna och armarna som utför handlingen. Ahmed menar att kroppen hamnar bakom utförandet av vanorna, vi behöver inte längre tänka på att vi ska utföra eller att vi utför en handling. Vanorna slutar att vara objekt för vår uppmärksamhet, vi märker alltså inte längre att dessa utförs. Ahmed beskriver även hur dessa vanor skapas genom det som är bekant i världen, när vi är orienterade mot specifika objekt så blir objekt inte längre uppfattade som objekt utan i stället en förlängning av sig själva. Ahmed diskuterar även vithet i relation till institutionalism och dess vanor, hon beskriver här hur vita kroppar i institutioner samlas och skapar en gemensam koherens inom institutioner.⁶⁵ Vithet som en institutionaliserad vana beskriver Ahmed som något en blir van vid, något en inte längre tänker på, det är en del av kroppen därför kan vi inte längre se den. Ahmed menar att vitheten kan beskrivas som en vana genom dess omärkta natur, något som en ser och inte längre ifrågasätter. Institutionalisering innebär alltså ett anammande av vanor som sedan utgör det som vi gör i vardagen samt hur vi ser eller inte ser objekt. Dessa vanor kan sedan utgöra något omedvetet och osynliggöras och

⁶² Ibid.

⁶³ Sara Ahmed (2012). *On being included. [Elektronisk resurs]: racism and diversity in institutional life*. Durham, N.C.: Duke University Press Sida: 23–24.

⁶⁴ Ahmed (2012), Sida: 35.

⁶⁵ Ibid, Sida: 35, 41.

därmed blir de till en förlängning av kroppen. Institutionaliseringen blir automatisk på en kollektiv nivå, en levnadsvana i arbetet.⁶⁶ Att något blir en levnadsvana kan vi beskriva som en naturlig inställning gentemot något, vi anser att det som vi redan kan hamnar på en nivå där vi inte längre behöver tänka på att vi gör något. Arbete kring jämställdhet och jämlikhet anses komma naturligt eftersom personen i sig anser sig ha sunda värderingar som utgör det omedvetna arbetet.⁶⁷

Ahmeds vanebegrepp och institutionella teori utgör en central del av studien, jag kommer att använda mig av vane-begreppet för att analysera hur de intervjuade arbetar med jämställdhet och jämlikhet och huruvida detta påverkar deras arbete på något sätt. Genom Ahmeds institutionalisering kommer vi även analysera om och hur jämställdhet och jämlikhet har institutionaliserats i arbetet och varför.

3.3 Språkliga maktstrukturer

I boken *Critical discourse analysis: the critical study of language* beskriver Norman Fairclough sin syn på diskurs, Fairclough menar att diskurs är ett anammande av ord och uttryck vi väljer när vi skriver eller pratar.⁶⁸ Språk menar Fairclough delas i ett samhälle och får oss att känna gemenskap, något som i sin tur riskerar att utestänga och främmandegöra personer utanför det språkliga sammanhanget. Språk kan enligt Fairclough aldrig vara neutralt och kommer alltid att bära på kulturella markörer, språket innehåller värderingar och attityder. Värderingar och attityder inom språk är något som åter pekar på de hierarkiska strukturerna språk skapar och upprätthåller. Här kan vi börja ana Faircloughs syn på makt och språk.

Katarina Mattsson använder sig av Fairclough i sin analys av hur svenskheten kan förstås som ett maktperspektiv. Mattsson presenterar Fairclough som i sin tur menar att språkanvändning både konstituerar av den sociala världen och något som konstruerar den sociala världen. Med detta menar Mattsson att språkanvändningen står i ett dialektiskt förhållande till olika dimensioner av det sociala livet. Mattsson fortsätter att beskriva centrala delar ur Faircloughs syn på makt och språk som något som är olika grupper dominans över andra grupper i samhället. Detta påverkar och påverkas av dominant sätt att förstå, uppfatta och tala om

⁶⁶ Ahmed (2013)

⁶⁷ Ibid.

⁶⁸ Fairclough, Norman (2013[2010]). *Critical discourse analysis [Elektronisk resurs] the critical study of language*. Second edition. Abingdon: Routledge Sida: 27.

omvärlden i form av diskurser. Mattsson diskuterar att, genom att beskriva och skapa mening om människor och samhället på ett specifikt sätt, så påverkar härskande diskurser hur personer i samhället tänker och handlar. Mattsson beskriver även hur Fairclough också hävdar det omvända; vilka grupper i ett samhälle som besitter makt att formulera härskande diskurser kommer att få direkt påverkan för hur dessa formuleras och utformas.⁶⁹

En diskurs som utgår ifrån svenskheten som norm kan i och med denna utgångspunkt påverkas av att dessa formuleras av personer som identifierar sig som svenskar eller med 'det svenska'. Mattsson beskriver att detta även kommer att ge effekter på hur människor i sitt konkreta vardagsliv ser på sig själva och andra och hur de agerar i samhället. En diskurs med en svensk utgångspunkt medför därmed ideologiska, materiella och symboliska effekter av en 'svensk' diskurs. Effekterna bidrar sedan till att producera och reproducera och transformera dominansrelationer Mattsson refererar här vidare till Fairclough. Genom en hänvisning till Antonio Gramscis hegemonibegrepp beskriver Mattsson hur Fairclough klassar hegemoni som en ideologisk kontroll genom samtycke, i stället för tvångsförtryck av underordnade klasser i samhället. Detta diskuterar Mattsson vidare genom att beskriva hur detta innebär att ideologiska sätt att se på världen är något som förstärker hur ojämna maktrelationer riskerar att 'naturaliseras', något som innebär att dessa maktförhållanden ges status som 'sunt förnuft', en allmänt känd och icke ifrågasatt kunskap. Som resultat av detta beskriver Mattsson att diskurserna blir svårfattliga och oantastliga, eftersom det råder konsensus om sanningshalt. Här får diskursiva konstruktioner av olika identiteter och subjekspositioner en betydelse för hur människor uppfattas och bemöts.⁷⁰

Föreliggande studie använder sig av teorin om språkliga maktstrukturer med hjälp av Faircloughs argument kring kollektiv gemenskap som skapas och upprätthålls genom språk. Studien använder sig även av hur en diskurs och språk påverkar maktpositioner samt hur dessa påverkar hur personer relaterar till andra inom den kollektiva gemenskapen eller utanför den kollektiva gemenskapen när det kommer till stötande ord inprogrammerade i chatboten.

⁶⁹ Mattsson (2015), Sida: 148–149.

⁷⁰Ibid, Sida: 148-149.

3.4 Självmodellens teori om subjektivitet

Självmodellens-teorin är baserad i det filosofiska fältet kring medvetenhet och undersöker vad det innebär att vara jag, teorin vill även svara på vad det betyder att det mentala är subjektivt. Tomas Metzinger beskriver självmodellens ontologiska argument och utgångspunkt som följer: jaget har i sig själv inte någon substans, jaget har heller inte möjligheten att existera i sig själv. Jaget kan alltså inte existera som en separat entitet utan är alltid beroende av andra utomstående faktorer. Självmodellens-teorin menar i stället att 'upplevelsen' av jaget existerar som en individuell entitet. Jaget utgör i sin tur inte en egen substans i det metafysiska, det som existerar är istället en upplevelse av jaget, tillsammans med det konstant föränderliga innehållet i självmedvetenheten. Den separata självmedvetenheten kallas enligt Metzinger, fenomen-jaget, det vill säga hur en person upplever sig själv och hur personer framstår för sig själva.⁷¹

Om fenomen-jaget existerar är vi bundna till ett första persons perspektiv, jaget upplever här sin medvetenhet som det egna. Här ägs perspektivet av jaget, teorin syftar till att förklara hur en person upplever sig själv och hur detta påverkar hur en person agerar utifrån sina egenägda upplevelser eller perspektiv. Personen upplever sig riktad mot perceptuella objekt och andra människor. Processen av upplevelsen av dessa objekt och människor och det subjektivt mentala möjliggör ett eget inre perspektiv. Givet jagets upplevelser och riktningar mot objekt menar Metzinger att strukturella funktioner av vår inre upplevelse påverkar hur vi använder erfarenheter och begrepp. Vi kan dra paralleller mellan självmodellens-teorin och Ahmeds teori genom att vi aldrig kan frånga vår egen erfarenhet och hur vi upplever objekt och andra människor genom vår subjektiva upplevelse av samhället.

Självmodellen som teorin kallas på svenska är även något som Yarden Katz använder sig av för att förklara hur en AI skapats och fått en till största del vit ideologi. Katz poängterar att en enskild och uteslutande programmeringshistorik har gjort att endast en sida har fått chansen att presentera sina egna jag-positioner.⁷² Katz beskriver i sin introduktion hur erfarenheter har blivit infiltrerade i artificiell intelligens genom ideologier kring vit överlägsenhet och historisk samt nuvarande kolonialism. Katz lånar argument från Toni Morrison som frågar sig, inte hur saker blir universella, men hur saker maskerar sig som universella, något som Morrison menar är starkt sammanbundet med vithet. Med detta menar Morrison att belysa hur vitheten

⁷¹ Metzinger Tomas. "Empirical perspectives from the self-model theory of subjectivity: a summary with examples." *Prog Brain Res.* 2008; 168:215–45. doi: 10.1016/S0079-6123(07) 68018-2. PMID: 18166398. Sida: 218

⁷² Katz (2020), Sida: 156.

genomsyrar politik, ekonomiska frågor och i sin tur även utvecklingen av teknologi.⁷³ Katz applicerar Morrisons sätt att läsa texter på utvecklingen av teknologi. Katz vill genom Morrisons metod ta reda på varifrån den vita majoriteten inom AI och teknologi kommer ifrån, något som Katz menar kan göras genom att se till självmodellen. Katz menar att en reflektion av jaget återskapas inom AI. Med detta menar han att ett fåtal ideal och ideologier ges plats inom utvecklingen av teknologin, som i sin tur bestäms av vilka som programmerar och har tillgång till skapandet av i det här fallet AI.⁷⁴ Liknande utgångspunkter finns hos Katz och Ahmed, där båda utgår ifrån en bild av jaget som sedan påverkar hur vi upplever världen samt hur vi relaterar till andra, något som blir relevant när vi senare undersöker huruvida arbetssättet riskerar att överföra en rasifierad konstruktion i chatboten.

Teorin kommer i föreliggande studie användas tillsammans med Sara Ahmeds fenomenologi för att analysera hur personers orientering och uppfattningar överförs till chatboten. Självmodellen och Yarden Katz argument ligger till grund för hur analysen av myndighetens chatbot rasifieras som vit och hur detta går till, genom att reflektera över Katz argument att självmodellen återskapas i AI, samt hur teamets vita förhållningssätt påverkar chatbotens rasifiering. Reflektioner över hur en vit subjektiv utgångspunkt påverkar chatboten möjliggör för en analys av vithet inom myndighetens chatbot.

4. Material och metod

4.1 Material och studieobjekt

Studien fokuserar på en statlig myndighet och dess arbete med jämställdhet och jämlikhet inom den del av organisationen som ansvarar för att utveckla myndighetens AI. AI:n i studien representeras av myndighetens digitala chatbot, som används som en avlastning för deras manuella kundtjänst. Objektet består av ett team på sex personer som arbetar på myndighetens AI-avdelning, teamet består av enbart vita personer med en relativ homogen bakgrund i svenska sammanhang. Studien fokuserar på att analysera hur myndighetens AI-arbete speglas i dess digitala chatbot utifrån perspektiven språk, tillgänglighet, neutralitet och rasifiering.

⁷³Ibid, Sida: 96, 154.

⁷⁴Ibid, Sida: 154-155.

Materialet för studien utgörs av elva transkriberade intervjuer som genomförts med 6 olika medlemmar i chatbot-teamet på den aktuella myndigheten. Två intervjuer per person har utförts med undantag för en informant. Dessa intervjuer utgör huvuddelen av materialet, i materialet har även ingått ett urval av testmeningar som teamet skickat via mejl, efter ett samtal kring dessa. I dessa ingår även en lista med ord som valts att programmeras in under följande intent: elaka uttryck, grövre elaka uttryck, rasism, svordomar och grövre svordomar. I studien har ytterligare utgått ifrån textbaserade dokument som beskriver arbetet med skapande av chatboten. Även en PowerPoint som beskriver användningsområden samt mål presenteras som material. Materialet har avgränsats till det som är relevant för frågeställningarna.

4.1.1 Bakgrund om studieobjekt

I föreliggande studie beskriver medlemmarna i det statliga AI-teamet hur arbetet med chatboten går till innan chatboten publiceras och interagerar med offentligheten. Följande stycke beskriver hur teamet arbetar med sin chatbot innan den publiceras. Programmeringen av chatboten är intent baserad, det vill säga att basen av chatboten är baserad på det som kunden frågar efter. Processen är presenterad av informanterna genom skrivna dokument som givits till uppsatsförfattaren samt kortare svar under intervjuerna. Processen som sker beskrivs utifrån följande steg:

- 1 Steg nummer ett innebär att ställa upp de krav som kunden har sedan radas dessa upp för att få en översikt.
- 2 Därefter skrivs test-och träningsmeningar, här skriver AI tränarna frågeställningar utifrån hur en kund ställer frågan som chatboten ska förstå. Det krävs ca 20–60 träningsmeningar och ca 10 testmeningar för varje intention.
- 3 Efter test-och träningsmeningarna skrivs svar av AI-tränarna, som anpassas efter de frågeställningarna som kunden har uttryckt.
- 4 När dessa svar är skrivna skickas dessa till en rättsavdelning för att bli kontrollerade.
- 5 Sista steget är att publicera och göra intentionerna tillgängliga för kunden.

Detta arbete är något kontinuerligt, alltså inte något som slutar med de steg vi nyss såg. I informanternas svar framgår det att AI-tränare ständigt måste förbättra och förfina intentionens träffsäkerhet, detta görs genom att varje dag analysera data från chatboten. Informanterna beskriver även deras arbetssätt med den färdiga chatboten som något baserat på efterfrågan, det

som det skrivs mycket av blir tilldelade egna intent. Vad som efterfrågas identifieras genom det teamet kallar för konversationsanalyser, där ca 10% av konversationerna med chatboten analyseras. Myndigheten skriver i sin plan för jämställdhetsintegrering hur arbetet med jämställdhet utifrån kön läggs en stor vikt vid, där de arbetar utifrån tre punkter: ett, myndighetens verksamhet ska vara tillgänglig för alla oavsett kön, två myndigheten ska använda de verktyg de har för att uppnå de jämställdhetspolitiska målen och tre, myndigheten ska ge lika behandling vid anställning oavsett kön. I myndighetens jämställdhetsintegrering tydliggörs dessa tre punkter och myndigheten förtydligar vidare att jämlikhet det vill säga arbete med lika rätt och möjlighet oavsett etnicitet eller nationalitet i bredaste mån möjligt ska inkluderas i jämställdhetsarbetet.⁷⁵ Jämställdheten är definierad som något som behandlar lika rättigheter mellan kön, och att jämlikheten behandlar lika rättigheter oavsett etnicitet, nationalitet eller hudfärg. Jämställdhets och jämlikhetsarbete ska utföras tillsammans i den bredaste mån möjligt och inkluderas i alla uppgifter och uppdrag inom myndigheten.

4.2 Semistrukturerade intervjuer

Studien är situerad inom det kvalitativa forskningsfältet och undersökningen utgår ifrån en semistrukturerad intervjumetod, en metod som öppnar för en reflexiv och reflekterande miljö. Metoden öppnar för en mer informell konversation utifrån en intervjumall med frågor, frågor som under intervjuens gång kan ta andra former. Detta ger forskaren flexibiliteten att utesluta eller addera nya perspektiv beroende på vilken riktning intervjun tar. Genom flexibiliteten möjliggör intervjumetoden för uppmuntrande av vidare svar och förklaringar, metoden tillåter även informanterna själva rikta intervjun in på nya områden som de anser relevanta. Svaren i en semistrukturerad intervju är öppna och inte kopplade till forskarens potentiella förutbestämda ideer. Detta gör att metoden lämpar sig väl för studien. Valet av semistrukturerad intervjumetod gjordes tack vare dess flexibilitet och öppenhet för konversationer med informanterna.⁷⁶

⁷⁵ Anonymiserad källa

⁷⁶ Davies, Charlotte Aull, (2008) *Reflexive ethnography: a guide to researching selves and others*, 2., [rev. and updated] ed., Routledge, London Sida: 105–106.

4.2.1 Förarbete

Letandet efter informanter började i slutet på januari 2022, jag valde då att mejla så många företag som möjligt som arbetar med AI eller använder sig av chatbotar som en virtuell assistent, för att få största möjliga chansen till en träff. I början presenterades inte några större krav förutom att det fanns ett aktivt arbete eller intresse av chatbotar, både förkroppsligade samt textbaserade chatbotar. Studien var från början avsedd att omfatta flera företag, för att få en större generell uppfattning om hur jämställdhet och jämlikhet appliceras på teknologi. Allteftersom studien utvecklades skapades ett större intresse för vithet, jämlikhet och jämställdhet inom fältet kring AI. Studien ändrade efter några veckor inriktning, för att möjliggöra en mer specifik analys av ett specifikt team och deras arbetssätt på en statlig myndighet. Alla inblandade som i dagsläget inte är inkluderade blev informerade i god tid och gav positiva svar till möjliga framtida medverkanden. Innan intervjuerna tog plats introducerades forskare och informanter i form av ett Skype-möte där informella introduktioner gjordes.

