



Norsk
Utenrikspolitisk
Institutt

Meningsfull, menneskelig, kontroll?

Erik Reichborn-Kjennerud



NUPI Rapport
[6 / 2020]

Utgiver: Norsk Utenrikspolitisk Institutt 2020
Copyright: © Norsk Utenrikspolitisk Institutt 2020
ISSN: 1894-650X

Alle synspunkter står for forfatternes regning. De må ikke tolkes som uttrykk for oppfatninger som kan tillegges Norsk Utenrikspolitisk Institutt. Artikkene kan ikke reproduseres – helt eller delvis – ved trykking, fotokopiering eller på annen måte uten tillatelse fra forfatterne.

Besøksadresse: C.J. Hambros plass 2d
Postadresse: Postboks 7024 St. Olavs Plass
0130 Oslo, Norway
Internett: www.nupi.no
E-post: post@nupi.no
Fax: [+ 47] 22 99 40 50
Tel: [+ 47] 22 99 40 00

Meningsfull, menneskelig, kontroll?

Erik Reichborn-Kjennerud

Publisert av Norsk Utenrikspolitisk Institutt

Innhold

Innhold.....	3
Oppsummering	4
Introduksjon	6
Spørsmålet om «menneskelig» kontroll.....	9
Kentaureer og andre myter	13
Targeting og menneske-maskin relasjoner.....	16
Targeting, kontroll og fremtidens autonome systemer	19
Menneske-maskin og militære visjoner	23
Bibliografi	29

Oppsummering

Denne rapporten tar for seg konseptet «meningsfull menneskelig kontroll» som har dukket opp som det førende temaet i debatten rundt autonomi og våpensystemer. For å belyse spørsmålet om hva menneskelig kontroll er, hvordan det bør forstås og hvordan man kan «sikre» det må vi tenke nytt rundt forholdet mellom mennesker og maskiner. Ikke bare se maskiner som verktøy som mennesker kan bruke av fri vilje, men snarere se på autonome systemer i sin helhet der de gjensidige relasjonene mellom mennesker og maskin former hverandre og gir opphav til nye måter å tenke og handle på. Altså at både mennesker og maskiner forandres i møte mellom dem. Gitt en slik vinkling åpner man også for muligheten for å belyse dagens debatt om autonomi og kontroll på nye måter. Ikke minst at den utøvende fasen av en operasjon, liv og død-avgjørelser, er et for snevert nedslagsfelt hvis vi skal forstå hva kontroll er og hvordan det utøves. Viktigheten av dette blir illustrert via en gjennomgang av prosessene og praksisene rundt militær targeting. En prosess der militæret definerer mål og beslutter operasjoner. Gjennom denne illustrasjonen ser man at beslutningstaking ikke bare er *ett* valg gjort på et bestemt tidspunkt, som avgjørelsen om å sende et missil mot et bestemt mål, men noe som er avledet av en rekke valg gjort over tid. Dette åpner opp for å se på beslutningstaking og dermed også kontroll som et resultat av en rekke praksiser og prosesser, distribuert over flere faser, mellom diverse elementer, der noen valg blir foretrukket fremfor andre. Et slikt fokus setter søkelys på de mer sammensatte og komplekse teknologiske systemene for datainnhenting, analyse og lagring, som spiller en nøkkelrolle i utformingen av valg og beslutninger, men som ofte blir neglisjert. Hvis vi bedre forstår hvordan disse systemene fungerer, hvordan de produserer kunnskap, hva de legger vekt på og hva de utelater, kan vi øke vår forståelse av kontroll og bedre «sikre» de etiske, legale og strategisk-politiske sider ved økende autonomisering av militære teknologier.

Introduksjon

Det er godt kjent at teknologier kan ha betydelige implikasjoner for hvordan relasjoner mellom stater er strukturert, sikkerhet forstått og opprettholdt, hvordan konflikt og fred blir forvaltet, politisk handling blir muliggjort og ikke minst hvordan kriger utkjempes. Derimot er det er åpnet spørsmål om nøyaktig hvordan teknologier påvirker, former og endrer relasjoner og handlinger. I den pågående debatten rundt sammenhengen mellom krig og teknologi dreier det store spørsmålet i dag seg om hvorvidt teknologier kan bli gjort autonome – at de kan handle uten menneskelig overvåking eller intervensjon – og hvis så, hvilken innvirkning dette vil ha. Mye av diskusjonene handler om brede og til dels luftige spekulasjoner om hvordan fremtidige kunstig intelligente og autonome systemer vil påvirke alt fra geopolitisk rivalisering og sikkerhetspolitikk til krigføring. Det blir ofte argumentert for at disse nye teknologiene vil endre alt; fra taktikk og operasjoner til strategi, etablerte sikkerhets- og forsvarspolitiske konsepter som avskrekking og beroligelse, offensiv og defensive balanse, og styrkesammensetning og organisering. I denne debatten er spørsmål om menneskelig kontroll over såkalte dødelige autonome våpensystemer (Lethal Autonomous Weapon Systems, LAWS) sentralt. Skal eller ikke skal (om mulig) man gi disse systemene kapasiteten til å handle uavhengig av mennesker.

Den store bekymringen er at mennesker skal miste kontroll over de våpnene de bruker og dermed hverken ha råderett over liv og død eller kontroll over handlingsforløp. Dette er ikke bare redsler etikere, jurister og menneskerettighetsaktivister har, men de deles i økende grad av forskere, militærpersonell, så vel som politikere (Williams & Scharre, 2016). I denne debatten har begrepet «meningsfull menneskelig kontroll» oppstått. Det er imidlertid langt fra opplagt hva begrepet betyr og innebærer. Ikke bare er det uklart hvordan autonomi skal defineres, men de ulike ordene meningsfull, menneskelig og kontroll er også begreper som defineres ulikt av forskjellige aktører enten man er ingeniør eller systemdesigner, jurist eller etiker, operatør eller politiker. Disse ulike synene på hva begrepet meningsfull menneskelig kontroll

inneholder gir dermed også opphav til vidt forskjellige meninger og vekting. Hva som er meningsfullt for én aktør er ikke nødvendigvis det for en annen. Hvordan definerer vi og forstår mennesket? Er mennesket i det hele tatt autonomt fra de tingene vi omgir oss med og bruker? Hva er kontroll og hva mener vi med det? Kan man overhode si at man har «kontroll» over handlingsforløp? Og hva skal kontrolleres? Våpenet i seg selv, dets muliggjørende og kritiske funksjoner eller hvert enkelt angrep våpenet utfører?

