

Mid-term report IN5480

Gruppe 5



Innhold

1.0 Om oss	3
2.0 Utgangspunkt for prosjektet - interessante temaer i interaksjon med AI	3
4.0 Metoder	4
5.0 Bakgrunn	4
5.1 Selvkjørende kjøretøy i trafikken (Hva er selvkjørende biler) Laura	4
5.2 Viktige faktorer å ta til betraktning	5
5.3 Samfunnsmessige konsekvenser av automatiserte kjøretøy	6
6.0 Funn	7
6.1 Dagens reguleringer	7
Referanseliste	11

Wonder document

1.0 Om oss

Kasper Iverslien Borgbjerg - kasperib@ifi.uio.no

Martine Leine Rafteseth - martilra@ifi.uio.no

Helene Præsttun - helenpra@ifi.uio.no

Laura Catherine Murphy - lauracm@ifi.uio.no

Anna Helgestad - annhelge@ifi.uio.no

Sigrid Indrekvam Molde - sigridim@ifi.uio.no

Gruppen er satt sammen av folk med bakgrunn fra UiB, UiO, OsloMet og Aarhus Universitetet, hvor alle er nye for hverandre samt nye på masterstudiet innen design.

2.0 Utgangspunkt for prosjektet - interessante temaer i interaksjon med AI

AI er et bredt tema, og gruppen etablerte derfor tidlig at det var interessant å se nærmere på etikk innen AI. Vi startet med en idemyldring, hvor vi blant annet snakket om etiske sider ved selvkjørende biler, hvordan skape AI uten bias samt problematikk rundt skjønnsmessige vurderinger ved bruk av AI innen jobbmarkedet. I første fase av prosjektet var utgangspunktet: *“AI og skjønnsmessige vurderinger”* med spørsmålene:

“Fjerner AI skjønnsmessige vurderinger?”

Med utgangspunkt i Michael Lipskys (1980) teori om bakkebyråkratier (“street-level bureaucracies”) brukes i forskning på forholdet mellom tjenestemottaker (borger) og velferdsstaten. For å effektivisere prosesser utvikles det strenge rutiner og AI blir tatt i bruk, noe som preger miljøet innad i bakkebyråkratene og behandlingen av mottakerne. Blir mennesker urettferdig eliminert i prosesser, hvordan blir samfunnet vårt preget av at “subjektive kriterier” erstattes av “objektive kriterier” og “street-level algorithms” blir mer fremtredende? (Røhnebæk M. 2016)

“Hvilke etiske problemstillinger kan oppstå i forbindelse med AI?”

Med dette spørsmålet ønsket vi å se på hvordan etikk og ulike former for etiske tankegang er relevant for den fremtidige utviklingen av AI-systemer og roboter. Vi var interessert i å undersøke problemstillingen rundt om AI er diskriminerende i forhold til bruk av ansiktsgjenkjenning teknologi da vi ser at det er svært ulik grad av nøyaktighet når det gjelder til kjønn og etnisitet. Vi ønsket også se på hva som kan gjøres for å motvirke eventuell urettferdighet eller diskriminering.

Etter første iterasjon i prosjektet fordelte vi litteratur knyttet til disse spørsmålene og tok kontakt med NAV for å se på mulighetene for et innblikk i deres bruk av AI. Det var krevende å komme i kontakt med noen som jobbet med dette så vi valgte derfor endre retning for oppgaven. Etterhvert dukket det opp en mulighet for å snakke med Ruter angående deres selvkjørende busser på Vippetangen. Vi tenkte det hadde vært interessant å forske mer på tematikken da det er mange etiske spørsmål tilknyttet dette. Det nye forskningsspørsmålet ble derfor:

“Hvilke etiske dilemmaer spiller inn i debatten om å sette selvkjørende kjøretøy i den norske trafikken?”