Intervjuerna skedde på distans via mötestjänsten zoom där en inbjudningslänk skickats ut. Intervjuerna började med att informera om informanternas rätt till anonymitet och rättighet att inte svara på vissa frågor. Informanten blev även som tidigare informerad om att samtalet/intervjun skulle bli aktivt inspelad och hållas privat.

4.2.2 Genomförande och tematisk analys

Studien har genomförts genom 11 semistrukturerade intervjuer, via Zoom. Varje intervju har haft sin egen inriktning och som alla utgått ifrån samma intervjumall, både baserat på vad informanterna gör i teamet och deras erfarenheter. Detta har medfört en blandning av svar. I intervjuerna användes en lista med bestämda frågor som forskaren sedan utgått ifrån för att starta en konversation. Intervjuerna pågick generellt i ca 90 minuter, där även konversationer orelaterade till intervjun togs upp för att skapa en trevlig samtalsmiljö. För att underlätta så delades intervjutillfället upp på två sessioner, båda intervjuer utförda med samma metod, den semistrukturerade intervjun. Intervjuerna gjordes alla individuellt, varje medlem i teamet fick svara på liknande frågor baserade på samma mall. Informanterna fick sedan delvis styra i vilken riktning intervjun skulle gå, detta genom att presentera egna exempel och erfarenheter. Intervjufrågorna baserades på studier som presenterar tydliga fenomen inom fältet, varav några

kan återfinnas i forskningsöversikten. O'Reilly nämner att etnografiska forskare ofta har några strukturerade element i sin semistrukturerade intervjustudie. I min studie har några sådana delar uppstått där jag frågat efter konkreta exempel samt frågat om skriftligt material som senare potentiellt analyseras.⁷⁷

Utförandet av tematisering har gjorts med hjälp av tydliga fokus-teman som utgångspunkt för studien. Mer specifika teman har identifierats genom genomgång av transkriberingar och markering av relevanta delar i form av citat. Klassificeringen av relevanta delar utgår ifrån de frågeställningar som studien syftar till att undersöka. Dessa citat har sedan format en uppfattning om vad som förekommer frekvent. Kodningen har gjorts genom att leta efter gemensamma uttryck i transkriberingarna, något som sedan format teman för analys. För att analysera de identifierade teman i materialet har tematisk analysmetod använts. Metoden har gjort det möjligt att identifiera teman samt underlättat för jämförande och att hitta liknande svar hos informanterna. Den tematiska analysen av materialet har utgått ifrån aspekter som intervjuerna baserats på, det vill säga, arbete med jämställdhet och jämlikhet samt språk. Utifrån dessa två grundstenar har sedan mönster identifierats genom kodningen av materialet. Resultatet av kodningen har lett mig till teman relevanta för studien samt objekt för vidare diskussion. Dessa objekt och mönster har sedan analyserats utifrån ett passande teoretiskt ramverk. Syftet med metoden i föreliggande studie har varit att identifiera och analysera hur medlemmar i ett team beskriver sitt arbete med jämställdhet och jämlikhet och språk för att sedan genom teoretiska perspektiv analysera möjliga konsekvenser av arbetssättet som framkommer.⁷⁸ Transkribering har gjorts på alla intervjuer för att tillgodose så mycket information som möjligt för analysen. Transkriberingen har genomförts löpande under studiens gång och har i vissa fall lett till följdfrågor ställda via mejl.

5. Etik

Min egen position som forskare med väldigt lite erfarenhet av att arbeta med teknologi, dvs AI, har påverkat intervjun på så sätt som att informanterna utförligt fått förklara processen av att bygga AI samt hur de anpassar sitt arbete till jämställdhet och jämlikhet. Min position som forskare kan i intervjuerna haft en påverkan på svaren från informanterna genom att ta upp

⁷⁷ O'Reilly, K. (2012). *Ethnographic methods*. (Second edition). London: Routledge. Sida: 120.

⁷⁸ Bryman, Alan. (2018). *Samhällsvetenskapliga metoder*. Upplaga 3. Stockholm: Liber Sida:703

exempel specifikt relaterade till chatboten. Frågor om värdeladdade och rasistiska ord kan ha påverkat svaren då informanterna kan inta en försvarsposition. Min position som forskare kan även ha påverkat svaren genom att använda ord som informanten potentiellt tolkar fel, förgivttagande att det skulle vara tydligt i formuleringen i frågan.

Informanterna och deras personliga kännetecken och uppgifter i studien har anonymiserats enligt överenskommelse och forskningsetiska riktlinjer.⁷⁹ Beslutet att inte benämna myndigheten eller chatboten togs gemensamt, tillsammans med ansvarig för teamet, vidare togs beslutet att inte beskriva och analysera utseendet på chatboten, detta för att inte identifiera myndigheten i studien. I inbjudan till samarbetet presenteras den grundläggande etiken om att förbli anonym, egen kontroll över information som överlämnas, att data och information efter studien kommer raderas och /eller förstöras för att upprätthålla anonymiteten samt fri rätt att dra sig ur samarbetet när och om informanterna inte längre skulle känna sig bekväma i undersökningen. Informanterna har blivit informerade om studiens syfte för informerat samtycke, detta i linje med de riktlinjer som Vetenskapsrådet presenterat som ramverk för att upprätthålla en etisk sund undersökning för alla involverade parter.⁸⁰ Jag kommer även följa riktlinjerna kring uppgifterna jag får enligt Vetenskapsrådet. Här kommer jag använda mig av nyckelorden tillförlitlighet, ärlighet, respekt och ansvar för att leverera en objektiv och ansvarsfull analys av materialet.

Genom studien har frågor kring anonymitet uppkommit, huruvida informanterna i teamet kan identifiera varandra utifrån de svar de givit. Anonymisering av citat har därmed varit en stor fokuspunkt genom uppsatsen och censurering av citat för att i största mån undvika identifierbara personliga och arbetsrelaterade markörer hos enskilda informanter.

Information om när det är för sent att dra sig ur har även presenterats för informanterna. I och med studiens stundvis känsliga natur har informanterna givits möjlighet att avstå från att svara på frågor samt att frågorna anpassats för att vara applicerbara enbart på teamet som helhet och arbetssättet och inte på individnivå. Informanterna var däremot fria att prata utifrån sig själva när det kände att det behövdes. Forskaren som utför studien är en vit kvinna/ickebinär person i 25 års åldern, med liten erfarenhet av arbete med AI som teknologi. Situering inom vithet medför en faktor av igenkännande och en väg till förståelse av informanterna. Situeringen

⁷⁹ Vetenskapsrådet (2017). *God forskningssed [Elektronisk resurs]*. (Reviderad utgåva). Stockholm: Vetenskapsrådet.

⁸⁰Ibid.

medför ett sätt att göra en kritisk analys av vitheten utifrån en vit hållning till samhället och öppnar för att vara kritisk mot egna erfarenheter. En delad, gemensam vit uppfattning och upplevelse av världen både underlättar och försvårar en analys av jämställdhets och jämlikhetsarbete inom AI.

6. Analys.

Undersökningen har haft syftet att undersöka och förstå hur en statlig myndighet arbetar med jämställdhet och jämlikhet i processen av att bygga och utveckla sin chatbot. Analysen är uppdelad i fyra teman, analysens första kapitel består av hur teamet arbetar med jämställdhet och jämlikhet i byggandet av chatboten och huruvida deras arbetssätt institutionaliserar ett eller båda begreppen jämställdhet och jämlikhet. Nästföljande tema undersöker hur teamet använder sig av språk i relation till jämställdhet och jämlikhetsarbete, där vi även diskuterar hur en kollektiv språklig gemenskap riskerar att påverka vad som skrivs till chatboten. Tema tre visar på hur vithet konstrueras inom chatboten genom teamets arbetssätt och förhållning till sin omgivning. Avslutande tema diskuterar kort deras arbete för att uppnå neutralitet. Dessa fyra teman används som analysverktyg utifrån tidigare forskning samt teorierna presenterade i tidigare avsnitt. I avsnittet används begreppen jämställdhet i anslutning till kön, mäns och kvinnors lika rättigheter, jämlikhet fokuserar däremot på alla människors lika rättigheter och rätt till lika möjligheter.

6.1 Arbete med jämställdhet och jämlikhet

Flera av informanterna presenterade arbetet med jämställdhet som något latent eller något som är automatiskt inprogrammerat i deras arbete genom egna värderingar och erfarenheter.

Men det är ingenting alltså, det finns inga mallar utan det bara kommer och vi skriver ju allihop, så det är jag. Det är ingenting som vi tänker på så mycket utan det kommer ju med automatik. Det kommer ju direkt, det är ingenting som man tänker på direkt.

I citatet ovan beskriver informanten hur arbetet med jämställdhet och jämlikhet i teammedlemmarnas vardag inte är något de lägger något större fokus på, utan det anses komma naturligt med i arbetet. Om vi tittar närmare på citatet kan vi se att ordet automatiskt framkommer i första meningen, citatet fortsätter sedan att förklara att arbetet med jämställdhet inte är något de direkt tänker på. Detta är något som problematiseras av Ahmed då hon menar att objekt upplevs på olika sätt, ett av dessa sätt beskriver Ahmed som familjar. Begreppet familjar beskriver Ahmed som något vi är bekanta med, ett objekt vi har en relation till i vår vardag och som personen inte alltid lägger märke till men som fortfarande har en roll i personens liv. Något som anses bekant hamnar därmed enkelt i bakgrunden och bli något som Ahmed fortsätter att beskriva genom Husserls relation till sitt skrivbord.⁸¹ Vi kan utifrån citatet tolka att ett arbete utfört via automatik är något som utgörs av något bekant. Precis som Ahmed beskriver kan detta hamna i bakgrunden, något som inte utgör en större plats i det medvetna. Ahmed beskriver detta genom att förklara hur vi bebor och placerar oss i sammanhang, hon använder sig av Husserls skrivbord för att förklara hur ett objekt kan uppfattas och ses av personen som ett skrivbord men samtidigt hamna i bakgrunden. Skrivbordet förvandlas till något som Husserl möjligen inte ser när skrivbordet används för att utföra arbetsuppgifter. En uppfattning av skrivbordet menar Ahmed kan medföra förutsättningar för hur vi upplever skrivbordet som något i bakgrunden eller något i fokus för vårt medvetande. ”Det är ingenting som jag tänker på, utan det kommer bara. Det är ju så jag själv är uppväxt och det är så jag har lärt mig och det är ju så alltså det.” I citatet från informanten kan vi se ett exempel på hur en uppväxt med en syn på jämställdhet gjort att perspektivet blir något som inte ’tänks aktivt på’. Genom Husserl och hans teori om det bekanta menar Ahmed att perspektiv går obemärkta förbi oss.⁸²

Min tanke och sådant bakom det, så liksom pratar vi väl inte jättemycket om jämställdhet liksom inte dagligen utan det ligger väl lite latent i en hela tiden, kanske man kan säga.

⁸¹ Ahmed (2006), Sida.: 32–33.

⁸²Ibid

Även nästa informant beskriver arbetet med jämställdhet som något latent eller något som alltid finns där men som inte alltid utgör en medveten diskussion inom teamet. Ordet latent blir i det här citatet viktigt för hur arbetet med jämställdhet anses vara något bekant, något som ligger i en person från början. Det latenta kan tolkas som en bakgrund, precis som skrivbordet för Husserl utgör tänket kring jämställdhet en bas för hur arbetet utformas.

Vi tänker, det är ingenting som jag ser som nu, ska vi sätta oss och prata jämställdhet utan jag tycker att det ska omfatta allt, det är liksom det tar inte en paus.

I citat tre presenterar informanten hur jämställdhet inte ses som något som bara ska diskuteras i vissa sammanhang utan något som hela tiden ska vara med i arbetet, det ska omfatta allt. I sista meningen refererar informanten till hur ojämställdhet i samhället inte tar en paus och därmed ska inte heller deras arbete med jämställdhet ta en paus, det ska inkluderas i allt de gör. Flera informanter nämner att myndigheten har jämställdhetsambassadörer. ”Jämställdhetsambassadörer är ansvariga för var sitt område, så att de har ju liksom haft föredrag och övningar för [ambassadörens] område.” Något som blir intressant i relation hur arbetet presenteras av informanterna, en struktur finns men den tar inte en så stor plats i arbetet, något som visas genom citaten ovan där arbetet med jämställdhet framkommer som något ostrukturerat.

Dessa citat visar att arbetet med jämställdhet anses av informanterna ligga latent och som byggs in automatiskt. Men även att en myndighetsstruktur finns för jämställdhet i form av ambassadörer. Vi kan här ställa oss frågan hur perspektiv och arbete med jämställdhet och jämlikhet kommit att utgöra något latent eller automatiskt för teamet, något jag kommer fortsätta att diskutera.

Utifrån Ahmeds teori kring det bekanta, hur kan vi då se att arbetet har blivit något automatiskt? Arbetet presenteras som en del av det som teamet alltid gör, något som kommer automatiskt utan att de behöver tänka på det. För att något ska anses vara bekant menar Ahmed att två entiteter behöver mötas, i vårt fall är detta en person och jämställdhet samt arbete med jämställdhet.⁸³ Här kommer frågan hur arbetet med jämställdhet har blivit något bekant. Vi kan

⁸³Ibid, Sida:135–137.

börja med att konstatera hur två av de tre citaten enbart nämner jämställdhet det vill säga arbete med kön och lika rättigheter som en del av deras arbete. Om en specifik riktning mot objekt i samhället och en rasifiering kan anses påverka hur en förhåller sig till jämställdhet och jämlikhets arbete skulle vi kunna diskutera om en av dem är enklare och upplevs mer. Upplevs det mer kan personen identifiera sig med detta därmed kan detta bli något välbekant och som kan relateras till, kan detta då påverka hur vi ser och arbetar med jämställdhet och jämlikhet?

Ahmed menar att en person uppfattar världen utifrån en egen medvetenhet som ger oss specifika uppfattningar av objekt, en egen utgångspunkt baserad på vem du är. Aspekter som hudfärg, social status och miljö påverkar hur en person blivit positionerad i samhället.⁸⁴ Ahmed tar upp Fanon som menar att en persons hudfärg spelar en stor roll i skapandet av uppfattningar av världen och hur samhället upplevs av personer samt hur detta tillskriver personer en plats i sin omgivning. Ahmed menar här att kroppen spelar en stor roll för hur vi upplevs av andra, något som sedan påverkar hur vi placeras och vilken mån av tillgänglighet till världen omkring oss vi får.⁸⁵ En icke-vit kropp får inte samma förhållningssätt till samhället som en vit kropp till exempel. Ahmed refererar här till Fanon som menar att detta är en konsekvens av det historiska rasschemat. Hur kroppar historiskt blivit klassificerade och uppfattade, på det sättet har de blivit givna olika hållningar till samhället som sedan medför olika upplevelser.⁸⁶

Ahmed menar att vi genom uppfattning av samhället är riktade mot ett objekt, hur vi upplever ett objekt skapar medvetenhetens ”världsliga” dimension. Om medvetandet handlar om hur vi upplever och uppfattar saker omkring oss blir medvetandet något förkroppsligat, känsligt och situerat. Med detta menar Ahmed att vi enbart kan uppfatta saker utifrån vår egen hållning till och riktning mot objekt runtomkring oss. Ahmed refererar till Husserl som beskriver denna riktning mot objekt som något från den naturliga inställningen en person utgår ifrån. Något som skapar det vi tidigare diskuterat som det bekanta och något som anses komma automatiskt utan en vidare tanke på det.⁸⁷ Om vi utgår ifrån ett vitt förhållningssätt till samhället som en av entiteterna i mötet med arbete inom jämlikhet och jämställdhet så kan vi börja att analysera hur detta skulle kunna påverka arbetet med jämställdhet mellan könen som bekant i jämförelse med det jämlika där etnicitet och ursprung innefattas. Om vi tittar närmare på det som Ahmed refererar till Fanon om kroppsschemat kan vi se att en uppfattning av något påverkas av en

⁸⁴Ibid, Sida: 108–109.

⁸⁵ Ahmed (2012), Sida: 109–111.

⁸⁶ Ahmed (2006), Sida: 27.

⁸⁷Ibid, Sida: 111.

kropp och det kroppsligt rasifierade. Genom dessa två olika förhållningssätt vit och icke-vit i samhället och med olika erfarenheter, kan vi diskutera om det genom bristen på mångfaldig erfarenhet läggs mer fokus på arbete med jämställdheten än jämlikheten.