Ved siden av de mer klare etiske, moralske og legale implikasjonene ved å overføre menneskelig kontroll til maskiner i spørsmål om liv og død handler dette også om interessante spørsmål rundt militær og politisk kontroll over handlingsforløp. Men alle disse spørsmålene beror på hva vi mener med menneskelig kontroll. Svarene på disse spørsmålene blir som oftest grunnet utfra en ingeniørbasert forståelse av at det eksiterer en binær separasjon mellom menneske og maskin der autonomi er definert som graderinger eller nivåer av kontroll i forholdet mellom operatør og våpen. Altså, jo mer autonomi som blir gitt til et våpensystem, definert som dets evne til å ta avgjørelser på egenhånd, jo mindre kontroll har mennesker og vice versa. Utfra disse synspunktene er meningsfull menneskelig kontroll noe som mennesket besitter, som det kan velge å oppgi, gjennom design og konstruksjon av maskiner. Ideen om kontroll strukturerer dermed hvordan systemer blir designet og konstruert som igjen former relasjonene mellom menneske og maskin, operatør og våpen og valg og handling. Utfra disse synspunktene handler spørsmålet rundt meningsfull menneskelig kontroll i stor grad om hvor mye eller lite kontroll mennesker skal avgi til maskiner og hva som skal være den førende vektingen for dette valget; etikken, jussen, operasjonell og strategisk kontroll eller det teknologisk drevne racet for å «ut-automatisere» rivalene? Men er det så enkelt? Kan vi simpelthen «sikre» kontroll gjennom å regulere hvor mye autonomi menneskene overgir til maskinene?

Denne analysen vil argumentere for at vi trenger å tenke nytt rundt forholdet mellom menneske og maskin for å bedre belyse spørsmålet om menneskelig kontroll. Det er et behov for å begynne å se på de autonome systemene i sin helhet, der de gjensidige relasjonene mellom menneske og maskin former hverandre og gir opphav til nye måter å tenke og handle på. Ikke bare er dette viktig for å forstå hvordan teknologier

påvirker, former og endrer relasjoner og muligheter for handling, men det er essensielt for å forstå hva kontroll er og innebærer. Det vil også bli argumentert for at autonomi og kontroll debatten må ta inn over seg at den utøvende fasen av en operasjon, liv og død-avgjørelser, er et for snevert nedslagsfelt hvis vi skal forstå hva kontroll er og hvordan det utøves. Dette vil bli illustrert via en gjennomgang av prosessene og praksisene rundt militær targeting eller metodisk målbekjempelse – en prosess som produserer kunnskap om fienden, operasjonsmiljøet og hvem eller hva som er mål og hvordan disse skal bekjempes – som igjen leder til valg og beslutninger om hvordan man skal føre krig. Gjennom denne illustrasjonen vil det bli vist at beslutningstaking ikke bare er *ett* valg gjort på et bestemt tidspunkt, som avgjørelsen om å sende et missil mot et bestemt mål, men noe som er avledet av rekke valg gjort over tid. Dette åpner opp for å se på beslutningstaking som et resultat av en rekke praksiser og prosesser distribuert over flere faser og mellom ulike former for menneske-maskin relasjoner, der noen valg blir foretrukket fremfor andre. Et slikt fokus setter søkelys på de mer sammensatte og komplekse teknologiske systemene for datainnhenting, analyse og lagring, som spiller en nøkkelrolle i utformingen av valg og beslutninger men som ofte blir neglisjert i debatten. Hvis vi bedre forstår hvordan disse systemene fungerer, hvordan de produserer kunnskap, hva de legger vekt på og hva de utelater, kan vi øke vår forståelse av kontroll og bedre «sikre» de etiske, legale og strategisk-politiske sider ved konseptet. Til slutt vil rapporten diskutere en rekke mulige utfordringer menneskelige kontroll står overfor i samspillet mellom visjoner og forestillinger om fremtidens krigføring og utviklingen av autonome systemer.

Spørsmålet om «menneskelig» kontroll

Selv om forkjempere og motstandere av autonome våpen har forskjellige syn på hva meningsfull menneskelig kontroll betyr, så eniges de om at en viss form for menneskelig kontroll er nødvendig og at dette kan «sikres» gjennom ulike former for juridiske, byråkratiske og designgitte reguleringer. Tidligere ble kontroll, i den grad dette begrepet gir mening i den kaotisk situasjon som krig er, «sikret» gjennom byråkratiserte og standardiserte rutiner, driller, trening og så videre, hvor tillitt til andre elementer i systemet var essensielt. Vestlige militære har i lengere tid har operert med hva de kaller «oppdragsbasert ledelse» – en desentralisert ledelsesfilosofi der kommando- og kontroll er sentralisert mens selve utførelsen av en operasjon er delegert ned til den med best situasjonsforståelse. Dette betyr i praksis at «kontroll» er distribuert, mellom nivåer og mellom mennesker, basert på tillit til hverandre. Det vi ofte glemmer i denne sammenheng er hvilken rolle maskiner og systemer spiller i disse beslutningstakingssløyfer. Kontroll er også i stor grad distribuert mellom mennesker og maskiner og i økende grad maskiner imellom. Vi kan dermed allerede i dag se at spørsmålet om menneskelig kontroll ikke er så entydig som ofte forstått. Den forventede utviklingen av autonome systemer, som i stor grad baserer seg på maskin-til-maskin interaksjoner i innhenting og analyse av data, synes bare å forsterke den distribuerte kontrollen og således bringe med seg en rekke nye spørsmål. Ikke minst for hvordan militæret tenker rundt kommando- og kontroll filosofi, struktur og arkitektur.

Den nevnte ingeniørbaserte forståelsen av autonomi ser på autonomi og kontroll som nivåer eller skalaer hvor mennesket er enten «i, på eller utenfor» den militære beslutningssløyfen. Det ingeniørbaserte konseptet «nivåer av automatisering (levels of automation, LOAs)» refererer til en balanse mellom automatisering og menneskelig kontroll i beslutnings- og handlingsselektering. LOA modellene spenner fra fullt ut automatisert der den menneskelige operatøren er helt utelatt fra beslutningen (nivå 10) til minimale nivåer av automatisering der

systemet gir anbefalinger eller bare filtrerer informasjon (Nivå 2-3). Disse modellene blir brukt av systemingeniører for å forstå hvordan de skal allokere roller og funksjoner mellom mennesker og maskiner når de skal designe og konstruere raske og effektive systemarkitekturer for beslutningstaking. Disse modellene er igjen basert på en rekke antagelser om hva styrkene og svakhetene til mennesker og maskiner er. Mennesker er generelt sett på som mer fleksible og innehar den nærmest ikke-vitenskapelige egenskapen «sunn fornuft» som gjør dem overlegne i situasjoner som er uoversiktlige og krever improvisering, som i krigføring. Maskiner på den andre siden er bedre til å prosessere mer informasjon raskere og på andre måter enn mennesker og til å respondere raskt og mer nøyaktig. Gjennom denne forståelsen av styrkene og svakhetene til mennesker og maskiner kan man så designe systemer der nivået av menneskelig kontroll er høyt eller lavt (Cummings, 2004).