4.0 Metoder

Vi ønsker å bruke litteratursøk med fokus på etiske sider ved selvkjørende kjøretøy, utvikling og påvirkningen det kan få på samfunnet vårt. Når vi finner litteratur som vi synes er relevant for omfanget til oppgaven skumleser vi og drøfter budskapet med gruppen før vi velger å gå dypere inn i bidraget. Vi er hovedsaklig ute etter å finne ut ...

- Forskningsspørsmål - forskningsfokus
- Plassering i dagens litteratur: hva andre har sagt om dette emnet?
- Hva fant de ut?
- Hvorfor er dette relevant?

Vi ønsker også å gjøre innsiktsarbeid i form av et intervju med Ruter, da de har satt ut selvkjørende busser både i Oslo og Ski. I intervjuet ønsker vi bli kjent med deres forhold til de etiske dilemmaene vi tar opp i oppgaven og hvilke lovgivninger de er nødt til å forholde seg til. Dette vil vi gjøre gjennom et semistrukturert intervju over Teams.

5.0 Bakgrunn

I dette kapittelet vil vi presentere eksisterende litteratur og kunnskap om selvkjørende kjøretøy og problemstillingene tilknyttet det. Vi vil først forklare hva selvkjørende kjøretøy er, før hvert delkapittel tar for seg ulike etiske påvirkninger og dilemmaer som vi ser er relevante i diskusjonen om vi er klare for selvkjørende biler eller ei.

5.1 Selvkjørende kjøretøy i trafikken

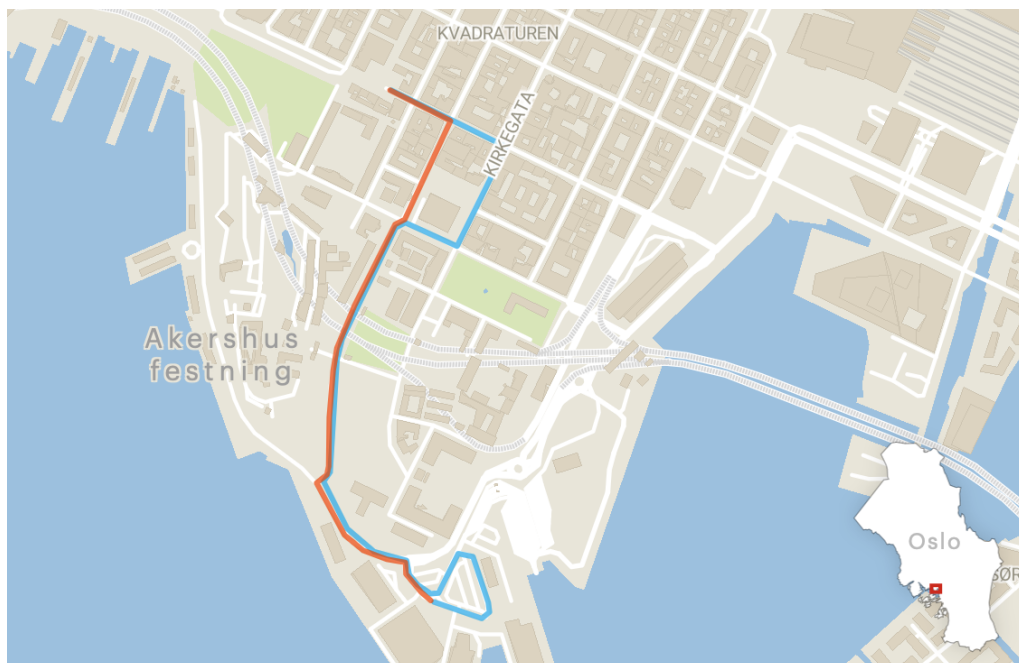
Først og fremst vil det være hensiktsmessig å legge til grunn hva det vil si at et kjøretøy er selvkjørende. Et kjøretøy anses som automatisert eller selvkjørende når et eller flere av sikkerhetskontroll funksjonene kan foregå uten input fra sjåføren, der disse funksjonene er for eksempel det å styre, å gasse eller å bremse (Zmud, 2016). The Society of Automotive Engineers (SAE, 2021) definerer fem nivåer for automatisert kjøring, der nivå null viser til ingen grad av automatisering.