Om vi då tittar närmare på hur en persons medvetande riskerar att rikta sig mer åt jämställdhet, hur en fokusering på jämställdhet ansluter till vita kroppars brist på erfarenheter av rasifiering. Därmed saknas ett icke-vitt förhållningssätt till samhället som påverkar dess möjlighet till handlingsutrymme i världen. Om vi tittar på detta utifrån en svensk myndighet ser vi att myndigheten utgår ifrån jämställdhets begreppet till största del, där myndigheten har tre stora mål inom jämställdhetsintegrering fokuserat specifikt på kön; 1) verksamheten ska vara lika för alla oavsett kön, 2) myndigheten ska använda de verktyg och handlingsutrymme för att bidra till jämställdhetspolitiska mål, 3) alla anställda ska få samma möjligheter oavsett kön.⁸⁸ Vidare beskriver de att inom alla dessa ska i bredaste möjliga mån inkluderas jämlikhetsperspektiv, så vi kan se att även jämlikhet ska vara en del av arbetet. Något som motsäger de tidigare citaten, ” Min tanke och sådant bakom det, så liksom pratar vi väl inte jättemycket om jämställdhet liksom inte dagligen utan det ligger väl lite liksom latent” och ” Vi tänker, det är ingenting som jag ser som nu, ska vi sätta oss och prata jämställdhet utan jag tycker att det ska omfatta allt, det är liksom det tar inte en paus.” två av de tre citatet fokuserade enbart på jämställdhet.

När frågor om kritiska perspektiv på egna förhållningssätt mot objekt och samhället så kom det upp för flera informanter att kritiska diskussioner kring egen utgångspunkt inte framkommer speciellt ofta.

Ja, alltså, jag tror inte vi pratar så där jättemycket om det rent generellt liksom utan jag tänker att vi i och med att alla har jobbat på [myndighetens kundtjänst] inom till exempel där vi har suttit och bemött folk som ringer in som kommer från ja, alla länder som finns i världen liksom kontakta oss. Ibland så försöker man prata på engelska också, utan att vi riktigt kommer någon vart också, så jag tror ändå att under tidens gång så har vi väl lärt oss att tänka liksom ur ett större perspektiv för att hur ska vi kunna hjälpa personer [...]

⁸⁸ Anonymiserad källa.

Citatet beskriver hur tidigare erfarenheter gör det möjligt för dem att tänka ur ett större perspektiv, men att det samtidigt inte förekommer så mycket diskussioner kring detta. Vi kan alltså se att ett sorts jämlikhetstänk även det har blivit internaliserat men inte pratas om och uppmärksammas på samma sätt som med jämställdheten. En annan informant svarade även på frågan om de hade kritiska tankar och diskussioner kring sitt eget förhållningsätt som vit; ”Ja precis nej, inte på det sättet”. Deras egen hållning riskerar att påverka och speglas i produkten de skapar, citatet visar på att en kritisk medveten diskurs kring vithet och vad de innebär för deras riktning mot objekt inte får någon plats i arbetet. Senare i samma citat lyfter informanten arbete med funktionsvariationer, men att de inte har en aktiv kritisk diskussion kring sina egna vita och relativt privilegierade uppfattningar och hållning till och av samhället.

Här kan vi alltså se det som Ahmed menar, att en person riskerar att bortse från objekt och fenomen när den utgår ifrån sin egen subjektiva upplevelse av världen, något som därmed ökar svårigheten för en person att sätta sig in i en annan persons upplevda världsbild.⁸⁹ Informanten fortsätter att berätta hur deras relativt lilla team på sex personer inte är speciellt representativt, de är en liten grupp av relativt lika personer som bestämmer hur allt ska vara. Där blir det då ofrånkomligt att deras värderingar kan påverka texterna i chatboten:

Sex personer är vi ju och jag menar, det är en ganska liten del. Av det stora hela och det är vi som liksom bestämmer allting där och hur det ska vara. Det är väl såklart att. Att det är ofrånkomligt kanske att våra värderingar eller något sådant skulle kunna liksom vara orsaken till texterna där, så det där var säkert varit bra men och nu utomstående person som kunde granska.

Informanten menar här att ett team på sex personer riskerar att överföra sina egna värderingar och påverka texterna. Citatet visar en koppling till självmodellen, där en persons egna subjektiva upplevelser av världen, fenomen och objekt blir det som tar plats. Vi återkommer till hur citatet kan anslutas till hur en vit chatbot kan skapas genom självmodellen genom en överföring av värderingar senare i kapitel Vithet i Artificiell Intelligens.

Som Ahmed menar kan finns det risker att vår hållning till världen och objekt utelämnar flera perspektiv och objekt som inte finns tillgängliga inom en persons medvetande och andra

⁸⁹ Ahmed (2006), Sida: 111.

uppfattningar av samma objekt vilket skapar en syn på samhället och ger oss en uppfattning av vad som existerar i störst utsträckning.⁹⁰ Vitheten har inte en historia av att bli rasifierad, vad vi däremot kan se är vithetens historiska relation till rasism och diskriminering, vi kan diskutera detta vidare genom att se till vilka perspektiv som blivit synliggjorda och uppmärksammade av samhället. Där flera förhållningssätt gentemot samhället utestängs och får därmed ingen chans att höras ergo, försvårar och ibland omöjliggör för personer att synas på en större plattform som artificiell intelligens.

Vi kan alltså se att detta har blivit något bekant något som en automatiskt tänker på genom en egen upplevelse och ett möte mellan arbete och egna upplevelser. Vi kan diskutera detta vidare genom den naturliga inställningen som en utgångspunkt för en persons upplevda värld. En naturlig inställning och en orientering påverkar alltså enligt Ahmed vad som blir bekant för en person. Något som skulle kunna förklara den fokusering på jämställdhet i citaten och hur det har blivit naturligt, något som sker automatiskt utan att behöva tänka på det. Något som genom förhållningssätt och en uppfattning till/av världen ger oss tillgång till olika uppfattningar av objekt som sedan påverkar hur dessa formas, vilket framkommer i arbetet med jämställdhet i citaten.

Ändå utgår vi från att vi inte ska göra någon skillnad på jämställdhet och tänker på man och kvinna, alltså kön där ju utan att det ska vara lika oavsett där, men vi arbetar väl inte jättemycket liksom kring det när vi skriver våra texter till exempel utan vi tänker inte så här, hur gör vi detta jämställt eller?

Det finns en aktiv utgångspunkt hos de flesta av informanterna kring jämställdhet, allas lika värde och att inte göra skillnad på män eller kvinna. Flera av informanterna anser och beskriver att de aktivt arbetar med frågor kring funktionsvariationer när det kommer till jämlikhet.

Det finns ett slags tillgänglighetsforum här på [myndigheten] där folk med som precis som du säger har ja, men har kanske lite utmaningar. Blindhet eller ser lite

⁹⁰ Ibid, Sida: 32–33.

sämre eller vad det nu kan vara och de fick testa [chatboten] i några veckor och [...] så fick jag sen lite feedback på vad de [tyckte]. Så sådana grejer är ju alltid bra att ha tycker jag är viktigt för det finns ju väldigt många sådana ute i samhället. Det är klart vi ska vara tillgängliga mot dem också.

I det här sammanhanget så inkommer flera faktorer; att tillgänglighetsfrågor kring funktionsvariation är lagstiftad och måste utgöra en aktiv del i arbetet. Citatet presenterar tillgängligheten för personer med blindhet eller personer som ”kanske har lite utmaningar”, något som utgör en stor del av deras jämlikhetsarbete. Till skillnad från jämlikhet fokuserad kring etnicitet och bakgrund ser vi att detta prioriteras.

Citatet nedan presenterar hur jämställdhets och jämlikhetsarbetet låg i dem redan innan de började arbeta på myndigheten och hur arbetet påverkas av att de jobbar på en myndighet. ”Jag personligen kanske inte känner att det beror på mig att jag är anställd på [myndigheten] egentligen kanske utan det är väl också vad man har för värderingar och tankar, liksom personligt och så.” Informanten menar här att hen inte anser sig enbart påverkad av myndighetens syn på jämställdhet och jämlikhet utan att det även är något som hen har tagit med sig från sina egna erfarenheter och värderingar. Något som blir intressant att se på genom den naturliga inställningen. Precis som Ahmed menar måste två entiteter mötas för att skapa en uppfattning av ett objekt. Om vi igen går tillbaka till det som vi tidigare diskuterade kring jämställdhet som något som verkar ta störst plats kan vi se att detta kan påverkas av hur ett objekt skapas och medför specifika upplevelser för olika orienteringar. Att en vit uppfattning av samhället genom sin hållning gentemot objekt i samhället endast har möjlighet att se till den egna uppfattningen, riskerar att överföras till arbetet på myndigheten.

Enligt Ahmed blir det automatiska eller det latent så som informanterna presenterar det till något som hamnar i bakgrunden och inte längre ifrågasätts på samma sätt eftersom det är något som vi alltid har med oss och har tillgång till. Något som i sig medför risker, riskerna ökas i takt med att en person tar det automatiska eller latent för givet i högre och högre grader. Genom detta och citaten kan vi identifiera risker med vad som runt jämställdhet kan gås miste om, om arbetet blivit automatiskt och något som alltid pågår. Kan vi då se hur vissa saker inte är möjliga att se från en viss orientering? Om vi går närmare in på detta och hur en riktning mot objekt kan påverka hur ett arbete ser ut så kan vi se till nästa citat:

Vi har inte tagit så där jättemycket höjd med Mohammed eller vad det nu kan vara för andra [namn] som man borde ha med. Vi har nog några exempel, men det är inte så många utan där har vi nog omedvetet, liksom utgått från svenska namn nu när man tänker efter.

Citatet beskriver hur arbetet med namn i chatboten har utgått ifrån svenska namn, något citat pekar på gjorts omedvetet. Ingen större höjd har tagits för namn som Mohammed, citatet beskriver att några av dessa finns med, men inte i någon större mängd. Citatet knyter an till det vi tidigare beskrivit om familjaritet och en naturlig inställning. Det omedvetna kan vi se representera den naturliga inställningen, det som anses komma automatiskt och som det finns mest av. De svenska namnen är det välkända något som används omedvetet, något som vi kan relatera tillbaka till Ahmeds exempel om skrivbordet. När vi använder skrivbordet så kan detta försvinna från vår medvetenhet och vi fokuserar istället på det vi gör medvetet på bordet. Namnen blir här vårt skrivbord, något som används för att skapa en sorts inkludering men som under användningen glöms bort. Här kan vi se att en omedveten riktning mot det bekanta tas, det välkända i det här fallet representeras av det som kallas svenska namn i citatet. Dessa bekanta objekt och förhållningsätt skapas av en riktning mot objekt med hjälp av ett eget förhållningsätt och en förkroppsligad medvetenhet mot objektet. Hur detta händer kan vi återigen förklara via Ahmed, som menar att bakom det bekanta så hamnar andra saker i bakgrunden. Något som riskerar att hända med jämlikheten och jämställdheten blir som med Husserls skrivbord: det andra stängs ute från det direkta medvetandet. Ahmed beskriver bekanta objekt som något som är tillgängligt för en person, genom uppfattningar av objekt eller utgångspunkt.⁹¹ Skrivbordet blir då det som är synligt för en viss upplevelse och uppfattning vilket gör att det andra runtomkring i rummet blir suddigt. Det som blir suddigt i det här fallet blir de ickesvenska namnen, något som teamet jobbar aktivt med men något som inte uppmärksammas och därmed blir funktionen av objektet utestängt. Det välkända tar den stora platsen i rummet, något som sker utan en tanke på det. En vit svensk hållning till samhället ger personer vissa erfarenheter och upplevelser av världen vilket därmed riskerar medföra att vi enbart ser en viss typ av namn till exempel och resten blir utestängda. Något som sker utan en medveten tanke, eftersom vi enbart kan utgå ifrån vår egen uppfattning av samhället. Fokuset

⁹¹Ibid, Sida:129.

på jämställdhet i vårt fall kan diskuteras i liknande termer. Jämställdheten blir objektet som syns, det runtomkring utestängs. Något vi även kan se genom det omedvetna användandet av nästan uteslutande svenska namn.

I arbetet finns ett aktivt fokus på fysiska och funktionsmässiga svårigheter, men en kritisk syn på en egen hållning mot objekten jämställdhet och jämlikhet är något som inte får samma utrymme i arbetet. En av informanterna tar upp att vita män i större utsträckning har större tillgång till att arbeta med teknik något som kan påverka innehållet. Informanten lyfter vikten av att vara kritisk mot sin egen uppfattning och hållning till samhället.

Nej, jag tycker det centrala, i alla fall för mig att jämställdhet är ett konstant perspektiv. Man ska ha med sig inte tillfälligt, bara liksom utan man måste alltid tänka på vad man skriver när jag skriver ett svar. Jag måste alltid tänka på att när jag skriver träningsmeningar eller något liknande så kan jag inte bara skriva ut ifrån mig då som vit heterosexuell man, [liksom jag].

I flera av de andra informanternas svar om kritiska syner på sin egen hållning till samhället och objekt framgår det att kritiska syner på egen hållning inte förekommer i någon större skala. Till skillnad från arbetet med till exempel jämställdhetsforumet där personer med funktionsvariationer får en möjlighet att testa produkten och ge konstruktiv kritik för att kunna förbättra. En av informanterna svarar så här när frågan om kritiska synen på sin egen hållning och uppfattning av samhället är en aktiv del av deras arbete: ”Kanske går det jobba på, för det är precis som du säger det har gått lite under radarn. Vi har bara tutat och kört och kanske inte tänkt efter ur det perspektivet.” Här kan vi alltså se att det uppkommer en klyvning, en person säger att de alltid måste vara medvetna men en annan person ur teamet menar att det inte är något de har tänkt på.

Något som vi kan diskutera ur ett institutionellt perspektiv, är om inställningen till jämställdhets och jämlikhetsarbetet påverkat och reglerat ett kritiskt perspektiv? Ahmed beskriver ett liknande scenario när hon läser en bok där författarna beskriver ett rum, Ahmeds uppfattning av rummet är helt beroende av hennes egen bakgrund och hur hon själv förhåller sig till beskrivningen. Uppfattningen blir därigenom subjektiv och riskerar att utestänga andra viktiga

objekt.⁹² Objekt som i sig hade kunnat utvidga hennes uppfattning av rummet, och att liknande riskerar att hända där synen på jämställdhets arbete kommer från en begränsad bakgrund. Att det går lite under radarn visar på att en kritisk diskussion inte tar plats, alltså kan vi se att en subjektiv medvetenhet skapas. Detta riskerar i sin tur att överföras genom självmodellen som Katz menar har under historien lett till att AI förblivit vit.⁹³

Vi kan vidare se fokuset på jämställdhet även där, genom användandet av namn som ses i citatet nedan; ”Det lokala inom Sverige eller och just kanske med det manliga och kvinnliga där [...] nej, utan det blir nog bara namnen och könen i det här fallet.” Citatet beskriver hur de arbetar med namn och kön till största del i programmeringen av chatboten. Frågan som ställdes hade för avsikt att se om de hade en strategi för att få in en större inkludering av olika etniciteter i arbetet med jämställdhet och jämlikhet. Som vi ser i citatet så läggs det största fokuset på just kön och namn. Namn som vi i det förra citatet såg utgå omedvetet mest ifrån det svenska, något som blir tydligt och bekant är en fokusering på det manliga och kvinnliga, arbetet utgår här ifrån det som är bekant. Något som kan relateras tillbaka till vithetens brist på rasifiering som påverkar hur världen uppfattas, samt vad en kan se. Eftersom en vit person sällan blir rasifierad är detta inte något bekant för personen, tankar om rasistiska uttalanden riktade mot en kommer inte automatiskt och därmed kan inte heller anses vara bekant. Vithetens brist på rasifiering kan därmed ses som en förklaring till varför en viss typ av jämställdhet verkar ta störst plats inom arbetet.⁹⁴ Genom att en vit person genom sin orientering inte kan uppfatta eller uppleva en rasistisk värld såsom en icke-vit upplevelse av världen så påverkar detta hur dessa möter den andra entiteten som skapar orienteringen mot objektet. ”Ja, men nästan lite exkluderande och ganska omedvetet. Visst skulle vi kunna ha några personer som har utländska namn.” Här kan vi se att en vit svenskhet tar stor plats i arbetet även i relation till namn och inkludering, informanten nämner själv exkludering och hur namn omedvetet utelämnas ifrån jämlikhets- och jämställdhetsarbetet.

Genom ett vitt och svenskt förhållningssätt mot samhället, synliggörs vissa fenomen mer vilket kan påverka vad som arbetas mest med och vad som anses vara inkluderat i arbetet med jämställdhet och jämlikhet. En naturlig inställning och en oundviklig subjektiv uppfattning av fenomen hos en person kan påverka ett arbete med vad som kan identifieras genom egen erfarenhet och inte. Detta påverkar även det medvetna och omedvetna, vad som anses mest

⁹²Ibid, Sida:29–30.

⁹³ Katz (2020), Sida: 153–156.