Konteksten og graden av kompleksiteten til systemet og de funksjonene det skal ha, spiller en viktig rolle. Hvis «problemstillingen» er stor, som for eksempel et system for situasjonsforståelse i et større operasjonsområde, vil nødvendigvis også systemet være mer komplekst og inneholde en rekke elementer som, sensorer, analyseenheter og algoritmer som interagerer for å gi tidsriktig støtte til operatører. Behovet for menneskelig kontroll vil her være større enn i for eksempel et mindre komplekst system med et mer begrenset «problemstilling» som i automatiske luftvernssystemer. Slike vurderinger vil følgelig strukturere forholdet mellom autonomi og kontroll. Problemet med denne ingeniørbaserte tilnærmingen er imidlertid dets forenklete syn på både mennesker og teknologier, og enda viktigere, at disse systemene ikke blir sett på som et helhetlig ensemble av sammensatte elementer av både mennesker og maskiner. Ved å separere menneske og maskin gjennom graderinger og nivåer mister man forståelsen av hvordan relasjonene og interaksjonene mellom menneske og maskin gjensidig former hverandre, og dermed også hvordan systemet som helhet tenker og handler. Problemet med den nåværende debatten mellom autonomi og menneskelig kontroll er at den altfor ofte opprettholder denne ingeniørbaserte tilnærmingen. Forskning innenfor vitenskap- og teknologistudier (STS) har imidlertid imøtegått dette synet gjennom flere år med store empiriske studier av interaksjoner og relasjoner mellom menneske og maskin (Suchman, 2007). Kort fortalt har denne

forskningen funnet at mennesker og maskin gjensidig påvirker hverandre og at beslutningstaking og handling må sees på som et resultat av disse relasjonene og ikke som et resultat av noe iboende i enten mennesker eller maskiner. Den gjengse ingeniøroppfatningen om at kontroll er noe som mennesker besitter og kan velge å oppgi til fordel for mer autonomi til maskiner tar derfor ikke inn over seg hvordan relasjonene og interaksjoner mellom menneske og maskin fungerer i virkeligheten og gjensidig former hverandre. For å forstå kontroll må vi derfor studere systemet som helhet og hvordan relasjonene mellom de ulike elementene i systemet former hverandre og gir opphav til ulike former for beslutninger og handling.

Debatten om hva meningsfull menneskelig kontroll er og eventuelt hvordan man kan oppnå dette lider derfor under en rekke innsnevring og forenklinger som begrenser vår forståelse. Det første problemet er som nevnt den klare separasjonene mellom menneske og maskin. Vi mennesker har alltid brukt instrumenter og verktøy for å utvide vår intelligens, fra huletegninger til avanserte datamaskiner. Et verktøy gjør ikke bare ting enklere, det muliggjør nye måter å tenke på, å leve på, å være på, og derfor også hvordan vi handler. Alle teknologier inneholder en form for kunnskap, en måte å se verden på og handle i den (van Creveld, 1989), men å se på teknologiske verktøy og teknikker som «sorte bokser» som mennesket kan bruke som de ønsker blir fort snevert. Ved å delegere noen roller til teknologier er ikke lenger mennesker i fullstendig kontroll av den enkle grunnen at teknologier gjør mer enn å uttrykke menneskets vilje, det påvirker det, og ofte på uforutsette måter (Bourne, 2012). Vi mennesker blir formet av disse verktøyene og teknologiene og ofte på måter som ikke var påtenkt av de som designet og konstruerte disse. Litt forenklet så kan man si at systemet mennesket og øks ser på trær og skog som noe annet enn menneske uten øks. Da blir skogen en ressurs på en helt annen måte enn hvis man ikke hadde muligheten til å felle og bearbeide trær. Således er verktøyene vi utvikler og besitter med på å forme våre tanker og handlinger, men de gjør ikke alltid dette på en lineær måte eller i de baner de var tiltenkt. Selv om vi kan si at vi har kontroll over øksen vi har laget, og at det er vi som velger om vi vil bruke den og til hva, så har øksen allerede formet hvordan vi ser på ting. Det å avstå fra å bruke en ting er like mye et valg som å velge å bruke det.

Det er denne påvirkningskraften eller aktøregenskapen som teknologier har i samspillet mellom mennesker i komplekse systemer, som er viktig å forstå hvis vi skal skjønne hva menneskelig kontroll er og i hvilken grad vi «besitter» den. Hvis det er slik at både teknologier og mennesker blir noe annet i relasjon med hverandre så betyr dette også at enhver endring i sammensetningen av elementene i et system potensielt kan endre dette systemet og dermed de beslutninger og handlinger systemet som helhet tar og utfører. Man må derfor spørre seg om hvordan vi former teknologier og hvordan teknologier former oss og hva resultatet av disse interaksjonene og relasjonene mellom menneske og teknologier er. Denne innsikten er ikke bare av teoretisk eller akademisk interesse, men er viktig for hvordan vi skal forstå og sikre kontroll, og ikke minst for hvordan man kan planlegge. Selv om vi har en tendens til å mene at vi er autonome fra de teknologier vi designer og skaper, er faren at vi her ikke ser skogen for bare trær. At vi blir så opptatt av enkeltelementer i systemene at vi mister fokuset på helheten.

Det andre store problemet med dagens debatt er at det er et for stort søkelys på våpen og våpensystemer. Fokuset på den fysiske manifesteringen av autonomi, som bildet av autonome «killer robots» eller Terminator, tar oppmerksomhet fra de mer sammensatte og komplekse teknologiske systemene for datainnhenting (sensorer), analyse og lagring, som spiller en nøkkelrolle i utformingen av valg, beslutningstaking og handling. Utover en diskusjon av hvordan systemene kan eller ikke kan forholde seg til legale begreper som proporsjonalitet og distinksjon i den prosessen militæret kaller for targetting, handler denne debatten stort sett om man skal gi eller ikke gi våpensystemer muligheten til å ta egne valg. Søkelyset på våpen og den spisse enden av militære operasjoner, den utøvende fasen, neglisjerer dermed hvordan beslutningene som tas forut for disse operasjonene allerede er avledet av en rekke valgmuligheter som oppstår underveis. Fra de visjonene og kunnskapene som disse systemene bygger på, designen og konstruksjonen av disse systemene, til den distribuerte interaksjonen mellom mennesker og maskiner som foregår gjennom hele beslutningsprosessen, fra innhenting og analyse av data opp til avgjørelsen om å bruke voldsmakt. Det betyr at vi ikke bare kan se på menneskelig kontroll som noe som skjer i kun i den utførende fasen, men noe som skjer gjennom hele prosessen i samhandling med maskiner.