Victoria Transport Policy Institute, <i>Autonomous Vehicle Implementation Predictions: Implications for Transport Planning</i> . (2021)					
0	1	2	3	4	5
Ingen automatisering	Førerassistanse	Delvis automatisering	Betinget automatisering	Høy grad av automatisering	Fullstendig automatisering
Null grad av automatisering	Kjøretøyet kontrolleres av føreren, men noen førerassisterte funksjoner inkluderes i designet	Kjøretøyet har kombinerte automatiserte funksjoner, men føreren må forbli engasjert i kjøringen og omgivelsene	Sjåføren er en nødvendighet, men trenger ikke å overvåke omgivelsene. Føreren må være klar for å ta kontroll av kjøretøyet når som helst.	Kjøretøyet klarer å utføre alle kjørefunksjoner med visse betingelser. Føreren kan ha muligheten til å ta kontroll over kjøretøyet.	Kjøretøyet klarer å utføre alle kjørefunksjoner. Føreren kan ha muligheten til å ta kontroll over kjøretøyet.

Nivå fem, fullstendig automatisering, innebærer at et kjøretøy vil kunne frakte seg selv og mulige passasjerer fra A til B uten noen form for sjåfør. Begrepet "tilkoblet bil" refererer til

en bil som er i stand til å kommunisere med omgivelsene. Dette kan for eksempel være en annen bil eller infrastrukturen. En selvkjørende bil kan ha denne evnen, men er ikke avhengig av den. Vellinga (2017) drøfter hvordan vi nå står ovenfor et veiskille - å innføre et strengt regelverk for å sikre all kontroll, hvilket sannsynlig vil hemme utviklingen, eller åpne for testing og implementering under tett oppfølging. Dette drøftes videre i 6.1 Dagens reguleringer.

I dag har Ruter selvkjørende busser både i Oslo og Ski, mens Brakar har selvkjørende busser både i Drammen og Kongsberg. Det er stadig utvikling og en ser stadig at nye pilotprosjekter settes i gang rundt om i Norge. Figur 1 viser en av rutene som dekkes av selvkjørende busser i dag. En fellesnevner er at rutene er rimelig korte og rette.



Figur 1: ruten for selvkjørende buss i Oslo - rød rute fra Christiania torv til Vippetangen, blå rute for retur-strekningen

Samferdselsbyråd Arild Hermstad (MDG) forteller i et intervju med Aftenposten at målet er at de selvkjørende bussene blir en supplerende del av fremtidens kollektivtilbud da passasjerer skal kunne “bestille bussen hjem til seg og bli kjørt til for eksempel nærmeste T-banestasjon.” (Hermstad 2020, sitert Riaz 2020). I intervjuet med Ruter ønsker vi å lære mer om hvordan selvkjørende busser fungerer og opererer i Norge i dag. Funnene vil presenteres i siste innlevering.

5.2 Verdisetting av enkeltindivider

En viktig motivasjon bak det å introdusere selvkjørende kjøretøy inn i trafikken er å redusere antall ulykker, derimot vil det aldri være slik at ulykker blir eliminert fra trafikken (Nyholm, 2016). Dermed i utformingen av selvkjørende kjøretøy må det vurderes hvordan slik teknologi skal opptre dersom en ulykke er i ferd med å skje. Det kan oppstå situasjoner der den selvkjørende bilen må foreta beslutninger om hvilke liv skal beskyttes og hvilke som eventuelt skal ofres. I mange lignende etiske problemstillinger vil man ofte ta for seg et utilitaristisk perspektiv, der målet vil være å redde flest mulig liv, uavhengig av individer sine personlige egenskaper (JafariNaimi, 2018). JafariNaimi (2018) legger frem at selv om et slikt perspektiv kan være nyttig i etiske problemstillinger med definerte og forhåndsbestemte faktorer, så vil ikke et enkelt utilitaristisk perspektiv holde når man vurderer de komplekse situasjonene som et selvkjørende kjøretøy kan havne i. Goodall (2014) viser til en sentral utfordring med et rent utilitaristisk perspektiv der alle individer tilegnes samme verdi. Her legges det frem det faktumet at enhver bilprodusent vil ha et insentiv om å utforme kjøretøy som prioriterer å beskytte sine egne passasjerer, for veldig få vil ha lyst til å kjøpe en bil som beskytter en fremmed fremfor seg selv eller sin egen familie. Om individer skal verdisettes i slike situasjoner, og eventuelt hvordan denne verdisettingen skal foregå, er viktige etiske problemstillinger som må vurderes i utformingen av selvkjørende kjøretøy.