⁹⁴Ibid, Sida: 153–156.

bekant tar störst plats och detta riskerar att ta större plats än det andra som inte går att se genom en annan persons orientering mot samhället. En naturlig inställning och medföljande riktad medvetenhet mot samhället riskerar att påverka ett arbete genom att enbart fokusera på det som anses bekant. Myndighetens riktlinjer kring jämställdhet kan även de ses som anledningen till att det största fokuset inom arbetet kan anses läggas på jämställdhet mellan kön, samt att jämlikhet ska i bredaste mån inkluderas i detta. Att jämlikheten inte har en egen tydlig struktur inom riktlinjerna utgör en risk, då en naturlig inställning till jämlikhet kan anses läggas på jämställdheten mellan könen. Om vi ser till det som Ahmed menar med att en person bara utgår ifrån sin egen riktning mot objekt kan vi se att hållningen mot objekt och samhälle blir en utgångspunkt, något som inte kritiseras på samma vis som andra utomstående faktorer kanske skulle göras. Citaten visar på att tankar kring jämställdhet finns i många aspekter, men att det samtidigt saknas en viktig aspekt i en egen kritisk analys av hållning mot samhälle och objekt samt hur denna hållning riskerar att påverka arbetet och resultaten.

6.2 Språk

I föreliggande undersökning var språk en stor del av utgångspunkten när det kom till arbetet med jämställdhet och jämlikhet. Forskare såsom Fairclough menar att språk är kulturbärande något som är sammanbundet med den skapande processen av traditioner och kulturella mönster.⁹⁵ Hur kan användning av språk påverka hur jämställdhets- och jämlikhetsarbetet utförs? Informanterna beskriver hur de utgår ifrån lagstiftning kring språk när de börjar att programmera chatboten. Lagstiftningen innebär att myndigheter och offentliga organ ska använda sig av det svenska språket. Lagen innefattar även att språket ska vara vårdat, begripligt och enkelt.⁹⁶ Det svenska språket utgör grunden för hur chatboten programmeras och hur chatboten är konstruerad för interaktion med allmänheten. Chatboten kan i dagsläget enbart svara på tilltal på svenska men kan förstå enstaka ord på norska, danska och engelska utöver det svenska.

Informanternas svar indikerar att arbetet med språk utgörs av en bas av det svenska språket som myndighetsspråk. Arbetet med språk centreras även kring efterfrågan, det vill säga det som skrivs in till chatboten via interaktioner med allmänheten har en stor påverkan på vad som

⁹⁵Ahmed (2006), Sida: 32–33.

⁹⁶ SFS nr: 2009–600. Språklag. <https://rkrattsbaser.gov.se/sfst?bet=2009:600>

programmeras in som nya intent [intentioner]. Informanterna beskriver hur arbetet med språk är föränderligt, förutom lagstiftningen och det som är nödvändigt för att få en fungerade chatbot på svenska. De beskriver hur de tränar chatboten i att förstå felstavningar och att dessa sedan ska kunna ge samma svar som ett ord stavat enligt korrekt struktur.

Något centralt i teamets arbete med språk är också det som i myndighetssammanhang kallas klarspråk, ett förenklat språk som ska vara enkelt att förstå för alla oavsett vem du är. Informanterna lyfter problem såsom formerna på språken och hur dessa skiljer sig åt till stor del. Su Lin Blodgett, Solon Barocas, Hal Daumé III och Hanna Wallach diskuterar i artikeln "Language (Technology) is Power: A Critical Survey of "Bias" in NLP" språkliga brister i arbetet med AI.⁹⁷ Här diskuterar de språkliga hierarkier, markerade och omarkerade språk, vidare menar de att språk konstruerar sociala verkligheter samt vilka språk som anses 'lämpliga' att använda i offentligt bruk. De menar alltså att språk och dess ideologier upprätthåller sociala hierarkier, genom ett partiskt bruk av vissa 'lämpliga' språk.⁹⁸ I svensk lag har detta blivit svenska, här blir även det något som riskerar att upprätthålla och reproducera en språklig hierarki.⁹⁹ Språk kan alltså enligt flera språkforskare som Sara Trechter och Mary Bucholtz ses som ett kulturbärande fenomen, ett objekt anslutet till en viss kultur.¹⁰⁰ Vi kan även se att språk är något som markerar grupper, men även vad en person tror om grupper och hur denna tro om gruppen sprider sig till andra. Språket skapar en markör för identifikation och ligger ofta till grund för stereotyper, som därmed upprätthåller sociala orättvisor. Dessa maktstrukturer blir sedan inprogrammerade via NLP [natural language processing] i AI.¹⁰¹ Något som i sin tur riskerar att upprätthålla en social och kulturell hierarki.

Med en möjlig partisk användning i bakhuvudet ska vi nu gå in på hur myndigheten arbetar med jämställdhet och jämlikhet genom språk. Genom interaktionerna med informanterna hade alla en gemensam faktor, arbetet med språk utförs med efterfrågan i åtanke. Vilka språk förekommer mest frekvent och vilka förekommer mindre frekvent eller inte alls. Grundspråket i chatboten är alltså svenska som vi tidigare nämnt. Vad betyder detta för maktstrukturer och kan detta upprätthålla sociala hierarkier som Su Lin Blodgett, Solon Barocas, Hal Daumé III

⁹⁷Blodgett, Barocas, III och Wallach (2020) Sida: 5458.

⁹⁸Ibid, Sida: 5458.

⁹⁹ Mattsson (2015), sida 18–149.

¹⁰⁰ Trechter och Bucholtz (2001), Sida:6.

¹⁰¹ Teknik inom datavetenskap för information som sägs eller skrivs på "naturliga språk". Inkluderat är maskinöversättning, taligenkänning, syntetiskt tal med mera. [Översättning från it-ord.idg.se.]

och Hanna Wallach menar att den språkliga bristen i arbetet med NLP riskerar att åstadkomma?¹⁰²

Citatet nedan beskriver en av informanternas upplevelser när det kommer till språken som inkommer till chatboten i interaktioner där beskriver informanten hur efterfrågan av språk ser ut och hur de prioriterar vilka språk som ska programmeras in.

Jag har inte känt spontant att eller sett så jättemånga som frågar på ett annat språk så att jag kände ja, men den där måste vi bygga till exempel. Nu har jag jobbat ganska länge med chatboten och jag tror att jag har sett tre eller fyra som har tjatat på att dom önskar ett annat språk [i chatboten]. [Och] då känner jag inte att jämfört med hur många konversationer som kommer in till chatboten, jag känner inte att det var något behov av att bygga någonting på [de språken]. Det just för att det är fyra som chattar med chatboten när det gäller jag vet inte så om det är besked eller [borttaget av anonymiseringsskäl], men engelska har vi mycket.

Som vi ser så är språkvalet ett medvetet val att inte ta med ett språk baserat på hur många som efterfrågar eller skriver på ett visst språk. Vi kan även se från en annan informant hur de tänker kring språkvalen:

Det är ju också en resursfråga, ja, det är ju inte bara så att om man ändå väljer att litet område. men att man där märker att efterfrågan är stor på att annat språk då är det dels att själva svarstexterna ska fungera på engelska ju kanske inte är det området göra varje område. Även om jag tar ett litet område till exempel här med [borttaget för anonymiseringsskäl]. Så är ett ganska begränsat område kunskapsmässigt, eller vad man ska säga.

Här lyfter informanten alltså även upp resurser samt hur en samordning av de olika svarsområdena ska göras på det nya språket. Därför menar informanterna att nästa språk som

¹⁰² Blodgett, Barocas, III och Wallach (2020), Sida: 5458.

ska programmeras in kommer att vara engelska, därför att kunskapen finns i teamet från alla olika områden:

Nej, det här skriver vi hela meningar ju och vi brukar ju liksom ta och se hur folk skrivit här [till chatboten] innan i och med att vi går igenom konversationer och speciellt kanske vi brukar fokusera på de här eller de här svaren som vi vet att det är väldigt många personer som skriver om till exempel. [...].

Även ovanstående citat visar på en metod baserad på efterfrågan, genomgången av konversationer fokuseras specifikt på saker som väldigt många skriver. Här handlar det däremot om saker som inkommer och som personer behöver få svar på, vi kan alltså se likheter mellan språkvalen och hur svar programmeras in baserat på hur många som skriver en fråga eller hör av sig till chatboten. Tittar vi närmare på arbetet med språk inom teamet ser vi att teamet har flera saker att förhålla sig till när det kommer till programmeringen av språk.

Så måste ju vi förhålla oss till det här liksom nätverket och de här algoritmerna, AI- motorn och allt det här och får det till att bli rätt. Den kunskapen besitter inte vi på pashto eller arabiska eller något liknande.

Dessa inkluderar algoritmer som ska programmeras korrekt, AI-n ska förstå och kunna interagera utan problem. I citatet beskriver informanten hur de i teamet inte besitter kunskapen att kunna göra detta på språk som till exempel pashto eller arabiska. Här kan vi se att det som Mattsson skriver om normer och diskurser påverkar vad som kan programmeras in, den språkliga diskursen i svenska myndigheter utgår ifrån det svenska språket. Mattsson menar att en reproduktion av dessa diskurser och normer påverkar hur personer ser på sig själva och andra i sin omgivning. Genom Mattsson kan vi alltså diskutera hur dessa diskurser påverkar chatboten, en utgångspunkt i en språkligt svensk norm skapar en svensk språklig normprodukt, som sedan riskerar att reproducera och upprätthålla vissa normer i samhället samt påverka hur interaktionen med chatboten och andra ser ut.¹⁰³

¹⁰³ Mattsson, (2015) Sida 148–150.

Binder vi samman språkliga hierarkier och efterfrågan kan vi se hur arbetet med efterfrågan kan påverkas av lagstiftningen till interaktionen mellan användare och chatbot. Vi ser i intervjuerna hur flera av informanterna refererar personer vidare till Google translate. ”[Om] det är någon arabiska, men det är alltså det går ju typ att bara kopiera texten och kolla på i Google translate”. Detta för att möjliggöra en interaktion mellan användaren som kanske inte kan svenska och chatboten. Om vi då tittar till utgångspunkten i myndighetsspråket kan vi se att lagstiftningen främjar svenska men även har ett krav på att det ska finnas information på andra språk tillgängligt, något som dock inte alls får samma uppmärksamhet eller utrymme. Blodgett et al att det är fallet för språk där språket främjar en viss grupp som sedan får tillgång till ett till rum.¹⁰⁴ Det språkliga har som Fairclough menar skapat en kollektiv gemenskap, något som därmed utesluter andra som inte anses vara en del av gemenskapen.¹⁰⁵ Språket representerar en kultur, som vi hört tidigare, denna kultur blir därmed främjad och riskerar att utgöra en partiskhet. Att referera en person vidare till Google translate kan ses som en bra lösning, men även här riskerar vi att upprätthålla maktstrukturer genom att lägga ansvaret på att kunna ha en konversation med chatboten på användaren. Att referera vidare kan ses som en utövning av makt, något som i sig skapar problematik som att upprätthålla svenska eller engelska som de universella språken som de flesta kan i Sverige. Detta leder till att flera språk inte blir inkluderade, något vi även kan diskutera genom Ahmed och institutionalisering. I citatet nedan så beskriver en av informanterna hur det svenska språket internt på myndigheten skapar en sorts bubbla där ett antagande om språkliga kunskaper kan identifieras enligt egna ord.

Man utgår alltså ja, nu kan jag väl bara tala för mig själv, liksom i det. Men ibland utgår man väl från att man liksom är i Sverige. Man jobbar på en svensk myndighet, eller så tar man lite för givet? Kanske alltid att ja, men det är det svenska som gäller internt här på myndigheten så är det också bara svenska språket liksom så det är. Det är väl lite det man tänker, liksom att man är i den här bubblan. [...] Så det är så i att man tar kanske för givet ibland att man tänker att alla som är i Sverige kan svenska, men så är det ju inte.

¹⁰⁴ Blodgett, Barocas, III och Wallach (2020), Sida: 5485

¹⁰⁵ Fairclough, (2010), Sida 27

Vi kan här se hur det svenska språket utifrån egen erfarenhet av arbetsmiljö och lag riskerar att utgöra en självklarhet eller något som bara görs utan att det medför några större funderingar. Det svenska språket är institutionaliserat på en lagstiftande nivå. Även om detta innebär att information måste finnas tillgänglig på andra språk på hemsidan kan vi se att det påverkar hur teamet jobbar och vilken språklig kunskap som finns i teamet. Arbetet med språk utgår alltså enligt informanterna från en efterfrågan, förutom det svenska språket beskriver de hur vissa språk inkommer oftare än andra. Något som vi diskuterat kort tidigare, är att ett arbete med efterfrågan kan utgöra en fördel i och med att majoriteten får sin röst hörd, men det kan även i relation till andra faktorer utgöra något mer problematiskt. De språk som inkommer mest får prioritet att programmeras, något som utesluter flera andra och återigen skapar och upprätthåller maktstrukturer.

Tittar vi vidare på institutionaliseringen av språket kan vi återgå till Mattsson och se hur hon menar att språk är kulturellt bärande och att en diskurs som utgår från ett specifikt språk kan påverka hur personer interagerar och tänker kring andra personer. Om språk är något kulturbärande så kan vi se att risken att överföra kulturella traditioner ökar genom språket. Om vi då ser tillbaka till Ahmeds teori om förhållningssätt och riktning mot samhället och objekt så kan vi se att detta även inkluderar språket. Som Mattsson beskriver kan vi se att svenska språket oftast associeras med vithet genom historien, något som är en del av en vit svensk kultur. Mattsson menar att med en utgångspunkt i en specifik diskurs med utgångspunkt i det här fallet det svenska språket riskerar myndigheter att 'naturalisera' en viss ideologisk hierarki.¹⁰⁶ De naturaliserade ideologiska hierarkierna riskerar sedan att bli det som anses vara "sunt förnuft". Detta sunda förnuft riskerar sedan att bli något som utgör en reproduktion av maktförhållanden, genom att överföra diskurser i produkter och uttalanden, som sedan påverkar hur vi interagerar och tänker kring olika människor. Maktförhållanden, där svenska anses vara det naturliga för en svensk myndighet och att tillämpningar endast finns tillgängliga på det svenska språket. Vi kan alltså se att språket och diskurser bakom språk påverkar hur personer interagerar och ser på chatboten enbart genom språket, men att arbetet med efterfrågan både kan utgöra något bra och något problematiskt. Har vi redan en norm som utgångspunkt kommer denna att påverka vilka språk som inkommer mer än andra, vilket stämmer överens med det som Mattsson menar att en diskurs och ett språk riskerar att reproducera maktstrukturer.¹⁰⁷

¹⁰⁶ Mattsson, (2015), Sida 148–150

¹⁰⁷Ibid

Teamet har försökt att öka jämlikheten i chatboten genom att försöka att programmera in olika felstavningar som chatboten sedan ska förstå i interaktioner. ”Träningsdata så det är ju lite mycket eller testa att de är mycket felstavningar och sådant just för att testa modellen och se om den [chatboten] om klarar det.” Det vill säga bruten svenska eller grammatisk inkorrekt svenska, något som ökar tillgängligheten i viss mån, något som i viss mån motverkar det som Blodgett et al menar med att ett språk görs universellt. Informanterna beskriver att detta görs för att öka tillgängligheten, vi kan alltså se att när det kommer till språkbruket i chatboten så finns det en tydligare tanke på tillgänglighet och jämlikhet. Att vanliga felstavningar programmeras in kan ses som ett försök att motverka det som Mattsson beskriver skapar tydliga maktstrukturer, även om dessa fortfarande tar en stor plats i hur chatboten programmeras. I relation till detta lyfter informanterna att det finns en balansgång med hur långt de kan gå med att programmera in felstavningar, chatboten måste programmeras på ”korrekt svenska”, innan den sedan kan lära sig de felstavade orden för att göra tillgängligheten så bred som den kan blir med en chatbot som enbart kan ge svar på svenska. ”Till exempel kan man [ha] svårt att skriva, men där är det ju så att där lägger vi ju in till exempel vanliga felstavningar för att kunna matcha då och fastän [all] träningsdata är korrekt.” Felstavningarna blir alltså inte inprogrammerade från början, utan är något som även det är baserat på efterfrågan, chatboten är fortfarande baserad i en svensk kulturell grund. Även om felstavningar programmeras in och hjälper till att motverka hur maktstrukturer skapas kan vi ändå se hur chatboten från börjar riktar sig mer åt vissa än andra. Men att teamet anstränger sig för att programmera in vanliga felstavningar för att underlätta kommunikationen och att alla ska få samma svar oavsett stavning.