Kentaurer og andre myter

Forsvarsanalytiker Paul Scharre har vært en av dem som har prøvd å nyansere dagens debatt og styre den bort fra mennesker kontra maskiner og alle mytene rundt kunstig intelligens, mot spørsmål omkring integreringen av menneske og maskin (Scharre, 2016). Der debatten har brukt «Terminator» som et populært referansepunkt prøver Scharre, hvis vi skal fortsette Hollywood-analogiene, å få oss til å tenke mer på «Ironman». Ifølge Scharre så er ikke bare menneske-maskin lagspill mulig, men også foretrukket i fremtidens krigføring. Gjennom å dra på erfaringer fra såkalt «kentaursjakk» der mennesker jobber i lag med maskiner, utvikler Scharre en idé om kentaurskriføring der målet er å utnytte de beste egenskapene til både mennesker og maskiner. I menneske kontra maskin debatten er mennesker det tregeste elementet i beslutningssløyfen, mens maskiner fremdeles er fryktelig dumme når det kommer til uventede hendelser. Hvordan skal man løse dette problemet?

Hva Scharre argumenterer for er en slags menneske+maskin symbiose eller kentaurskriføring – halvt menneske, halvt maskin – der målet er en utvidet intelligens som drar nytte av både menneskers og maskiners styrke. Slike hybride menneske-maskin kognitive strukturer, som Scharre forestiller seg, vil ikke bare kunne utnytte maskinenes fart, pålitelighet og nøyaktighet, men også menneskers robusthet og fleksibilitet og kanskje viktigst av alt vår «sunne fornuft» – den evnen som gjør oss kapable til å tilpasse oss nye og uventede situasjoner. Selv om dette i teorien høres veldig fornuftig ut er det selvfølgelig ikke bare enkelt å overkomme svakhetene til begge ved å kombinere styrkene til begge. Mennesker vet vi hvordan vi lager, utvikler og trener opp, og forskning på kunstig intelligens og maskinlæring gjør stadige fremskritt i å gjøre maskiner «smartere». Problemet ligger i +’en. Scharre sin løsning er imidlertid basert på den nevnte ingeniørbaserte tilnærmingen til mennesker og maskin der man kan designe, konstruere og trene frem en + og dermed oppnå en sømløs symbiose med det beste fra hver. Det Scharre egentlig beskriver er ikke en symbiotisk kentaurskriføring, men snarere et system der mennesker og maskiner er separert og hvor

mennesker blir redusert til moralske agenter og feilsikringer (fail-safes), som brukerstøtte til maskiner, for å sikre etisk, legal og operasjonell kontroll over de raskere og mer nøyaktige maskinene.

For at Scharre skal kunne argumentere slik han gjør er han også nødt til å basere sin tilnærming til autonomi og kontroll på det vi kan kalle en ingeniørbasert tilnærming til krig og dets fokus på den utøvende fasen av krigføring. Dette er en ide om krig som noe som utspiller seg på en slagmark, begrenset i tid og rom, der handlingsforløpet i stor grad kan beregnes gjennom kalkyler, modelleringer og simuleringer, og der sekvensielle taktiske seiere vil lede til strategisk suksess. Dette forenklete og begrensede synet på krig, der krig blir simpelthen redusert til krigføring eller militære operasjoner, er Scharre langt ifra alene om. Det er dette som danner basisen for den vestlige tilnærmingen til krig med amerikanerne i spissen (Echevarria, 2004). At krig kan bli redusert ned til tekniske problemer som kan løses gjennom tekniske tilnærminger stemmer dessverre ikke overens med de empiriske erfaringer vi har gjort gjennom årtusener med denne menneskelige aktiviteten. Det er ikke tilfeldig at Clausewitz sammenliknet krig med kortspill snarere enn sjakk, som Scharre gjør, der tilfeldigheter og eventualiteter er rådende snarere enn kalkulert kontroll. Selv om sjakk er et komplisert spill med utallige valgmuligheter så er disse begrenset av klare regler og handlingsmønstre. Det er også viktig å huske at når Clausewitz skrev om krigens tåke, friksjon, tilfeldigheter og eventualiteter så gjorde han ikke det bare for å belyse at dette var en del av krigens natur. Han gjorde det som en advarsel til politikere om at krig ikke er et fenomen som kan reduseres til noe kontrollerbart ved hjelp av matematiske formler. En advarsel vestlige politikere og militære har valgt å neglisjere, all den tid de bruker på å redusere krig til et teknisk problem åpent for tekniske løsninger.

Scharre er dermed med på å opprettholde de to hovedproblemene med dagens debatt nevnt innledningsvis; separasjonen mellom menneske og maskin og det overdrevne søkelyset på roboter og våpen, og den utøvende fasen av militær operasjoner. For å komme bort fra dette vil rapporten nå sette søkelys på targetting, eller metodisk målbekjempelse som Forsvarets fellesoperative doktriner kaller det (Forsvarsstaben, 2014). Targetting er en militær operasjonell praksis som kan illustrere hvorfor både det forenklete synet på menneske-maskin

relasjoner og det snevre fokuset på den utøvende fasen av krigføring bør endres for bedre belyse spørsmålene rundt autonomi og menneskelig kontroll. Targeting er viktig fordi det er en prosess der kunnskap produseres, valg blir tatt og handling gjennomført. Gjennom en illustrasjon av targeting kan vi se hvordan relasjonene mellom ulike elementer, mennesker og maskiner, i et system er med på å forme beslutningstaking og også handling i forskjellige faser av prosessen. Dette er med på å belyse hvordan kontroll ikke bare er fordelt mellom de ulike fasene, som en tillitsbasert rekke av avgjørelser fra analyse til våpenbruk med beslutningstaking imellom, men snarere bør sees på som en prosess der valg og avgjørelser er distribuert i relasjonen mellom mennesker og maskin.