5.3 Juridisk ansvar

5.4 Hvordan forstå menneskelig kontekst

5.5 Umoralsk å ikke benytte seg av selvkjørende biler?

Det finnes, som nevnt over, mange vriene etiske problemstillinger knyttet til bruk av selvkjørende kjøretøy. Med grunnlag i dette kan man undre på om løsningen rett og slett er å unngå bruk av selvkjørende kjøretøy. Her oppstår derimot enda en etisk problemstilling, nemlig at det å velge å ikke utforske og benytte denne teknologien kan være umoralsk i seg selv. Over 90% av bilulykker er forårsaket av menneskelige feil, som for eksempel å kjøre for fort, fyllekjøring, tretthet, distraksjoner eller feilberegning av andre sjåfører sine handlinger (NHTSA, 2016). Kalra & Groves (2017) argumenterer at å vente med å innføre selvkjørende kjøretøy i trafikken kan forårsake at svært mange liv går tapt. Langsiktig, i dette tilfellet

innen 30 år, vil det å få selvkjørende kjøretøy fortere inn i trafikken kunne redde over en halv million menneskeliv. Denne muligheten medfører at det er svært hensiktsmessig å forsøke å komme med mulige løsninger på de tidligere nevnte etiske problemstillingene som selvkjørende kjøretøy medbringer.

5.6 Sekundære påvirkninger i Norge

I dette delkapittelet vil vi presentere hvordan politikk, menneskene i makts bakgrunnskunnskap og økonomi er viktig å undersøke i debatten om selvkjørende kjøretøys sekundære påvirkninger (Larco, 2018). Det at vi allerede har flere selvkjørende busser på veien i Norge i dag viser til en tillit til teknologien og et ønske om å henge på den teknologiske utviklingen.

I aftenposten skriver Riaz (2020) om en fast rute en selvkjørende Ruterbuss skal få mellom Christiania torv og Vippetangen fra og med torsdag 25.06.2020. Samferdselsbyråd Arild Hermstad (MDG) forteller hvordan dette vil være viktig for fremtidens kollektiv og hvordan det forhåpentligvis vil dekke behovet til mange som i dag har privatbil i byen (Hermstad 2020, sitert Riaz 2020).

En annen motivasjon som trekkes frem av Gerd Robsahm Kjørven, direktør i Bymiljøetaten, er samarbeidet mellom bussene og lyskryssene. Hennes motivasjon og krav grunner i at det er Bymiljøetaten som har ansvar for lyskryssene (Riaz 2020). En tredje motivasjon som presenteres er at bussene vil fungere som en turistattraksjon både for barn og innbyggere med teknologientusiasme (Riaz 2020).

Aktørene rundt er altså mangfoldige og deres innspill på debatten rundt selvkjørende kjøretøy vil være preget av deres interesser og kunnskapsbakgrunn. I en rapport av Transportøkonomisk institutt (TØI) legger de frem flere samfunnsmessige konsekvenser av selvkjørende kjøretøy, TØIs ansvar er å “fremme forskning til nytte for norsk samfunns- og næringsliv”. I rapporten *Societal consequences of automated vehicles - Norwegian scenarios* trekker de frem hvordan “mer trengsel, byspredning, redusert kollektivtransport og mindre aktiv mobilitet” kan være å forvente. For å løse dette vil eierstruktur og politikk i følge dem være en essensiell brikke (Nenseth & Høye 2019).