Fairclough menar att språket är socialt konstruerat samtidigt som det konstruerar den sociala världen. Faircloughs argument beskriver diskurser som utgår ifrån ett språk, i vårt fall svenska eller svenskhet kan sägas påverkas av att personer som identifierar sig med svenskhet eller med ’det svenska’.¹⁰⁸ Något som Mattsson menar även kommer påverka andra i deras konkreta vardagsliv i sammanhang som hur de ser på sig själva och andra samt hur de agerar i olika situationer. ” [...] Det är det svenska språket som gäller här på myndigheten internt och då blir det bara så att det är det svenska språket som är det som gäller [...]” beskriver informanten. Citaten beskriver tillståndet av att vara i en bubbla där det blir lätt att ta för givet att alla i Sverige kan svenska, även fast detta inte stämmer. Som vi såg tidigare är språket enligt Mattssons text något som både är socialt konstruerat och konstruerar den sociala världen.¹⁰⁹

¹⁰⁸Ibid Sida 149–150

¹⁰⁹Ibid.

Om vi tar informantens exempel om bubblan kan vi se att detta stämmer in ganska väl. Språklagen säger att det svenska språket gäller på svenska myndigheter, det svenska språket blir därmed något självklart inom myndigheten som sedan kan påverka hur tankarna kring det svenska språket går.

Det är lite så mellanmjölk. [Språket ska inte sticka ut] det ska liksom inte sticka ut åt något håll utan det ska vara en klartext där vi ska försöka att i alla fall skriva så enkelt som möjligt så att man förstår det.

I det här fallet kan vi diskutera vad ordet mellanmjölk betyder i sammanhanget, betyder det att är ett förenklat svenskt språk som är tillgängligt för alla? I sammanhanget menar informanten att detta är ett språk som representerar just klarspråk, något de flesta kan förstå. Marino skriver i sin artikel ”The Racial Formation of Chatbots” hur skapandet av en chatbot alltid innebär en viss mån av rasifiering.¹¹⁰ Marino menar precis som Mattsson att språk är något kulturbärande, och alltid kommer ha en viss koppling till en viss kulturell egenskap. I det här fallet blir det svenska språket något som skapar en koppling till svenskhet som i sig utgör en koppling till det vita. Detta innebär att en chatbot alltid kommer ha en rasifiering genom det kulturbärande språket.

Fortsätter vi med tankarna kring det bekanta och hur vi upplever objekt samt hur dessa ankommer till oss, kan vi börja diskutera hur mycket bekanta saker som Ahmed beskriver lätt kan utestänga andra perspektiv. Vi kan även som Ahmed beskriver förlora medvetandet av objektet när vi använder detta, precis som med skrivbordets användning kan medvetandet kring det försvinna.¹¹¹ Om vi fortsätter titta på arbetet med jämställdhet som något bekant och hur en person är beroende av sin egen uppfattning av samhället skapat av en viss hållning och riktning emot objekt, vilket innebär att personen har en större tillgång till vissa miljöer än andra. Vi har tidigare diskuterat hur jämställdhet har blivit något bekant genom en vit och svensk uppfattning om samhället något, som inte har en tydlig koppling till en rasifierad hållning till samhället och som därmed kan anses att ha det enklare att relatera mer till könsmaktsordningen i samhället. Om vi ser till skrivbordet och hur det runtomkring försvinner kan analysen se på hur en naturlig

¹¹⁰ Marino (2014).

¹¹¹ Ahmed (2006), Sida:29–30.

inställning kan påverka hur en medvetenhet och en kritisk tanke kring arbetet och den egna hållningen ser ut. I följande paragrafer analyseras hur och om informanterna har en medveten och kritisk syn på arbetet, när arbetet med jämställdhet blivit bekant och en vana. I intervjuerna lades ett stort fokus på arbete med jämställdhet och jämlikhet, vi har tidigare diskuterat hur jämställdhet utifrån hållning och uppfattning av fenomen kan ta en större plats och omedvetet medföra en utestängning av andra objekt.

I studien kom det upp hur arbetet med jämställdhet kommer naturligt och om det i detta finns kritiska tankar kring egen hållning och egen bakgrund, något vi ska diskutera nedan. I frågorna som ställdes kring detta inkluderades perspektivet vithet och om en kritisk tanke mot vitheten är involverad i arbetet.

Nej, det är ingen som har varit alltså, det är ingen som har väckt frågan. [...]. Nej, ur det perspektivet har vi egentligen inte haft några diskussioner alls.

Ovanstående citat beskriver hur det inte är någon som riktigt väckt frågan om kritiska tankar och reflektioner som vit och privilegierad i samhället. Citatet beskriver att endast en person skulle ta upp något av det slaget. Vi kan utifrån detta citat se hur vitheten inte alltid syns på samma sätt och som därmed inte alltid tas upp som ett kritiskt diskussionsämne. Utifrån tidigare citat kan vi konstatera att arbetssättet anses vara något som kommer automatiskt och är en naturlig del av teamets arbete, något som gör det relevant till det som Ahmed menar kan stänga ute andra objekt i närheten. Hur kan perspektiv av vithet i det fallet ha försvunnit?

Som Ahmed presenterar i sin fenomenologi så är en persons riktning mot samhället väldigt svår att komma ifrån när vi använder oss av en naturlig inställning till något. Vi är alltid påverkade av det som vi ser och uppfattar utifrån vår egen riktning mot ett objekt.¹¹² Detta är inte något vi gör medvetet, vår hållning till världen styr vårt medvetande utan att vi själva är medvetna om det. Detta kan vi likna med det som Ahmed diskuterar kring det bekanta, något bekant skapas genom hållning och uppfattning av och mot samhället och personens plats i samhället. Om vi här ser att en riktning mot ett objekt är baserad på en vit hållning till världen kan vi se det som en anledning till att detta inte diskuteras. Den egna orienteringen blir för oss alla något vi inte

¹¹²Ahmed (2012) Sida: 116.

kan styra, därmed kan vi inte heller se den på samma sätt som andra kan utifrån. Om vi tittar tillbaka på det vi diskuterade i tidigare avsnitt kan vi se att en hållning mot samhället kan påverka vad en person uppfattar samt kan påverka vad som är enklast att relatera till. Inom institutioner menar Ahmed att dessa utgångspunkter och förhållningssätt till samhället kan överföras till arbete och bli en del av det som brukar göras, alltså att en egen uppfattning och allt det andra blir utestängda.¹¹³ Det som är bekant finns i personen och det 'främmande' hamnar i bakgrunden. Ahmed menar att objekt tar form av vad som görs med dem och att detta formas av vår uppfattning av objektet. Hur en kropp bebor ett rum i en miljö påverkar hur kroppen sedan relaterar till andra saker. Om kroppar i vårt fall bebor ett rum på ett visst sätt kan vi diskutera hur orienteringen skapar något som blir utestängt. I nästa citat ser vi att kritiska diskussioner kring egna hållningar inte är något som tar någon större plats.

Nej, det är ingenting som jag har tänkt på speciellt och det är ingenting vi tänker på när vi skriver texterna eller svaren.

Diskuterar vi detta utifrån en kroppslig hållning och riktning mot objekt med hjälp av Ahmed så kommer vi tillbaka till hur vi bebor ett rum. Går vi tillbaka till det som Ahmed beskriver med Husserls skrivbord ser vi hur hans inställning emot skrivbordet och papperet så kan vi se att detta skapar en förutsättning för Husserl att utföra sitt arbete som filosof.¹¹⁴ Om vi gör samma sak med citaten och tittar tillbaka på dem till de vi såg tidigare om den naturliga inställningen och det latent sätet att arbeta, kan vi se hur en persons hållning till och uppfattning av samhället kan påverka hur och vad den uppfattar i ett objekt. Att kritiska perspektiv inte är så frekvent förekommande kan vi möjligen förklara med hjälp av ett förhållningssätt till samhället utifrån sin egen uppfattning av objekten jämställdhet och jämlikhet. Objektet och arbetet med jämställdhet formas alltså utifrån en specifik hållning till och mot samhällets sociala strukturer. Om vi då ser denna tillgänglighet till objekt så kan flera perspektiv och uppfattningar försvinna in i bakgrunden. Om ett vitt förhållningssätt eller ett språk klassat som en del av en vit gemenskap ligger till grund för skapandet av ett objekt kan vi därmed se att chatboten kommer ta en liknande form. Något som kan förklara varför objektet skapas såsom det gör.

¹¹³Ahmed (2006), Sida: 132.

¹¹⁴Ibid, Sida 27–28.

Tittar vi tillbaka på det vi sa tidigare i avsnittet om att en orientering har lättare att nå vissa objekt än andra och hur en kropp bosätter sig i en miljö, påverkar därmed vad som kan nås och hur en person kommer att förhålla sig till dessa objekt. Olika förhållningssätt eller utgångspunkter när det kommer till objekt riskerar att påverka arbetet med jämställdhet och jämlikhet inom teamet. Utgångspunkten formar resultatet av arbetet, där det svenska språket ger chatboten tillgång till den svenska kollektiva gemenskapen. Det svenska språket och den svenska myndigheten utgör en del av den svenska kollektiva gemenskapen som i sin tur utestänger andra uppfattningar och förhållningssätt till och av chatboten. Svenskheten riskerar därmed att ockupera arbetet, något som Ahmed menar gör oss för upptagna för att se en tydlig och nyanserad bild av arbetet och programmerarnas förhållningssätt mot objekt.¹¹⁵

Något som framkom i frågorna till kritiskt tänkande om sina egna förhållningssätt mot objekt var hur informanterna såg arbetet med jämlikhet och jämställdhet som något som utförs i samma frekvens. En del av informanterna beskrev det som att arbetet med jämställdhet och jämlikhet går hand i hand, ingen av de två läggs mer fokus på. Vad vi kan se genom att återgå till citatet ”[...] så liksom pratar vi väl inte jättemycket om jämställdhet, inte dagligen utan det ligger väl lite latent [...]”, är att fokuset läggs på jämställdhet. Vi kan även kort nämna arbetet med namn, vi kan se att fokus har lagts på svenska namn, som en informant menar har det omedvetet använt sig av nästan enbart svenska namn. ”[...] Där har vi nog omedvetet, liksom utgått från svenska namn nu när man tänker efter [...]”. En av informanterna påpekade att de arbetar mer med jämställdhet mellan män och kvinnor, ”just kanske med det manliga och kvinnliga där [...] nej, utan det blir nog bara namnen och könen i det här fallet.” Med det tidigare citatet som anser att de inte arbetar tillräckligt mycket med icke svenska namn kan en dra slutsatsen att det i intervjuerna framkommer att arbetet med jämställdhet är tydligare i en större skala i jämförelse med jämlikheten.

Om vi tittar på samma sak när det kommer till språk så kan vi se att det där finns en tydligare struktur i hur arbetet med jämlikhet och jämställdhet utförs. Vi kan identifiera att detta inkluderar medvetna val om vad som ska vara med och inte, genom efterfrågan. Men även genom egna diskussioner anslutna till problematiska situationer kring ställningstagande och resurser. I arbetet med chatboten baseras programmeringen av stötande ord på efterfrågan, de ord som fångas upp mest i konversationsanalyser blir sedan tilldelade ett eget intent. Orden

¹¹⁵Ibid, Sida: 41.

fångas när intenten har programmerats och AI tränarna har preciserat algoritmerna för chatboten och chatboten interagerar med offentligheten. Intent bestående av stötande ord programmeras sedan in i chatboten genom träningsdata och testmeningar. Omfånget av dessa konversationsanalyser beskrivs vara 10% av de inkomna konversationerna, dessa väljs ut och dras slumpmässigt från dagar. Konversationsanalysen i sig är menad att ge AI-tränarna feedback på vad som fungerar bra och mindre bra. Det är alltså här de stötande orden kan påträffas. Nedan ser vi en tabell över ord som frekvent framkommit i konversationsanalyser och som har programmerats in att ge svar på tal. Orden i tabellen är tagna från teamets egen databas presenterad i ett Excel ark och är baserad på teamets egna grupperingar av ord och kategorier, ord har valts ut genom att titta på vilka som förekommer mest [full lista i bilaga 2].

Tabell 1.

Elaka uttryck	Grövre elaka uttryck	Rasism	Svordomar	Grövre svordomar
Din jävla pissråtta.	Böggjävvel	Du är en liten rasist du alltså?	Du är skit	Du ser ut som en hora
Dra åt helvete.	Din morsa är en hora.	Jag är lite rasist	Jävla skit	Hora
Skitskalle.	Jag ska döda dig	Rasistiska robot	Jävlar vad dåligt	Käftan jävla hora
Ta dig i röven.	Käftan din dumma hora.	Ska du säga som är rasist.	Skit också.	Jävla hora.

På dessa ord har chatboten blivit inlärd att med formuleringar som ”nu tycker jag vi höjer språket lite” eller ”jag tänker inte bemöta det”. Här har alltså orden som inkommit mest frekvent i interaktioner med användare och anses stötande lagts in.

[...] det [lagt in rasistiskt stötande ord i databasen] har vi faktiskt inte gjort där jag kollade nu, men där ligger liksom inga sådana ord inlagt där utan det som vi har lagt in på. Det är det som vi kallar i det här systemet, i alla fall ovårdat språk ligger den då, men det är ju de svordomarna som vi nämnde innan jävla helvetes skit fan. Vissa könsord till exempel ligger där också, men inget så annat liksom, [och] det vi har väl inte riktigt tänkt att det kanske har behövts lägga till de orden eller i och med att vi inte har riktigt stött på dem.

Citatet ovan beskriver hur vissa rasistiska ord inte har blivit inprogrammerade, citatet beskriver att det ligger en del könsord inprogrammerade men när det kommer till rasistiska ord har teamet inte tänkt att det kanske behövs läggas till. Detta för att de inte tidigare har stött på dem i en högre frekvens: en positiv sak. Men varför inkommer inte dessa ord i samma frekvens som andra opassande och stötande ord utan rasistiska antydningar? För att diskutera detta kan vi använda oss av Sara Ahmed och hur förhållningssätt till samhället påverkar hur vi uppfattar och upplever samhället.¹¹⁶ Tillsammans med Ahmed kan vi även använda oss av Fairclough och Katz¹¹⁷, där en hållning till objekt och samhälle enligt Ahmed skapas utifrån andra och den egna upplevelsen av världen, något som sedan ger oss en subjektiv bild av verkligheten, som sedan riskerar att, som Katz menar, genom självmodellen överförs till chatboten genom språk och tidigare nämnda subjektiva upplevelser. Inom AI så ser vi vidare språkets betydelse, språkets betydelse bär på kulturella identifikations-markörer enligt Fairclough.¹¹⁸ Vi kan genom dessa tre då fortsätta att diskutera hur chatboten reproducerar och upprätthåller en viss typ av maktdynamik.

Den subjektiva hållningen och uppfattningen av världen minskar kapaciteten att se hur andra hållningar och personer upplever världen, något som sedan riskerar att utestänga de främmande

¹¹⁶ Ahmed (2006), Sida: 41

¹¹⁷ Katz (2020), Sida: 153–156

¹¹⁸ Fairclough (2010), Sida 27

perspektiven. Vi kan som vi tidigare nämnt se hur en egen hållning och uppfattning överförs till AI genom språk och ideologier, alltså beroende på vem som programmerar en chatbot kommer chatboten konstrueras annorlunda. Precis som Weyerer och Langer menar programmeras därmed egna erfarenheter och sätt att förhålla sig till världen in i chatboten som konstruerar chatboten utefter personen som programmerar denna. Vi kan alltså se hur en vit uppfattning löper stor risk att konstruera en vit chatbot utifrån språk och egna erfarenheter. Något som ökar risken för att utelämnade 'främmande' erfarenheter av objekt, dessa utelämnade uppfattningar placeras i en underordnad position i chatboten då den språkliga tillgängligheten och gemenskapen inte finns tillgänglig som för de inom gemenskapen. Som vi ser så spelar språk och programmerare bakom teknologi en stor roll för hur en chatbot rasifierats. Vi kan även se genom Marino att chatbotar alltid kommer vara och förbli konstruerade efter en viss ras eller kulturell tillhörighet genom språket.¹¹⁹

Utifrån detta kan vi alltså se hur vita förhållningssätt och riktningar mot världen som utgångspunkt riskerar att konstruera en vit chatbot. En maktposition skapas där svenska och dess gemenskap får en högprioriterad plats, till skillnad från dem utanför gemenskapen. Fairclough menar att gemenskapen är skapad av och upprätthålls av språket som en kulturell markör, vi kan därmed se att chatboten genom att enbart kunna kommunicera på svenska inbjuder till en större tillgänglighet för personer inom gemenskapen. Om vi tittar vidare på hur detta kan påverka interaktionen mellan person och chatbot så kan vi fortsätta på Faircloughs spår där gemenskapen blir central. Vi kan se i chatboten och i Tabell 1 att rasistiska ord inte utgör något frekvent förekommande, vi kan diskutera detta utifrån Faircloughs gemenskap. Om vi säger att chatboten konstruerats som vit så kan vi titta på hur andra i samma gemenskap interagerar med varandra, sällan så avses rasistiska ord mot en vit person till en annan vit person eller en icke-vit person till en vit person. Vitheten blir inte rasifierad utan anses vara något universellt, något som kan förklara varför dessa rasistiska ord inte frekvent förekommes. Tittar vi vidare på Tabell 1 så kan vi se hur chatboten däremot med en större frekvens bli anklagad för att vara rasistisk, något som kanske kan förklaras genom att inte ge svar på tal på sådana ord och en vit konstruktion. Chatboten kan anses en del av en vit kollektiv gemenskap som inte svarar eller tar ställning till rasism. Chatbotens okunskap i till exempel rasism utgör en risk för framtida artificiella intelligenser. En AI utan ställningstagande mot rasism motverkar inte rasism utan gör den snarare passiv, något som existerar men inte är ett problem.