Targeting og menneske-maskin relasjoner

NATOs Joint Targeting doctrine, AJP-3.9 (NATO, 2016) definerer *targeting* som en prosess som kopler det strategiske nivå med taktiske aktiviteter gjennom målrettingscykluser på det operasjonelle nivå. Targeting fungerer dermed som en bro mellom det taktiske og strategisk-politisk nivå hvor fokuset er på å skape fysiske og kognitive effekter slik at man oppnår militære målsetninger og den ønskede slutttilstand. Dette er delt opp i seks faser.¹ De pre-definerte strategisk-politiske målsetninger er første ledd og danner grunnlaget for targetingsyklusen. Den kanskje viktigste, men sjeldent fremhevede, fasen er målutvikling og målutvelgelse der man identifiserer hvilke objekter eller individer som skal være mål for operasjonene og kopler disse sammen gjennom en systemisk analyse hvor de ulike delenes opptreden og gjensidig avhengighet blir fremhevet. Således skaper man seg et bilde av fienden og operasjonsmiljøet man befinner seg i, som igjen blir brukt til å skape sammenfallende effekter på og utenfor slagmarken. Basert på dette blir noen mål valgt ut og den neste fasen, som innebærer hvordan og med hva disse målene skal engasjeres, starter. Etter å ha tatt en avgjørelse beveger man seg så inn i planlegging og utførelsesfasen og til slutt, etter endt operasjon, gjøres det en vurdering av måloppnåelse for å anslå om man oppnådde de effekter man ønsket seg eller ikke. Og slik fortsetter denne syklusen til man har nådd de militære målsetninger og ønsket slutttilstand. Selv om den doktrinnelle beskrivelsen av hvordan targeting prosessen skal fungere kan virke stilistisk, lineær, og i overkant sekvensiell, er den i realiteten mer dynamisk, iterativ og toveis og faser kan overlappe. Men som doktrinen til det Amerikanske luftforsvaret så

¹ Jeg skiller ikke her mellom planlagt og dynamisk targeting. Selv om det åpenbart er forskjeller er det etter min mening i stor grad bare tidsaspektet som skiller dem og således vil disse fasene, i mer eller mindre grad, være til stede i begge former for targeting. Et annet aspekt er at dynamisk targeting er mer responsivt enn planlagt targeting i den form av at det vil være objekter eller personer som blir oppdaget som ligger utenfor de planlagte målene som en beslutningstaker tar en avgjørelse på.

oppriktig beskriver det så er det lett å bli opphengt i de dagligdagse gjøremål og operasjonell tenkning:

It is easy for those caught up in the daily battle rhythm to become too focused on tactical-level details, losing sight of objectives, desired effects, or other aspects of commander's intent. When this happens, execution can devolve into blind target servicing, unguided by strategy (U.S.AirForce, 2019, p. 40).

Dette betyr ofte i realiteten at targeting-prosessen, spesielt den utførende fasen, er med på å forme og i noen tilfeller diktere de strategiske målsetningene som skulle danne utgangspunktet for prosessen. At mål, metoder og midler ofte former hverandre på uforutsette og ikke-lineære måter er ikke noe nytt. Ser vi derimot nærmere på targeting-metodologien utformet av amerikanske spesialstyrker i deres globale kontraterroroperasjoner, kjent under sitt akronym F3EAD (find-fix-finish-exploit-assess-disseminate), så er det klart at targeting er mer enn bare en prosess for mest effektivt å finne og «drepe» et fiendtlig mål. Gjennom en fusjon av etterretning og operasjoner har F3EAD (NATO, 2016, pp. Section III, 5-7) gjort targeting til en dynamisk prosess der targeting ikke bare er veien til målet men snarere en prosess for å lære fienden å kjenne. Ifølge tidligere CIA og NSA direktør Michael Hayden, pleide fienden å være lett å finne men vanskelig å drepe, mens nå er den veldig vanskelig å finne men lett å drepe (Hayden, 2016, p. 40). Selv om terrortrusselen kanskje ikke er like presserende som den en gang ble utgitt for å være, og man nå har returnert til «fiender» som er lette å se, så gjør deres tilsynelatende ukjente og flyktige kapabiliteter og operasjonsmønstre i hybride eller gråsonetekster operasjonsmiljøet mer og mer komplekst. I denne sammenheng har derfor targeting prosessen seilet opp som en av de viktigste praksisene for å forstå fienden, det stadig skiftende operasjonsmiljøet og operasjonaliseringen av dette. Som Mitch Ferry (2013, p. 61) forklarer det:

Targeting is used now not merely to neutralise threats or prepare the operational environment, but to learn about the adversary. Whereas learning had previously been a means to an end (to enable a strike), it is now an end in itself. *What targeting achieves* has changed, bringing opportunities and threats related to the dual outcomes. Targeting in a complex environment invites a tension between learning in order to affect and affecting in order to learn.

Sett gjennom disse linsene er det ikke bare slik at taktikk og operasjoner tar over for strategi og overordnede målsetninger, som militæret ofte blir kritisert for, men at *targeting*-prosessen i stor grad er målsetningen. Det blir et mål i seg selv fordi vedvarende operasjoner er sentrale ikke bare for å foregripe en situasjon, men viktigere, for å forstå fienden og operasjonsmiljøet. At etterretning er operasjoner og operasjoner etterretning, samlet under *targeting*, ser vi ikke bare i spesialstyrkers kontraterroroperasjoner, men i økende grad også i amerikanske cyberstrategier og operasjonelle konsepter i deres forsøk på å oppnå overlegenhet i cyberspace (Healey, 2019).

Det vil være en underdrivelse å si at disse *targeting*-prosessene er påvirket av ulike former for nettverksteknologier; sensorer, analyseapparater, databaser og kommunikasjonsutstyr. Uten disse ville man rett og slett ikke kunne drive disse operasjonene, ei heller tenke i de baner. I et hav av data er det bare datamaskiner som kan analysere tilstrekkelig fort. Dette baserer seg igjen på en visjon om at hvis alt kan datafiseres og analyseres så vil dataene avsløre viktige mønstre som vil gi viktig tidsriktig informasjon til å gjennomføre operasjoner som igjen vil føre til mer data og nye operasjoner. Det er i stor grad disse forestillingene og syklusene, og ikke overordnede målsetninger, som nå driver militære operasjoner. I så henseende kan man argumentere for at man er i den situasjonen den amerikanske luftforsvarsdoktrinen advarte mot. Men i hvilken grad dette er bevisst, et resultat av tilfeldigheter, eller et resultat av de interaksjonene og relasjonene som finnes mellom mennesker og maskiner i *targeting*, er ikke kjent. På den ene siden er både de sosiotekniske praksisene rundt databehandling og *targeting*-prosessene designet som evigvarende lærende prosesser, og det kan derfor argumenteres for at dette er en ønsket utvikling (Nordin & Öberg, 2015). På den annen side så er det også mulig å argumentere for at dette har ført til at mennesker ikke er i «kontroll» i den grad de er blitt «fanget» i sine egne systemer der stadig nye mønstre i data skaper stadig nye mål og evige nye operasjoner.