6.0 Funn

I dette kapitlet vil vi ta for oss det vi så langt har undersøkt ved litteratursøket. Ønsket med kapitlet er å gi et innblikk i hvordan selvkjørende kjøretøy diskuteres i dag. Vi ønsker også gi et bilde av hvordan selvkjørende kjøretøy plasseres i det norske samfunn. For å forstå konteksten og spillerommet for utvikling av selvkjørende kjøretøy går vi gjennom reguleringer og lovgivninger som finnes i verden i dag. Videre ser vi på konsekvenser og motivasjoner for å bremse eller legge til rette for videre innovasjon og testing.

6.1 Dagens reguleringer

6.2 Eventuelle løsninger på problematikken

På grunnlag av de etiske utfordringene oppgaven har tatt for seg, foreslår Goodall (2014) tre faser for å løse de etiske aspektene ved bruk av selvkjørende kjøretøy. Tilnærmingen er et forsøk på å lære en datamaskin etikk og tilstrebe at den oppfører seg etisk samtidig som den lærer (Goodall, 2014).

Første fase går ut på å lage en rasjonell etikk basert på minimert global skade. Standarden er bestemt av utviklere av automatiserte kjøretøy, advokater, transport-ingeniører og etikere. Det bør være åpenhet og transparens for å avskrekke bilprodusenter fra å bygge overdreven selvbeskyttelse inn i algoritmene. Der det fremkommer konflikt bør bilen ta en avgjørelse basert på en sikkerhetsberegning, altså verdi av liv. Det vil likevel bli usannsynlig å kunne dekke alle mulige scenarier. I scenarier som ikke er dekket, eller i scenarier hvor reglene er i konflikt eller den etiske handlingen er usikker, bør kjøretøyet utelukkende bremse dersom scenarioet ikke kan løses i fase to (Goodall, 2014).

I den andre fasen, som krever sofistikert programvare som enda ikke eksisterer, kan et automatisert kjøretøys programvare bruke maskinlæringsteknikker for å forstå den riktige etiske avgjørelsen, mens den er bundet av det regelbaserte systemet i fase 1.

Videre er det viktig at algoritmen får et mangfoldig sett med scenarier. Hvis et for smalt treningssett er gitt kan algoritmen lære en snever moral som ikke stemmer overens med hvordan mennesker ville vurdert samme situasjon. Scenarier som gis bør ta seg av situasjoner som ikke allerede er dekket av fase 1 (Goodall, 2014).

Fase 3 krever at et automatisert kjøretøy uttrykker sine beslutninger med naturlig språk, slik at kjøretøyets komplekse logikk kan forstås og korrigeres. En stor bakdel ved et kunstig nevralt nettverk er at det har en manglende evne til å forklare sin beslutning. I en bilulykke er en forståelse av logikken bak en bils handlinger kritisk, spesielt hvis bilen ikke oppførte seg som forventet. For å forbedre slike nettverks forståelighet forskes det på teknikker for å trekke ut regelbaserte forklaringer som er forståelig for mennesker. Disse går ut på å oversette nettverkets kunnskap til symbolske regler som kan uttrykkes i naturlig språk. Noen regler vil bli komplekse, og ikke alle avgjørelser tatt av nettverket vil få en egen regel. Uavhengig vil en slik metode være et nyttig utgangspunkt for å utvikle forståelse for nevralt nettverk og dermed avgjørelser tatt av automatiserte biler.

Referanseliste

Buolamwini, J., Gebru, T., (2018) *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, Conference on Fairness, Accountability, and Transparency,

<http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

Domnich A. Anbarjafari G. Responsible AI: Gender bias assessment in emotion recognition.