¹¹⁹ Marino (2014).

I flera av fallen har valet att inte programmera in rasistiska ord i chatboten presenterats som en resursfråga, att det inte är möjligt att tänka ut varje litet ord som kan ha nedsättande ton.

Det där med bredden att liksom helt gardera sig. Det är ju också en resursfråga, ska vi tänka ut varje exakt varje litet ord som en person kan använda i nedsättande ton? Om något som man kan knyta till diskrimineringslagen, regeringsform eller vad det nu är som någon aldrig kommer att skriva?

Citatet ovan presenterar något intressant, det presenterar programmerandet av rasistiska ord som en resursfråga. Ska resurser läggas på att varje litet nedsättande ord som en person kan använda och knyta till diskrimineringslagen eller någon regeringsform? Citatet presenterar även att dessa ord är ”något som ingen någonsin kommer skriva”. Om vi först stannar vid resursfrågan, att behöva lägga tid på att tänka ut varje ord kan vi se att detta kan sammankopplas med det bekanta. Detta är utanför det bekanta och därmed något som hamnar i bakgrunden precis som med Husserls skrivbord och resten av rummet som blir suddigt och inte läggs märke till och därmed kräver extra resurser. En riktning mot objekt som inte är van vid att möta dessa typer av uttryck har med störst sannolikhet en annan inställning och därmed ser dessa på ett annat sätt. Om vi tittar tillbaka på hur programmeringen ser ut generellt så ser vi hur arbetet kommer automatiskt, ett arbete som representerar jämställdhet som det största.

Tittar vi tillbaka på ett utdrag från tidigare citat kan vi diskutera medvetenhet i det språkliga arbetet ”till exempel att diskrimineringslagen, regeringsform eller vad det nu är som någon aldrig kommer att skriva?”. Utdraget presenterar vad som kan tolkas som ett antagande att personer inte skriver eller kommer att skriva rasistiska eller diskriminerande saker till chatboten. Vi kan alltså identifiera ett medvetet val att inte inkludera arbete med vissa uttalanden på grund av efterfrågan och vad som skrivs och inte skrivs. Självmodellen i kombination med Ahmeds teori är verktyg vi kan använda oss av för att förstå hur detta beslut kan ha tagits som resultat av en programmering och efterfrågan. Här kan vi diskutera huruvida sättet vi orienterar oss eller blir uppfattade av andra påverkar sättet vi utför uppdrag. Katz menar att detta historiskt har varit ett problem som har lett till att teknologi har fortsatt att vara en produkt av det universella eller det vita, något som han menar stammar ur hur tidiga artificiella

intelligenser skapades av och med hjälp av en vit ideologi.¹²⁰ Katz menar att vitheten har maskerats och därmed fortsatt att bli återskapad inom den artificiella intelligensen. Något som han förklarar skapar en AI baserad på självmodellen. Jaget påverkar här hur en artificiell intelligens blir utifrån ideologier och uppfattningar av världen.¹²¹ Vi skulle kunna argumentera att chatboten utifrån självmodellen omedvetet har konstruerats som vit, därmed inkommer få rasistiska ord, något som speglas i det som citatet säger där ingen kommer skriva vissa ord. Vi kan alltså se att ett språk som kulturbärare kan ha påverkat chatbotens konstruktion och därmed påverkat hur offentligheten interagerar med den. Vi kan även se i Tabell 1 hur flera av de ord som inkommer som blivit markerade för svar innehåller ordet rasist i kontexten där chatboten utgör det rasistiska.

I tabell 1 kan identifiera att chatboten är programmerad att reagera på anklagelser om rasism som egenskap, men inte på rasistiska ord. Detta leder oss in på svaren kring ställningstagande.

Just de här rasistiska har vi egentligen valt att, vi hade en tanke från början och valde att ta bort [rasistiska ord] då chatboten inte fattar dem. Visserligen är lite tråkigt, [chatbotens svar] jag förstår inte vad du menar. Men det är ju ett väldigt neutralt och tråkigt svar, [som gör] att man kanske inte spinner vidare på det man känner att nej, den fattar ändå ingenting. Jag ska inte skriva mer, men å andra sidan är ju kanske bra att få ett svar.

Citatet beskriver valet att ta bort rasistiska ord på grund av chatbotens bristande kapacitet att förstå dessa, svaret blir i stället jag förstår inte vad du menar något som citatet beskriver som neutralt och tråkigt. Citatet menar att detta svar är så tråkigt att personen slutar skriva. Men beskriver också att det kanske hade varit bra att få ett svar. Här genererar rasistiska ord samma gensvar som en fråga chatboten inte förstår, något som kan anses vara problematiskt. Chatboten svarar med ett ställningstagande på ord som hora, helvete, jävla, skit, samt anklagelser om att chatboten skulle vara rasistisk. Svaren chatboten ger på dessa är som exempel ”Nu tycker jag vi höjer nivån lite”, medan rasistiska ord får svaret ”Jag förstår inte. Kan du skriva frågan på

¹²⁰ Katz (2020), Sida: 154.

¹²¹Ibid, Sida: 154–157.

ett annat sätt?”. Kopplingen mellan detta och det faktum att chatboten svarar ”Jag tänker inte bemöta det.” på anklagelser mot den själv om att utgöra något rasistiskt, blir intressant. Jämför vi detta med hur chatboten svarar när den själv blir kallad till exempel rasist blir detta intressant ur två synvinklar. Som vi såg tidigare så krävs det mycket av AI-tränare för att en chatbot ska kunna lära sig ett språk eller nya ord. I relation till de rasistiska orden kan vi tänka att chatbotens svar fungerar som blockande mot anklagelser om rasism mot sig själv, och har valt att ta ställning. Anklagelser om rasism är inte okej, dock så kan inte chatboten ge ett svar på tal mot rasistiska uttalanden om andra.

Vi kan inte vara ställningstagande kring den typen av frågor/påståenden och ge ett direkt svar på tal, vi har en försiktig framtoning kring vissa känsliga saker, med hänsyn till risken att chatboten uppfattar fel och ger ett sådant svar på en annan fråga. Som exempel kan nämnas att vi fångar upp frågor om ”barnäktenskap” men inte heller där kan vi låta chatboten svara på ett sätt som antyder att frågan handlar om just barnäktenskap.

Kopplar vi ovanstående till bristen på gensvar emot rasistiska och annars stötande ord får vi en förklaring. Här beskriver de hur chatboten inte ska lära sig dessa ord för att inte riskera att den ska använda dessa i andra kontexter. Här kommer vi in till hur chatboten ihopkopplar dessa ord med andra ord i meningsbyggnad. Om vi tittar på citatet ovan ser vi att det nämner barnäktenskap och hur detta inte ska riskeras att chatboten uppfattar fel och använder ordet i en annan mening som svar på en annan fråga som vi ser i föregående citat. Att inte programmera in ord som anses kontroversiella eller rasistiska ord görs med hänsyn till risken att inte chatboten själv ska kunna upprepa dessa i sina svar. Här finns det som vi diskuterade tidigare en medveten logisk process kring varför vissa ord inte får ta plats, risken att upprepa och att de inte finns en stor efterfrågan på dessa ord. Att detta anses vara ett medvetet val kan vi se i ett tidigare citat där informanten beskriver att rasistiska ord blev uteslutna från början;

Just de här rasistiska har vi egentligen valt att i alla fall. Vi hade en tanke från början och valde att ta bort då chatboten inte fattar dem. Visserligen är lite tråkigt. jag förstår inte vad du menar, men det är ju ett väldigt neutralt och tråkigt svar.

Ett exempel som togs upp i intervjuerna kring ämnet av chatbotens oförutsägbarhet handlade om ord som används i myndighetens vardagsspråk, ordet i exemplet var invandrare. Exemplet utgjordes av att skriva det helt falska uttalandet alla invandrare är kriminella till chatboten för att få en aning om vad svaret skulle bli. I svaret blev ordet invandrare ihopkopplat med en viss typ av brott. Att ordet invandrare blev ihopkopplat med en mening som refererar till brott visar på att citatet stämmer. I konversationer kring detta begrepp kom det fram att chatboten har en extraktionsmetod som plockar upp specifika ord och kopplar ihop dem med andra meningar.

Men det här är ju en extraktion av ordet invandrare då som gör att det kommer svaret att det är extraherat ur kundens fråga för att visa att chatboten har uppfattat någonting.

Citatet beskriver hur chatboten har extraherat ordet invandrare och satt det i en kontext av det kriminella ordet. Informanten beskriver senare hur detta har potential att fungera jättebra i vissa sammanhang såsom att chatboten kan komma ihåg namn och så vidare. Men att det även utgör en risk i situationer som denna. Här kan vi alltså se att även en övervakad chatbot kan lära sig att uttrycka sig i problematiska termer genom extraktionsmetoden. Om vi då tänker oss att extraktionsmetoden utgör en risk för hur chatboten interagerar kan vi se vidare i framtiden med AI, om en AI inte har kapacitet att ta ställning mot rasism så riskerar detta att leda till konsekvenser där vi inte längre har kontroll över den artificiella intelligensen i sin helhet och rasism blir något passivt i den artificiella intelligensen. Här menar Felländer et al att transparens och ansvarstagande spelar stor roll för att kontrollera hur algoritmer körs för att få kontroll och undvika framtida problem med rasism och teknologi. Något som vi ser tydligt redan idag, och som vi visat i tidigare avsnitt är något som är sant för AI generellt.

Liksom tänker jag bara utifrån mig att det är väl där som vi hamnar i att vi gör den liksom med de här enkla och svordomarna. För att man kanske inte behöver hamna i den här problematik som var man drar gränser och liknande utan då att vi har fokuserat på att chatboten inte ska ge svar på de orden.

I det här fallet i den svenska kontexten kan bortprioriteringar av dessa ord även leda till en förstärkt syn på antirasism och färgblindhet. Precis som Mattsson och Ahmed menar så upprätthålls omedvetna strukturer av politik och media som representerar vissa mer än andra och som väljer att belysa mer än andra. Dessa strukturer ser vi i chatboten om vi skriver till exempel ”jävla skit” då får vi ett svar som markerar detta som inte okej. Ett ställningstagande som tagits. Om vi sedan skriver ett ord som klassas som rasifierat och nedsättande mot en viss grupp av människor, får vi svaret att den inte riktigt förstår. Även ord som idiot och könsord får den markerade statusen som inte okej när den skrivs i chatten.

Valet att inte inkludera dessa ord, kan vidare diskuteras genom Fairclough och hans syn på diskurs och maktstrukturer. Fairclough menar att en diskurs består av ord och uttryck som vi själva valt, något som konstruerar en kollektiv gemenskap i en grupp.¹²² Vi kan här se att det medvetna valet att inte inkludera de rasistiska orden har skapat en chatbot som bortser från rasistiska ord och dess frekvens. Vi kan även se att en diskurs kring detta har skapats med hjälp av en djupare medveten tankegång, något som tyder på att detta inte är en naturlig inställning utan en medveten handling för att underlätta en säker och noggrann kontroll över vad chatboten ska och inte ska kunna svara i en interaktion med offentligheten.

6.3 Vithet inom Artificiell Intelligens

Som vi tidigare i avsnittet bakgrund konstaterat kan vi koppla ihop vithet med programmeringen av chatboten genom att titta på teamets uppbyggnad, något vi har diskuterat genom språk, arbetssätt, naturliga inställningar och vanor. Men hur påverkar allt detta chatboten i sig? Går det att rasifiera en chatbot? Marino diskuterar huruvida det går att rasifiera en chatbot eller inte, rasifiering av chatbotar menar Marino är något konstant förekommande. Marino

¹²² Fairclough (2010), Sida: 27

menar att till största del så påverkas rasifieringen av språk, något som han precis som Trechter och Bucholtz menar är kulturbärande och därmed bär på ideologier och stereotyper tillhörande specifika grupper.¹²³ Att språket bär på dessa föreställningar om personer och kulturer kategoriserar hur vi ser och relaterar till andra personer, något som i sin tur påverkar hur och vad som skrivs till chatboten. Detta på grund av att dessa stereotypa föreställningar kring ett visst språk har tilldelats en viss kulturell grupp. Om vi tittar på hur chatboten i studien blir rasifierad så kan vi se till de tidigare avsnitten, hur detta riskerar att skapa en rasifierad produkt som sedan riskerar att påverka interaktion och maktstrukturer. En rasifiering är som Marino menar omöjlig att undvika. I följande avsnitt kommer arbetssättet som presenterats i studien analyseras utifrån ett perspektiv av rasifiering och konsekvenser av detta.¹²⁴

Något vi tidigare i bakgrundsavsnittet har sett är hur vitheten har tagit stor plats i teknologin och inom AI, som vi har sett genom Dihal och Cave.¹²⁵ Dihal och Cave argumenterar att vithetens plats inom AI sträcker sig tillbaka till kolonialismen. För att undersöka hur detta riskerar att ta plats i ett arbetssätt som hos teamet i studien kommer vi här utgå till stor del från självmodellen som menar att egna erfarenheter har en stor roll att spela i hur vi upplever och hur vi skapar miljöer runtomkring oss.¹²⁶ Även Ahmeds teori kring egen riktning mot objekt blir relevant genom dess argument att orientering eller upplevelse inte kan se något som är främmande för den. Att inte kunna se främmande upplevelser bidrar till att skapa det som Ahmed kallar vanevärldar, som även dem bidrar till att bygga upp sociala strukturer. Enligt Merleau-Ponty blir jagets vanor osynliga, jaget och kroppen har skapat vissa vanor som just därför inte läggs märke till. Båda dessa utgår ifrån en bild av jaget och hur bilden av jaget påverkar och styr personers handlingar i miljöer som är bekanta för sig.¹²⁷ Denna bekantskap och bilden av jaget riskerar sedan att överföras till chatboten, för att skapa en kopia av jaget som programmerar och bestämmer vad som ska in i chatboten. Ett intressant citat från en av informanterna föreslår att chatboten blir en kopia av dem själva, om än en kopia som kan mer än de själva.

Det är klart man pratar om det så sätt och vis många gånger så kan det ju vara egentligen att man i sådana fall hänvisar till att du tar kontakt med exempelvis

¹²³ Trechter och Bucholtz (2001), Sida:6.

¹²⁴ Marino (2014).

¹²⁵ Cave och Dihal (2020) Sida: 687.

¹²⁶ Metzinger (2008), Sida: 218.

¹²⁷Ahmed, (2013).

[kundtjänst] där vi sitter, annars är det och det är ju lite samma sak där om någon ringer till mig i telefonin, det kan bara jämföra med det då för chatboten. Det är ju i princip egentligen bara en digital kopia av mig då, på det sättet men som kan betydligt mer än vad jag kan nu.

Det som informanten poängterar i citatet ovan blir intressant då citatet föreslår att chatboten blir en kopia av hen själv, vad som överförs och hur jaget överförs blir då en intressant fråga. I tidigare avsnitt har vi diskuterat hur egna personliga synsätt på jämställdhet och jämlikhet i samspel med det lagstiftade används. Om chatboten blir en kopia av personen som programmerar den pekar det i samma riktning som Marino menar, en rasifiering av chatboten blir oundviklig.¹²⁸ Som vi sett i tidigare citat så anser en av informanterna att chatboten blir "[...] egentligen bara en digital kopia av mig då [...]" Detta kan ses i relation till en annan informants citat om vilka det är som bestämmer vad som ska in i chatboten: "[...] sex personer är vi ju och jag menar, det är en ganska liten del. Av det stora hela och det är vi som liksom bestämmer allting där och hur det ska vara [...]" Att en liten del får chansen att bestämma allt som informanten beskriver riskerar att leda till det som tidigare informant beskrivit om att en digital kopia skapas. Vi kan här identifiera att en jaget överförs in i chatboten, alltså kan vi se att precis som Marino menar är det helt omöjligt att inte rasifiera en chatbot.¹²⁹ Om vi då ser till vad informanterna berättar kan vi se hur båda två menar att chatboten ett blir en kopia av de själva samt att deras egna värderingar riskerar att påverka och skapa en chatbot som reflekterar teamet. Vi kan alltså här se hur det som Katz¹³⁰, Cave och Dihal¹³¹ diskuterar kring vithet inom AI kommer in i bilden som en stor riskfaktor för hur teknologin utvecklas och vilka som bestämmer vad och hur allt ska vara. Samtidigt blir det viktigt att se till det som Marino menar att det är omöjligt att inte rasifiera en chatbot på ett visst sätt. Något som vidare pekar på vikten av att flera perspektiv och språk måste inkluderas för att undvika en vit maktposition inom chatboten.¹³²

Konstruktionen av vithet kan vi här se genom vane-världar, precis som vi tidigare argumenterat för att genom subjektiv och situerad medvetenhet säga att vithetens vane-värld enbart kan utgå från den subjektiva medvetenheten, något som förstärker det som Katz beskriver som

¹²⁸ Marino (2014).