Targeting, kontroll og fremtidens autonome systemer

Targeting-analyse, som ikke bare danner grunnlaget for utførelsen av bestemte angrep, men som også produserer kunnskap og forståelse av hva trusselen er, hvordan den ser ut, og hvordan man best skal hankses med den, er vitalt for å skape situasjonsforståelse og hvordan man skal operere. Siden dagens systemer, der vi mener vi har meningsfull menneskelig kontroll, er lagt opp som sentralisert kommando- og kontroll (C2) med desentralisert utførelse, så gir det liten mening å argumentere for at operatøren av et våpensystem skal ha «kontroll» så lenge denne personen ikke har full tilgang til nøyaktig den samme informasjonen som de som har tatt avgjørelsen om bruk av makt har. I disse systemene er kontroll «sikret» gjennom menneskelig tillitt og det er ikke noe prinsipielt i veien for at denne tilliten ikke kan forlenges til maskiner. Vi kan til og med gå så langt å si at dette i stor grad også er dagens situasjon, der stadig mer av innsamlingen, prosesseringen, analysen og lagringen av digitale data er maskin-til-maskin-basert utenfor menneskelig observasjon, innflytelse eller «kontroll».

Selv om det finnes mange variasjoner av kunstig intelligens, er den største utfordringen når det gjelder menneskelig kontroll knyttet til nettopp de systemer som analyserer data og produserer kunnskap. Alle autonome systemer, og noen vil si mennesker også, følger mer eller mindre den samme grunnleggende kognitive prosessen; de samler data, de prosesserer og analyserer data, og de tar valg og handler. Dette likner prosessene som «manuelle» militære planleggingsmetodologier som targeting følger, og det er dermed her vi både ser og kommer til å se den rivende utviklingen i autonome systemer først. Når denne prosessen allerede er, i stor grad, drevet av svært omfattende automatiske og autonome datainfrastrukturer for innhenting, analyse og lagring av data er det betimelig å spørre om mennesket har kontrollen, og i hvilken grad det er autonomt og har frie valg? Slike spørsmål åpner opp for en mer informativ debatt som ikke bare har å gjøre med de etiske og juridiske sidene ved bruken av makt, eller mulige fremtidige «killer robots», men

i større grad om kontroll og en bevisstgjøring av hvordan de valg vi tar i dag er med på å forme fremtiden.

I debatten rundt menneskelig kontroll blir det ofte argumentert for at mennesker bør være kognitivt bevisste på hvordan maskiner har kommet frem til spesifikke valg (Roff & and Moyes, 2016). Men spørsmålet om kognitiv bevissthet og forståelse av situasjonen er ikke bare et problem som dukker opp i forbindelse med autonome systemer, men som også er et generelt spørsmål i en kaotisk situasjon som den på slagmarken. Det er lite sannsynlig at en soldat på bakken eller en pilot i luften har en slik overordnet situasjonsforståelse gitt dagens C2 systemer, men det er heller ikke meningen. Her er beslutningstakingen sentralisert mens utførelsen er desentralisert under den ledelsesfilosofien vi i Norge kalle oppdragsbasert ledelse. Kontroll og situasjonsforståelse ligger dermed hos den sentraliserte kommandostrukturen og ikke hos den enkelte operatør selv om disse selvfølgelig har muligheten til å gjøre seg opp en mening om situasjonen på bakken eller i luften gitt den overordnede informasjonen de er gitt. Kontroll er distribuert, mellom mennesker, mellom mennesker og maskin, mellom maskiner og i alle faser av for eksempel targeting og andre beslutningstakings og kunnskapsproduserende systemer. I så henseende kan vi si at det hverken finnes autonome systemer, og ei heller autonome soldater, piloter eller sjømenn. Det spiller for eksempel liten rolle, bortsett fra rent juridisk rundt plasseringen av ansvar, om det er et menneske som til syvende og sist trykker på knappen hvis denne avgjørelsen er tatt på bakgrunn av valg som er utledet og/eller diktert av en maskin som menneskene har liten eller ingen interaksjon med eller ikke forstår. Når en beslutning er avledet av en lang rekke med valg som oppstår underveis i interaksjonen mellom mennesker og maskin må også spørsmålet om kontroll ta dette innover seg. Spørsmålet om menneskelig kontroll bør derfor ikke dreie seg om en operatør i den utførende fasen av en operasjon har nok informasjon til å ta en informert beslutning, og heller konsentrere seg om spørsmål om man «stole» på eller i det hele tatt forstå de valgene som er skapt av autonome systemer.

Kunstig intelligens og autonome systemer kommer med en lovnad om å ikke bare forenkle og redusere arbeidsmengder, men også med et løfte om å se og oppdage ting vi mennesker ikke kan oppdage, og forutsi hendelser før de skjer. Ikke bare er denne lovnaden problematisk i seg

selv, men utviklingen av avanserte dype nevralt nettverk og andre maskinlæringsystemer bringer med seg et annet stort problem. Nemlig det at de innehar en iboende ugjennomsiktighet. Dette er et resultat av at disse systemene endrer seg selv basert på input og output, det de lærer, og ikke som et resultat av hvordan mennesker trener dem. Gitt at disse endrer seg og gitt at vi ikke kan «åpne» dem og se hvordan de kom frem til de svarene de kom frem til, vil vi mennesker heller ikke forstå dem. Maskinene vil gjennom læring og basert på de dataene de har, begynne å gjøre seg opp en mening om hvordan «verden» fungerer og således skape sine egne forutinntattheter og ideer. Selv om det nå er slik at de beste systemene også er de minst forståelige (Bornstein, 2016), er det mulig at det er nettopp disse systemenes ugjennomsiktighet som gjør dem lite attraktive i for eksempel targeting-analyser. Der disse systemene i stor grad er skapt for å unngå menneskelige ideer og forutinntattheter skaper de seg sine egne som vi ikke nødvendigvis kan ha kontroll over.