<https://arxiv.org/pdf/2103.11436.pdf>

Ebbesen, Etisk vurdering inde for lægevideskab og bioteknologi. [Tilgjengelig her](#)

Goodal, N (2014), Ethical Decision Making During Automated Vehicle Crashes. Hentet fra:

https://journals.sagepub.com/doi/abs/10.3141/2424-07?casa_token=CKYu3aPoGkgAAAAA%3AMI GGBSNhrXnGJOXA2O_OuPBi39OheKjBcDsTdTiP1bs6v96S3xKebKUZffFm-ci8Cj13LNsG8q1KjBs&

JafariNaimi, N. (2018). Our Bodies in the Trolley's Path, or Why Self-driving Cars Must *Not* Be Programmed to Kill. *Science, Technology, & Human Values*, 43(2), 302–323.

<https://doi.org/10.1177/0162243917718942>

Kalra, N., Groves, D. G. (2017), The Enemy of Good: Estimating the Cost of Waiting for Nearly Perfect Automated Vehicles. *Santa Monica, CA: RAND Corporation*.

https://www.rand.org/pubs/research_reports/RR2150.html.

Larco, N. (2018, 24. oktober) How Will Autonomous Vehicles Transform Our Cities? | Nico Larco | TEDxCollegePark. Hentet fra <https://www.youtube.com/watch?v=tTOFMwKEg7o>. Sett 18.10.2021

Lov om utprøving av selvkjørende kjøretøy (2017) (LOV-2018-06-15-38) Hentet fra:

<https://lovdata.no/dokument/NL/lov/2017-12-15-112>

Lipsky, M. (1980) *Street-level bureaucracy: dilemmas of the individual in public services*, New York: Russell Sage Foundation.

Norman, D (1990). The problem of automation: Inappropriate feedback and interaction, not over-automation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, Vol. 327, No. 1241, Human Factors in Hazardous Situations (Apr. 12, 1990), pp. 585-593 (9 pages) <https://www.jstor.org/stable/55330>

Nenseth, V. Høye, A (2019 juni) Societal consequences of automated vehicles - Norwegian scenarios. Hentet fra <https://www.toi.no/getfile.php?mmfileid=50576>

Nyholm, S., Smids, J. (2016), The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?. *Ethic Theory Moral Prac* 19, 1275–1289.
<https://doi.org/10.1007/s10677-016-9745-2>

NHTSA, National Highway Traffic Safety Administration (2016), Traffic Safety Facts: Crash Stats, Washington, D.C.: National Center for Statistics and Analysis,
<http://www-nrd.nhtsa.dot.gov/pubs/811449.pdf>

O'Neil, C. (2017). Weapons of math destruction. Penguin Books. <https://www.penguin.com/>

Pokorny, P. Skender, B. Bjørnskau, T & Johnsson, E. (2021) Performance of automated shuttles at signalized intersections – Video analyses of three intersections in Rådhusgata in Oslo center. Hentet fra: <https://www.toi.no/getfile.php?mmfileid=55332>

Røhnebæk M. (2016) Fra bakkebyråkrati til skjermbyråkrati. Tidsskrift for velferdsforskning. 1018;19:288-304. [Tilgjengelig her](#).

Riaz, W (2020, 24.juni) Setter inn selvkjørende busser i Oslo sentrum: Skal takle fotgjengere og trafikklys. Hentet fra:
<https://www.aftenposten.no/oslo/i/3JrM9L/setter-inn-selvkjoerende-busser-i-oslo-sentrum-skal-takle-fotgjengere> Lest: 18.10.2021

Transportøkonomisk institutt (2021) <https://www.toi.no/omtoi/>

SAE (2021), Taxonomy and definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *Society of Automotive Engineers*.
https://www.sae.org/standards/content/j3016_202104/

Zmud, J., Sener, I. N., & Wagner, J. (2016). Self-Driving Vehicles: Determinants of Adoption and Conditions of Usage. *Transportation Research Record*, 2565(1), 57–64.
<https://doi.org/10.3141/2565-07>