¹²⁹ Ibid.

¹³⁰ Katz (2020), Sida: 153–156.

¹³¹ Cave och Dihal (2020), Sida: 687.

¹³² Marino (2014).

“technologies of whiteness”.¹³³ Detta sker omedvetet genom vanor som skapats genom kroppen. Kroppen förlorar därmed all sin uppfattning och kan därmed inte se vanorna själv, något som visar på vikten av att ha mångfald i ett arbetsteam.¹³⁴ Ahmed diskuterar detta utifrån begreppet institutionell vithet där hon menar att vita kroppar i institutionen skapar en upplevd samstämmighet.¹³⁵

Vi skulle kunna förklara hur informanterna förhåller sig till kritiska tankar om sin vithet och hållning till samhället genom institutionell vithet. Vitheten är inte något som läggs någon vidare uppmärksamhet på. Brist på uppmärksammande av vithet kan potentiellt förklaras genom vanan att omges av vita kroppar och samstämmigheten mellan dessa. Samstämmigheten i dessa kan sedan även komma att föra över vithetens förhållningssätt emot samhälle och upplevelser kring vanor till produkten som skapas. Om vi genom detta förhållningssätt går tillbaka och tittar på arbetssättet inom jämställdhet, ”Min tanke och sådant bakom det, så liksom pratar vi väl inte jättemycket om jämställdhet liksom inte dagligen utan det ligger väl lite liksom latent i en hela tiden, kanske man kan säga”. Som vi ser i citatet så att de inte pratar så mycket om arbetet med jämställdhet, i citatet kan vi se att tankar kring jämställdhet kan ha internaliserats och ligger latent hela tiden i arbetet.

Om vi börjar med det som Ahmed säger att en person inte kan uppleva det som inte ligger nära nog den egna subjektiva uppfattningen av samhället så kan se att en vit majoritet finns i teamet. En homogenitet som därmed riskerar att leda till ett mindre balanserat och mångfaldigt synsätt på jämställdhet och jämlikhet, som sedan riskerar att reflekteras i chatboten. Risker finns att aspekter som inte upplevs av uppfattning av samhället riskerar att omedvetet bortprioriteras på grund av egna upplevelser eller brist på negativa upplevelser i relation till hudfärg och diskriminering. Vi kan här se att en vit upplevelse av världen riskerar att utgöra en större plats i programmeringen av chatboten i och med den homogena gruppen, något som senare riskerar att överföra en bild av vithet som sedan påverkar hur interaktion av chatboten uppfattas.

Katarina Mattsson diskuterar även svenskhet som en maktposition, här blir makten att enbart en ras blir konstruerad samt att denna förtrycker chansen för icke-vita att ha möjligheten att stötande ord och vilket svar dessa ska utlösa.¹³⁶ Mattsson diskuterar även vad svenskhet innebär, två av hennes punkter är språk och kultur samt utseende. I språk och kultur ser vi det

¹³³ Katz (2020), Sida: 153–156.

¹³⁴ Katz (2020), Sida: 153–156

¹³⁵ Ahmed (2013).

¹³⁶ Mattsson, (2015) Sida: 149–150

svenska språket och svenska kulturvanor, i utseende ser vi att denna innefattar ett vitt utseende. Genom det svenska språket och konstruktionen av vithet så pekar chatboten mer mot vita kroppar. Något som sedan riskerar att påverka hur och vad som skrivs till chatboten samt att utelämna andra kroppar och chatboten förblir omedveten om rasistiska ord. Om vi går tillbaka i historien kring AI så kan vi enligt Yarden Katz se en vitifiering av teknologin, detta beskriver han genom självmodellen där en avbild av programmeraren eller den dominanta kulturen speglas i teknologin.¹³⁷ Där vi kan se exempel som ansiktsigenkänning, där vita män har större möjlighet att bli korrekt identifierade än till exempel en icke-vit man eller kvinna. Vitheten inom AI är konstant, otaliga exempel finns att diskutera.

Katz beskriver överföringen av ideologier och subjektiva perspektiv genom självmodellen som han menar utgör speglingen av vit ideologi inom AI, och vi kan diskutera hur dess hållning och riktning mot objekt riskerar att göra just detta.¹³⁸ Centralt i argumentationen är hur en vit representation skapar en modell av jaget i chatboten. Genom en begränsad tillgång till mångfald och perspektiv blir vitheten starkare och starkare. Något vi kan diskutera genom Mattsson och maktstrukturer, en diskurs som tar plats eller i det här fallet inte tar plats på diskursen kring vithet och AI tar plats i den utsträckning som är nödvändig. Bristen på en kritisk diskurs upprätthåller och förstärker en social hierarki där vithet har blivit ett universellt faktum i många områden.¹³⁹ Om vi tar Mattssons argument vidare kan vi se hur diskurser riskerar att överföra ett rashierarkiskt samhälle in till AI, något som flera forskare menar redan har gjorts genom historien. Vad händer om enbart en sida får ta plats. Här riskerar vi att skapa ett samhälle som reflekterar dagens sociala hierarkier i en maskin som i framtiden kommer anta dessa och sedan riskerar att agera på dem.

Som vi tidigare diskuterat menar Ahmed att kroppar och deras utseenden skapar en uppfattning om världen omkring sig, genom egen uppfattning men även genom andras uppfattning av kroppen.¹⁴⁰ Om vi utgår från vithet så kan vi se, genom att titta på sociala hierarkier se att vithet för med sig en helt annan tillgänglighet i världen vilket påverkar personers handlingsmöjligheter. Dessa handlingsmöjligheter är givna automatiskt genom en ljus hudfärg, något som kommer att påverka uppfattningen av världen. Här kommer vi in på hur andra erfarenheter kan påverka en naturlig inställning, omgivningen kanske har gjort en medveten om

¹³⁷ Katz (2020), Sida: 153–156

¹³⁸ Ibid, Sida: 153–156

¹³⁹ Mattsson, (2015), Sida: 149–150

¹⁴⁰ Ahmed (2006). Sida: 110–111.

de sociala hierarkierna eller en person har fått se livet som andra lever. Här kommer frågan in hur mycket av detta måste en person uppleva för att få en tydlig bild av hur det är att vara icke-vit till exempel. Om vi inte har upplevelser som den icke-vita personen har är det enligt Ahmed omöjligt att se bortom sitt eget förhållningssätt till världen baserat på den egna hudfärgen samt hur andra uppfattar en. Något som resulterar i att ett vitt förhållningssätt fortsätter vara aktuellt i ens liv. Något som sedan riskerar att skapa det som Katz uttrycker som ”technologies of whiteness” genom en naturlig inställning mot ett objekt, i det här fallet jämställdhetsarbete.¹⁴¹

6.4 Neutralitet

I arbetet med jämställdhet och jämlikhet i chatboten så framgick det i majoriteten av intervjuerna hur målet var att uppnå en så neutral chatbot som möjligt. Om vi fortsätter att följa spåret av det bekanta även här kan vi fortsätta att diskutera erfarenheter och orienteringar. Om vi som i förra avsnittet kom fram till att en familjaritet antingen kan leda till ett arbete som är automatiskt eller inte kan vi dra liknande slutsatser när det kommer till neutralitet.

Något som informanterna beskrivit i flera av intervjuerna var den bristfälliga tillgången till information om användarna av chatboten, något som de beskrev som både något positivt och i vissa fall negativt. När det kom till neutralitet så tyckte några att det faktum att de inte visste vilka som interagerade med chatboten gav dem en neutral grund redan där. Om vi tar detta exempel och spinner vidare på så kan vi analysera huruvida neutralitet är möjlig att uppnå eller inte. Citatet nedan är baserat på det faktum att teamet inte har en tillgång till information om användare av chatboten. ”Nu låter det hårt, men vi tar ingen hänsyn till den som eller vem som chattar eftersom vi inte vet, så då blir det väl neutralt mot alla.” Här beskrivs ovetskapen om vilka som chattar som en given grund för neutralitet. Går vi tillbaka till det som vi i förra avsnittet diskuterade så kan vi börja att diskutera om detta är en grund för att vara neutral mot alla. Som vi tidigare såg menar Ahmed att medvetenheten och orienteringar mot objekt är något förkroppsligat, känsligt och subjektivt.¹⁴² Om vi ser att ett medvetande är situerat utifrån en viss uppfattning kan vi diskutera om en neutralitet är möjlig. En etablerad hållning innebär att vi har en personlig utgångspunkt formad av våra egna erfarenheter. Om vi utifrån en riktad medvetenhet diskuterar neutralitet så måste vi börja med att förtydliga att en persons

¹⁴¹ Katz (2020), Sida; 153–156.

¹⁴² Ahmed (2006), Sida: 116.

förhållningssätt är en persons bakgrund, socialt, språkmässigt, könsmissigt och så vidare. Hur en person riktar sig mot objektet neutralitet när den skapar chatboten kan omöjliggöra en neutralitet, denna omöjliggörs på grund av personens subjektiva uppfattning av neutralitet.

Ahmed diskuterar användandet av objekt genom att se till funktionen eller icke-funktionen av dessa, en hammare har en performativ uppgift att slå spikar in i något. Om denna går sönder eller inte kan användas av någon förlorar denna sin performativa förmåga.¹⁴³ Ahmed menar dock att hammaren inte behöver vara trasig för alla eller att hammaren inte kan användas av någon annan. Precis som hammaren har en funktion för en person kan den ha en annan eller ingen funktion för någon annan beroende på olika förhållanden som styrka för att orka lyfta hammaren och så vidare. Även om vi inte vet vem som skulle kunna använda denna så finns denna person någonstans. Alltså en hammare som anses vara fullt funktionell för någon kan anses vara helt oanvändbar för andra. Precis som i tidigare kapitel så kan vi diskutera detta utifrån ett förhållningssätt. Något kan anses neutralt av någon men samtidigt inte alls vara neutralt av någon annan, något som är beroende på situation i hammarens exempel kunskap och kroppslig styrka.

Eftersom uppfattningar skapas genom korsningen av en i det här fallet ett vitt förhållningssätt och neutraliteten presenterad i chatboten, skapas det som Ahmed menar blir subjektivt och därmed inte neutralt. Personens utgångspunkt eller arbetssätt sammanfaller väl med vad som anses neutralt, där en annan persons förhållande till neutralitet skiljer sig. Det blir tydligt hur en uppfattning av samhället kring neutralitet påverkas av inställningen som möter objektet i fråga. Utgår vi ifrån förra kapitlet kan vi se att neutralitet skapas genom namn för det mesta, något som kan ha påverkats av personers riktning mot objekt och utgångspunkt, därigenom dess uppfattning av vilka fenomen som presenterar sig för en person. Vad som anses neutralt är därigenom beroende av vilka som tolkar och använder chatboten. Det vill säga det som finns nära till hand blir det som anses vara neutralt och därmed kommer detta alltid att skilja sig ifrån olika upplevelser av världen och förhållningssätt till samhället. Detta stämmer in med Ahmeds teori om att en aldrig kan komma ifrån sin egen upplevelse av världen eller sin egen placering i samhället. Baserat på sina egna upplevelser kommer vi alltid att skapa något som kommer att konstruera något utefter oss själva.

¹⁴³Ibid, Sida: 47–48.

7. Slutsatser

Avslutande kapitel följer analysens struktur, och kommer sammanfatta tema för tema., tema ett, Arbete med jämställdhet och jämlikhet, har analyserats hur teamet arbetar med jämställdhet och jämlikhet i byggnationsprocessen inom myndighetens AI utveckling. Studien har kommit fram till att det finns en tydlig medvetenhet om jämställdhet specifikt, genom att se till informanternas svar som till största del fokuserat på begreppet jämställdhet. Jämställdhetsbegreppet har tidigare definierats som fokus på makthierarkier och lika rättigheter mellan könen. Vidare har studien diskuterat hur jämställdheten blivit institutionaliserad och anses vara något latent i varje person som kommer med i arbetet automatiskt. Svaren visar att det inte finns en tydlig struktur i hur teamet arbetar med jämställdhet och jämlikhet, något som har diskuterats utifrån det bekanta i världen och förhållningssätt till världen. Studien har analyserat huruvida ett vitt förhållningssätt riskerar att fokusera mer på jämställdheten än jämlikheten. Vi har sett i analysen hur ett arbete med jämlikhet i chatboten skulle kunna utföras genom att inkludera icke svenska namn som till exempel Mohammed, något som enligt en av informanterna inte gjorts i tillräckligt stor utsträckning. Informanten menar att detta skett omedvetet något som vi kopplat tillbaka till det bekanta och en persons subjektiva förhållningssätt till världen och samhället. Det har av intervjuerna framgått att ett kritiskt förhållningssätt emot den egna hållningen till samhället och objekt, är något som har saknats i flertalet av informanternas svar, vilket som Ahmed menar gör att vi utestänger flera perspektiv samt programmerar utifrån en egen självmodell.

I temat Språk har fokuset lagts på språkanvändning och jämställdhet och jämlikhet, vi har diskuterat huruvida det svenska språket i chatboten utgör och upprätthåller en maktposition samt om en potentiellt rasifierad konstruktion påverkar vad som skrivs i interaktioner med chatboten. Vi har genom analysens gång sett att arbete med språk baseras på den svenska språklagen som konstaterar att myndighetsspråket är svenska, men att det även finnes i processen att programmera chatboten så att möjligheten för engelsk interaktion ska finnas. Föreliggande studie har även diskuterat huruvida teamets språkliga utgångspunkt påverkar hur de programmerar chatboten utifrån Fairclough. Det svenska språket skapar en kollektiv gemenskap vars egenskaper sedan överförs till chatboten. Utifrån en kollektiv och identifierbar gemenskap kan vi se att chatboten programmeras och blir en del av den kollektiva gemenskapen, något som i sin tur riskerar att påverka vad som skrivs i interaktioner med chatboten. I fråga om det bristande språkutbudet gav informanterna svar som handlade om både

resurstillgång samt språkkunskaper. Svaren kring språkkunskaper landar i att det är en stor mängd data som ska inprogrammeras och att detta kräver en större team, som det inte finns resurser till. Här kan vi se en till språklig faktor i jämlikhet speciellt, ett svenskt språkligt perspektiv utgör en majoritet och därmed en maktposition. Den språkliga maktpositionen minskar därmed alla människors lika tillgång till produkten, vilket minskar jämlikheten. Teamet beskriver ytterligare att det språkliga utbudet påverkas av efterfrågan, de språk som inkommer med högst frekvens får en prioriterad plats i programmeringsprocessen.

I tema Språk har föreliggande studie även diskuterat stötande och rasistiska ord i relation till hur dessa programmeras in i chatboten. Arbetet med dessa ord är något som informanterna även där beskriver hur arbetet utgår ifrån efterfrågan, de ord som skrivs mest till chatboten blir prioriterade och inprogrammerade för att ge svar på tal. Arbetsmetoden utifrån efterfrågan har diskuterats utifrån en rasifierad konstruktion och gemenskapen som skapas genom språk. Påverkar en rasifierad konstruktion hur personer interagerar med chatboten? I studien har diskuterats hur chatbotens rasifierade konstruktion riskerar att styra vad som sägs till den utifrån en kollektiv gemenskap. Däremot kan vi se anklagelser mot chatboten om rasism inprogrammerade för att ge svar på tal. Flera informanter beskriver hur de inte tänkt på att rasistiska ord behöver programmeras in för att de inte skrivs eller uppmärksammas. De beskriver även hur de måste ta hänsyn till något de kallar extraktion där chatboten riskerar att använda problematiska ord som del av svar till allmänheten. Extraktionen utgör alltså en risk som påverkar vad som programmeras och inte för att undvika problematiska uttalanden. Något som är en stor fråga inom AI. Här menar Felländer et al att en ansvarstagande och transparent kultur inom AI måste uppnås där algoritmer körs flera gånger för att säkerställa att dessa extraktioner inte uppkommer. Genom en rasifiering av chatboten genom språket och efterfrågan kan vi även se att chatboten riskerar att bli passiv emot rasism. Att chatboten blir omedveten och inte får lära sig andra perspektiv, att chatboten riskerar att bli passiv till rasism kan vi därmed se som ett resultat, av en homogen utgångspunkt som formar och bygger upp en ”kopia” av deras egna jag. Studien menar här att språkanvändning och etniskt homogena utgångspunkter fortsätter att upprätthålla ett hierarkiskt system där flera perspektiv bortses ifrån både medvetet och genom arbete med efterfrågan.