Men vi trenger ikke trekke frem nevralt nettverk som lærer uten menneskelig intervensjon for å poengtere hvordan autonome maskinlæringsystemer kan påvirke menneskelig kontroll. De maskinlæringsystemene vi nå bruker og bygger i stor skala, blir lært opp til å se verden på en bestemt måte etter bestemte mål og forutinntattheter som vi mennesker har. Disse maskinene og algoritmene vil alltid kunne finne mønstre i data, og deres lovnader om at de vil finne og forutse viktige ting vil dermed forsterke dette spesifikke synet, og følgelig ikke bare strukturere hva vi «ser», men også hvordan vi forstår verden og handler i den. Hvis man handler på bakgrunn av at det er mulig å predikere verden, så hjelper det lite at kritikken rundt effektbasert tenkning og targeting og den ingeniørbaserte tilnærmingen til krig har satt spørsmålstegn ved mulighetene ved å kalkulere og forutse effekter. Vi kan selvfølgelig velge å ikke bruke disse analyseverktøyene, men så lenge vi utvikler og bruker dem så former de måten vi ser på verden og handler i den, mens de ekskluderer andre muligheter. Slike systemer danner det forskeren Paul Edwards i sin gjennomgang av den amerikanske utviklingen av digitale kommando- og kontrollsystemer under den kalde krigen kalte for en «lukket verden». En selvrefererende lukket verden som skulle vise seg å være instrumental i å forme og omskape den politiske konteksten disse systemene hadde sitt utgangspunkt i (Edwards, 1996).

Når amerikanerne nå er i ferd med å bygge nye avanserte kommando- og kontrollsystemer som Advanced Battle Management System (ABMS) og Joint All-Domain Command & Control (JADC2) kan vi igjen se konturer av disse selvreferendene lukkede verdenene. Der den kalde krigen produserte avanserte automatiske kommando- og kontrollsystemer for å detektere Sovjetiske missiler og fly eller etterretningsinnhentingsystemer som «Igloo White» under Vietnamkrigen (Correll, 2004), så er dagens visjoner om nettverksbasert krigføring av en annen skala og kompleksitet. I dag er autonomi sett på som nøkkelen til endelig å muliggjøre drømmene fra visjonene om revolusjon i militære anliggender (RMA) fra 1990-tallet om å samle og nettverke *alt* i et gigantisk elektronisk nervesystem eller økosystem (Tucker, 2017). En fremtidsvisjon der alt er sensorer, alle ting og mennesker opererer i et sømløst informasjonsdelingssystem hvor slagmarken blir transparent, ingenting kan gjemme seg, og krigens tåke har lettet. Selv om det amerikanske forsvarsdepartementets direktiv 3000.09 og uskrevne policy er klar på at mennesker ikke vil bli fjernet fra beslutninger om å drepe (U.S.DoD, 2012) så muliggjør disse systemene «krig på autopilot». Uansett hvordan disse systemene vil bli seende ut i fremtiden, så tyder mye på at kontroll vil i stor grad bli mer distribuert enn før. Den sømløse delingen av informasjon og behovet for flere og raskere avgjørelser gjør sannsynligvis at det ikke vil være hensiktsmessig å sentralisere kontroll, selv om kommandostrukturen fremdeles vil være sentralisert. Dette kaster nytt lys over debatten om hvem eller hva som skal kontrollere avgjørelser.

Selv om forestillingene og visjonene for disse systemene er preget av de forventningene kunstig intelligens og autonomi bringer, ser vi samtidig at opphavet til de systemene vi nå bygger og de visjonene vi har for fremtidens krigføring er dypt forankret i historiske ideer om og kunnskaper rundt krig og krigføring. For å forstå hvordan disse systemene blir bygget og deres eventuelle konsekvenser er det derfor viktig å forstå samspillet mellom spesifikke teknologier og deres opphav i militære og kulturelle forestillinger om nettverksbasert krigføring. Med andre ord: Våre forestillinger om fremtidens krigføring og de teknologier vi ser for oss er gjensidig forsterkende og bidrar dermed til å sementere spesifikke framtidsscenarier mens de ekskluderer andre.

Menneske-maskin og militære visjoner

Automatisering av militære oppgaver har en lang tradisjon som strekker seg tilbake til andre verdenskrig og utviklingen av luftvern, datamaskinen og kybernetikken til utviklingen av den «elektroniske slagmarken» og presisjonsvåpen under Vietnamkrigen. At denne automatiseringen har økt i takt med den tekniske og teknologiske utviklingen er ikke spesielt oppsiktsvekkende. Å erstatte mennesker med maskiner for å øke effektivitet og presisjon, eller for å ta de kjedelige, skitne, farlige og vanskelige jobbene, er i seg selv fullt forståelig. Men i våre fremtidsvisjoner om kunstig intelligens og autonome våpensystemer, enten man vil ha dem eller ikke, har det nå etablert seg en «sannhet» i sikkerhetspolitiske kretser om at dette er en «militære nødvendighet». Slik nødvendighet oppstår når man ensidig ser på teknologier som problemløsende verktøy snarere enn å prøve å forstå hvilke mulige implikasjoner utviklingen av visse systemer har utover det umiddelbare bruksområdet det er ment til. Dette er spesielt viktig i debatten rundt menneskelig kontroll over automatisering og autonomi i våpensystemer.

Visjonene og drømmene, men også frykten rundt kunstig intelligens er at disse teknologiene lover en slagmarkstransparens som vil muliggjøre en nær perfekt targetting av fiendtlige styrker med en fart man ikke kunne ha drømt om for kun få år siden. Her vil mennesker bare være en hindring for den friksjonsløse operasjonen til maskinkrigføring. Mennesket, synes det, er nå blitt overflødig i krig. Forskjellen fra tidligere er at nå er ikke bare maskiner sett på som en støtte til mennesker i ulike typer arbeid, men snarere at mennesket, hvis ikke tatt helt ut, blir en støttefunksjon til maskiner. Det er som vi har sett strengt talt ikke nødvendig med mennesker hverken i innhenting eller analyse av informasjon, i valg av handlingsmønstre, eller i implementeringen av disse valgene, med mindre man ønsker en viss form for menneskelig kontroll. Alt dette er maskiner kapable til gitt muligheten. Ikke bare viderefører dette en lang tradisjon med å fjerne mennesket fra

slagmarken ved bruk av avstandsvåpen som fly og missiler, men det gjelder også i beslutningstaking på alle nivåer. Nå er det ikke bare maskinautonomi og dets iboende evner til å prosessere informasjon raskt, eller ønsket om å skape en friksjonsløs slagmark uten mennesker, som er de eneste drivkreftene som trekker utviklingen i retning av flere og flere maskiner som handler på egenhånd. Erkjennelsen av at kommunikasjonene mellom mennesker og maskin, som er essensielt for menneskelig kontroll, og for de krigføringssystemene vi har bygget, er særdeles sårbare overfor ulike typer av elektronisk krigføring og cyberangrep har også ført til argumenter for økt autonomi til ulike våpensystemer som dagens fjernstyrte UAVer. Uansett, autonomi blir forklart som en teknisk løsning på vår reduksjon av krig til et teknisk problem.