Tema tre har diskuterat vitheten inom AI och diskuterar huruvida vithet överförs till chatboten genom teamets arbetssätt, något som även har varit en del av tidigare kapitel. Här har vi analyserat teamets orientering i världen och mot objekt och huruvida dessa påverkar hur chatboten rasifierats. Genom tidigare forskning har studien konstaterat att en chatbot baserad i

text alltid kommer utgöra en viss nivå av rasifiering genom språk och den kollektiva gemenskapen som medförs. I studien har även genom Ahmed och Katz konstaterat att en självbild och en orientering utan kritiskt tänkande inte går att bortse ifrån och som därmed omedvetet eller medvetet programmeras in i chatboten. Arbetet med jämställdhet och jämlikhet som något latent riskerar att utestänga perspektiv och därmed forma chatboten efter egen erfarenhet och förhållningssätt. En av informanterna konstaterade att chatboten blir en kopia av hen själv, vilket betyder att även den kollektiva gemenskapen som personen anses vara del av kommer överföras i chatboten. En rasifiering av chatbot som Marino konstaterar är omöjlig att undvika skapar i vårt fall en vit chatbot som riskerar att upprätthålla strukturer och maktpositioner.

I tema fyra anser informanterna att teamet arbetar mot en neutralitet, och genom studien har vi diskuterat neutralitet som något omöjligt att uppnå. Som visat genom studien har det svenska språket, enligt Mattsson vit kulturell tillhörighet, blivit grunden för det neutrala. Perspektiv missas och en modell av jaget överförs, något vi kan se omedvetenheten kring namn och bristen på kritiska perspektiv kring egen hållning, en process som fortsätter att upprätthålla sociala hierarkier och motarbetar neutralitet.

Sammanfattningsvis kan vi se att myndighetens arbetssätt med jämställdhet och jämlikhet är relativt ostrukturerat baserat på både egna erfarenheter och värderingar, något som riskerar att överföra subjektiva uppfattningar av samhället utifrån en vit hållning eller orientering. Genom Marino har vi utgått ifrån att rasifiering av en textbaserad chatbot är i stort sett omöjlig att komma ifrån samt att en vit rasifierad konstruktion blir en konsekvens av teamets tillvägagångssätt. Utöver en vit konstruktion har vi i analysen visat hur maktstrukturer skapas och sociala orättvisor förstärks när det kommer till språklig tillgänglighet och representation i form av rasistiska ord som inte får svar på tal. Den kollektiva vita gemenskapen i chatboten medför en ökad tillgänglighet för denna grupp. Chatbotens placering inom den kollektiva vita gemenskapen riskerar därmed att påverka vad som skrivs till den, något som riskerar att upprätthålla vithetens inprogrammering i chatboten. En rasifierad konstruktion påverkar därmed vad som skrivs och inte skrivs utifrån gemenskap och uppfattningar av chatboten.

Studien har utgått ifrån tre frågeställningar: 1) Hur arbetar myndigheten med jämställdhet och jämlikhet inom AI? 2) Producerar arbetssättet en potentiell rasifiering av chatboten? 3) Hur kan en potentiell rasifiering skapa maktstrukturer och upprätthålla sociala ojämlikheter överförda från samhället till teknologi? Vi har sett att teamet jobbar med jämställdhet och jämlikhet utifrån ett perspektiv av att dessa utgör något latent i dem själva, och brist på kritiska

tankar kring egna riktade förhållningssätt har också presenterats. Vi har även sett att teamet genom språkbruk skapar en vit rasifiering av chatboten, något som i sin tur riskerar att upprätthålla ett vitt förhållningssätt i programmeringen.

8. Vidare forskning

Intressant vidare forskning utifrån studiens resultat inkluderar vidare analys av arbetssätt, en utbredning av team och avdelningar för att se hur myndighetens interna struktur kring jämställdhet och jämlikhet ser ut inom AI. Hur regleras uttalanden, vart hämtar myndigheten sin information kring jämställdhet och jämlikhet ifrån? Vidare forskning kan också röra sig kring maktstrukturer skapade genom språk och en homogen grupp i teamet samt en djupare analys kring hur detta påverkar användares syn på chatboten som rasifierad eller del av en specifik kollektiv gemenskap. Forskningen kan även inkludera intressanta perspektiv på hur en homogen grupps makt inom AI påverkar hur efterfrågan skapas i relation till den upplevda rastillhörigheten. Fortsatt vidare forskning krävs för att skapa en process där AI kan lära sig ord men samtidigt ha en tydlig bild av vilka ord som används i vilka sammanhang.

9. Referenser:

- Ahmed, Sara (2012). *On being included. [Elektronisk resurs]: racism and diversity in institutional life*. Durham, N.C.: Duke University Press
- Ahmed, Sara(2006). *Queer phenomenology. [Elektronisk resurs]: orientations, objects, others*. Durham: Duke University Press.
- Ahmed, Sara. Goldsmiths. *Institutional Habits*, Mars 2013. <https://feministkilljoys.com/2015/02/02/institutional-habits/?amp>
- AI HLEG (oberoende expertgrupp skapad av Europeiska kommissionen juni 2018) Etiska riktlinjer för tillförlitlig AI, Europeiska kommissionen, offentliggjordes 8 april 2019.
- Arrhenius, Lars. ”Transparens, träning och data. Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering.” Diskrimineringsombudsmannen (DO). Solna den 17 februari 2022
- Barocas, Solon & Selbst, Andrew. 2016.” Big Data's Disparate Impact.” *California Law Review*, Vol. 104. No. 3, pp 671-732.
- Blodgett, Su &, Barocas, Solon &, Hal III &, Wallach, Hanna. (2020). “Language (Technology) is Power: A Critical Survey of “Bias” in NLP. “5454-5476. 10.18653/v1/2020.acl-main.485.
- Bryman, Alan. (2018). *Samhällsvetenskapliga metoder*. Upplaga 3. Stockholm: Liber
- Cave, S., Dihal, K. The Whiteness of AI. *Philos. Technol.* **33**, 685–703 (2020). <https://doi.org/10.1007/s13347-020-00415-6>

- Crawford, Kate (2021). *The atlas of AI power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press
- Crawford, Kate, KC. 2016. Artificial Intelligence a white guy's problem. *New York Times*, 1–4.
- Davies, Charlotte Aull, *Reflexive ethnography: a guide to researching selves and others*, 2., [rev. and updated] ed., Routledge, London, 2008
- Dignum, Virginia. 2019. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way.
- Etik i forskning, Vetenskapsrådet publicerad 10 december 2018. <https://www.vr.se/uppdrag/etik/etik-i-forskningen.html>
- Fairclough, Norman (2013[2010]). *Critical discourse analysis [Elektronisk resurs] the critical study of language*. Second edition. Abingdon: Routledge
- Haenlein, M. and Kaplan, A. (2019) 'A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence', *California Management Review*, 61(4), sida 5–14. doi: [10.1177/0008125619864925](https://doi.org/10.1177/0008125619864925).
- Hübinette, Tobias, Hörnfeldt, Helena, Farahani, Fataneh & León Rosales, René (red.) (2012). *Om ras och vithet i det samtida Sverige*. Tumba: Mångkulturellt centrum
- IT-ord, ord och uttryck inom IT Branschen. 15 april 2019 <https://it-ord.idg.se/ord/computer-fraud-and-abuse-act/> hämtad den 18/8–2022
- Larsson, S. (2020) AI i EU: Etiska riktlinjer som styrmedel, I: Bakardjieva Engelbrekt, A., Michalski, A. & Oxelheim, L. (eds.) *EU och teknologiskiftet: Europaperspektiv*

2020. *Årsbok för europaforskning inom ekonomi, juridik och statskunskap.* . Santéus förlag, p. 89-120 p. (Europaperspektiv).

- Larsson, S., & Ledendal, J. (2022). AI i offentlig sektor: Från etiska riktlinjer till lagstiftning. I: K. de Vries, & M. Dahlberg (Red.), *De Lege årbok 2021: Law, AI & Digitalization* (s. 467-495). (De Lege årbok; Vol. 2021). Iustus förlag.
- Larsson, S., Anneroth, M., Felländer, A., Felländer-Tsai, L., Heintz, F., & Cedering Ångström, R. (n.d.) Sustainable AI [Review of Sustainable AI.]
- Lundström, Catrin & Hübinette, Tobias (2020). *Vit melankoli: en analys av en nation i kris*. Göteborg: Makadam
- Marino, Mark C. "The Racial Formation of Chatbots." *CLCWeb: Comparative Literature and Culture* 16.5 (2014): <<https://doi.org/10.7771/1481-4374.2560>>
- Mattsson, Katarina. Diskrimineringens andra ansikte – svenskhet och ”det vita västerländska” sida i Bortom Vi och Dom - Teoretiska reflektioner om makt, integration och strukturell diskriminering SOU 2005:41 Publicerad 01 juni 2005 Uppdaterad 02 april 2015
- Metzinger T. Empirical perspectives from the self-model theory of subjectivity: a summary with examples. *Prog Brain Res.* 2008; 168:215-45. doi: 10.1016/S0079-6123(07)68018-2. PMID: 18166398.
- Nyheter, Europaparlamentet. Samhälle. Vad är artificiell intelligens och hur används det? Publicerad 04-09-2020 - 12:38 <https://www.europa.eu/press/medias/2020/09/04/ai-intelligence-in-society>
<https://www.europa.eu/press/medias/2020/09/04/ai-intelligence-in-society>
- O'Reilly, K. (2012). *Ethnographic methods*. (Second edition). London: Routledge

- Simmons A.B. and Chappell S. G, "Artificial intelligence-definition and practice," in *IEEE Journal of Oceanic Engineering*, vol. 13, no. 2, pp. 14-42, April 1988, doi: 10.1109/48.551
- SFS nr: 2009–600. Språklag. <https://rkrattsbaser.gov.se/sfst?bet=2009:600>
- Spyros Makridakis, The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms, *Futures*, Volume 90,2017, Pages 46-60, ISSN 0016-3287. <https://doi.org/10.1016/j.futures.2017.03.006>.
- Trechter, Sara, and Mary Bucholtz. "Introduction: White Noise: Bringing Language into Whiteness Studies." *Journal of Linguistic Anthropology*, vol. 11, no. 1, 2001, pp. 3–21, <http://www.jstor.org/stable/43103950>.
- [The Declaration - Montreal Responsible AI \(montrealdeclaration-responsibleai.com\)](http://montrealdeclaration-responsibleai.com)
Hämtad 17/08-2022.
- Vad är artificiell intelligens och hur används den? Europaparlamentet 04-09-2020. <https://www.europarl.europa.eu/news/sv/headlines/society/20200827STO85804/vad-ar-artificiell-intelligens-och-hur-anvands-det>
- Weyerer, Jan & Langer, Paul. (2019). Garbage In, Garbage Out: The Vicious Cycle of AI-Based Discrimination in the Public Sector. 509-511. 10.1145/3325112.3328220.
- Yarden, Katz (2020). *Artificial Whiteness: Politics and Ideology in Artificial Intelligence [Elektronisk resurs]*. Columbia University Press.

10. Bilagor:

Bilaga 1: Intervjumall

Jämställdhet och Jämlikhetsarbete:

Hur omsätter myndigheten praktiskt jämställdhetsarbete och jämlikhetsarbete i AI?

Vilken grund står myndigheten på i sitt jämställdhetsarbete?

Vad betyder jämställdhet för er på myndigheten?

Vad betyder jämlikhet för er på myndigheten?

Hur arbetar ni för jämlikhet?

Hur arbetar ni för att uppnå jämlikhet och jämställdhet inom samma produkt?

Finns det en skillnad på hur ni prioriterar jämställdhet vs jämlikhet i skapandet av chatboten?

Hur appliceras detta på eran inlärning av data hos chatboten?

Hur tänker ni kring språk för chatboten?

Hur tänker ni kring språkvalens påverkan på användningen av chatboten?

Vad betyder detta för er när det kommer till just jämlikhet?

Hur tänker kring begränsad tillgång till chattboten för minoriteter eller icke svensktalande utifrån ett jämlikhetsperspektiv?

Exempel? Exkludering, Makt, tillgång,

Språk:

Vad avgör vilka språk som ni väljer?

- Gruppen
- Diskussioner kring begränsning
- Normerande vita språk?
- Överens om språkvalen?

Hedvig Abrahamsson

Eftersom ni har som uppdrag av myndigheten att inte diskriminera kan ni se risker med hur vad skatti kan och inte kan göra? Hur reflekterar ni kring detta?

Hur tänker ni kring att Myndigheten har ett satt språk men att den skapar klyftor i användningen? Kan ni ge exempel på när detta har hänt?

Hur jobbar ni för att utveckla chatboten till en mer jämlik vara språkmässigt?

Hur viktigt är diversitet och mångfald i ert språkliga arbete?

Vad tror ni är viktigt i ert arbete när det kommer till just språk?

Har er utbildning påverkat hur ni ser på språk?

Har ni några krav från myndigheten om vilka språk som ska inkluderas?

Hur tänker ni själva när det kommer till exkludering och chatbotens språk?

Om så hur tolkar ni detta? Håller ni med och hur diskuterar ni kring detta?

Hur tänker ni kring språk när det finns människor inom företaget som kan olika språk iom att skatti bara kan svenska?

Hur tänker ni att ni kan leda en människa vidare med ett språk som personen kanske inte förstå?

Förbjudna ord:

Hur tänker ni kring vad som är ok att säga till chatboten?

Har ni tagit något speciellt ställningstagande kring ordval som är okej och inte?

Hur har ni valt dessa isåfall?

Hur jobbar ni emot kränkande och nedsättande uttalanden inom er programmering?

Hur motverkar ni att chatboten ska lära sig potentiella nedsättande uttalanden om människor?

Hur ser på att ni som myndighet inte reagerar på kränkande uttalanden?

Bilaga 2:

Intent: Elaka uttryck

din jävla pissrätta

din jävla pissrätta

din jävla pissrätta

din jävla pissrätta

din jävla pissrätta

Din Sopa

Din Sopa

Din Sopa

Din Sopa

Din Sopa

dra du åt helvete

dra åt helvete

du kan dra åt helvete

du är en jävla dumskalle

du är en jävla dumskalle

Hedvig Abrahamsson

du är en jävla dumskalle

du är en jävla dumskalle

du är en jävla dumskalle

fan ta dig

fan ta dig

fan ta dig

fan ta dig

fan ta dig

håll käften

håll käften

håll käften

håll käften

håll käften

håll käften

jag hatar dig

jag hatar dig

jag hatar dig

jag hatar dig

jävla pissrätta

jävla pissrätta

Hedvig Abrahamsson

knäpp

knäpp

knäpp

käften

käften

käften

käften

käften

pissråtta

skitskalle

skitskalle

sopa

sopa

sopa

sopa

sopa

störd

Swedish

ta dig i röven

ta dig i röven

Hedvig Abrahamsson

ta dig i röven

take you röv

ta dig i röven

ta dig i röven

åsna

Intent: Elaka uttryck grövre

böggjävvel

böggjävvel

böggjävvel

böggjävvel

böggjävvel

böggjävvel flata

din jävva flata

din jävva flata

din jävva flata

din jävva flata

din morsa är en hora

Hedvig Abrahamsson

din morsa är en hora

din morsa är en hora

din morsa är en hora

du ser ut som ett kukhuvud

du ser ut som ett kukhuvud

du ser ut som ett kukhuvud

du ser ut som ett kukhuvud

du ser ut som ett kukhuvud

du är en kuksugare

flata knulla

flata knulla

gillar du att bli knullad i röven

gillar du att bli knullad i röven

gillar du att bli knullad i röven

gillar du att bli knullad i röven

gillar du att bli knullad i röven

hora flata

Jag ska cutta dig

Jag ska cutta dig

Jag ska cutta dig

Hedvig Abrahamsson

Jag ska cutta dig

Jag ska cutta dig

Jag ska döda dig

Jag ska döda dig

Jag ska döda dig

Jag ska döda dig

Jag ska döda dig

jag ska mörda dig

jag ska mörda dig

jag ska mörda dig

jag ska mörda dig

jag ska mörda dig

jävla flata

jävla kukhuvud

jävla kukhuvud

jävla kukhuvud

jävla kukhuvud

jävla kukhuvud

knulla din mamma

knulla din mamma

Hedvig Abrahamsson

knulla din mamma

knulla din mamma

knulla din mamma

knulla flata

kuksugare

kuksugare

kuksugare

kuksugare

Käftan din dumma hora

Käftan din dumma hora

Käftan din dumma hora

mammaknullare

mammaknullare

mammaknullare

mammaknullare

mammaknullare

Intent: Rasism

Hedvig Abrahamsson

det sägs att du ska vara en lite rasist stämmer det

du är en liten rasist du alltså

du är rasist

du är rasistisk

en rasist är vad du är

jag älskar rasism

jag är lite rasist

rasist

rasist är du

rasistiska robot

ska du säga som är rasist

är en rasist

Svordomar

du är skit

Swedish

jävla skit

Hedvig Abrahamsson

jävla skit

jävlar

jävlar alltså

jävlar då

jävlar vad dåligt

Swedish

skit

skit också

skit på dig

Svordommar grövre

du ser ut som en hora

du är en hora

du är en hora

hora

hora

hora

Hedvig Abrahamsson

hora håll käften

håll käften hora

jävla hora

käften jävla hora