Der berettigelsen for automatisering tidligere var basert på menneskelig feilbarlighet og mangler og ønske om å fjerne det fra risiko, er det nå automatisering i seg selv som blir brukt som grunnlaget for hvorfor man trenger mer automatisering og autonomi. Dette er ikke ulikt dynamikker man tidligere har studert under konseptet våpenkappløp, der en stats nye våpensystemer berettiger økte investeringer og utvikling fra andre stater. Ikke bare ser man på utviklingen av autonomi som en eksponentiell og irreversibel prosess, men også noe man i stor grad ikke har kontroll over, og som vil lede mot hva som nå blir kalt «slagmarkssingularitet» (Kania, 2017). En tilstand der de tradisjonelle militære beslutningsprosessene vil foregå i maskinhastighet og dermed presse mennesket ut av beslutningssløyfene. Denne «militære nødvendigheten» av mer automatisering skaper sin egen sirkulære logikk der utvikling og bruk av autonome systemer med liten eller ingen menneskelig innblanding leder til stadig kortere tidsrammer for beslutningstaking som igjen krever mer autonomi. Dette er med andre ord en prosess ute av (menneskelig) kontroll, der vestlige stater må «tilpasse seg eller dø» i konkurranse med andre teknologiske avanserte stater som Kina og Russland. Ikke bare er dette med på å sette i gang et våpenkappløp som i seg selv kan ha katastrofale følger, men det er også med på å forme og videreføre en forståelse av fremtidens krigføring basert på en teknoideologi og de utfordringer og sårbarheter dette bringer med seg.

Denne logikken inneholder også et stort paradoks og dilemma. På den ene siden blir vi fortalt at utviklingen av kunstig intelligens og autonomi er en prosess utenfor kontroll. På den andre siden snakker vi om hvordan vi kan utnytte denne prosessen best mulig. Hvis man så aksepterer en viss form for kontroll i form av at man kan utnytte utviklingen, oppstår det et dilemma. Utviklingen av mer autonomi i systemene vil ikke bare presse mennesker ut av beslutningssløyfen, men også gi krig sin egen logikk der politisk og strategisk kontroll over krigføring vil være minimal. Som dresultat kan vi gjenoppta den akademiske debatten om hvorvidt «krigens uforanderlige natur» faktisk er forandret.

Et slikt ekstremt tankeeksperiment rundt den spiraliserende effekten autonomi har på visjoner om fremtidig krigføring er i stor grad muliggjort av hvordan vestlige land forstår krig. Mye av den militære debatten, og til dels også den etiske og legale debatten rundt autonomi og krig, legger til grunn en snever forståelse av krig og krigføring der slagmarken og den spisse enden av krigføringen får stor plass. Dette gir opphav til en forståelse av krig som noe som skjer på den avgrensede slagmarken og som i stor grad kan beregnes gjennom å sette søkelys på de kalkulerbare siden av krigføring, det vil si våre materielle kapasiteter kontra fiendens. Således kan vi si at i tillegg til de mer tekniske problemstillingene man står overfor i utvikling av nye overgripende kommando- og kontroll systemer, systemer for situasjonsforståelse og datainnhenting og analyse, vil også disse systemene og deres design og utvikling i stor grad være styrt av visjoner og forestillinger om hva fremtidens trussel er, hvilke funksjoner systemene må ha for å forholde seg til denne trusselen, ulike historiske betingelser, måter å operere på, men kanskje viktigst, våre ideer om hva krig er og hvordan disse skal vinnes. At man i stor grad har redusert krig til en teknovitenskap der man har fremhevet de målbare sidene av krigføring, muliggjør også visjonene om «slagmarksingularitet». Hvis man derimot ser på krig som et menneskelig fenomen formet av en rekke faktorer som ikke bare motsetter seg kalkuleringer, men som også i aller høyeste grad er underlagt tilfeldigheter og eventualiteter så vil ikke autonomi forandre dette. Dette betyr ikke at visjonene om slagmarkstransparens og forkortede tidsrammer ikke vil være resultatet av en fryktdrevet teknologisk utvikling, men derimot at krig handler om så mye mer enn slagmarken. Og det er her den store utfordringen ligger. Hvordan kan

man forene den tilsynelatende operasjonelle militære nødvendigheten av stadig økende behov for autonomi med etisk, juridisk, og strategisk-politisk menneskelig kontroll? Paradoksalt nok er det nettopp frykten for «slagmarkssingulariteten» som gir opphav til to vidt forskjellige syn på denne saken; enten et forbud mot autonome våpensystemer eller behovet for å ligge i forkant av utviklingen av disse systemene. Begge synene er virker å være like meningsfulle for de «stridende».

Å forene det oppfattede behovet for autonomi med behovet for kontroll begynner med å revurdere ideen om meningsfull menneskelig kontroll. Det betyr kort fortalt at vi må se på teknologier som noe mer enn bare instrumenter som viderefører menneskelig vilje, og skape en forståelse som tar innover seg at teknologier er en aktiv deltager som former hvordan vi tenker og handler. Mennesker, samfunnet og teknologier samproduseres i et fellesskap og endrer hvordan vi ser på verden og hvordan vi handler i den. Teknologier er hverken nøytrale eller determinerende, men de fremmer noen handlinger og ekskluderer andre. Kort fortalt så inneholder slike sosiotekniske systemer ikke bare en viss form for kunnskap, men også politikk og ideer om fremtiden som blir bakt inn i disse systemene, som igjen er med på å forandre denne fremtiden. Derfor må vi også tenke dypere på menneske-maskin relasjoner og interaksjoner i systemer, og innse at kontroll og beslutningstaking er noe som skjer hele tiden, i alle faser fra design til utførelsesfasen, og mellom operatør og våpen.

En slik forståelse av hvordan menneske-maskin systemer og prosesser fungerer synes å være fraværende i mye av debatten rundt våpensystemer. Ved å ta utgangspunkt i at mennesker og maskiner allerede er konfigurert sammen, at våre evner til å tenke og handle og vår autonomi er effekter av våre samhandlinger, snarere enn noe som er iboende i mennesker eller i maskiner, så kaster vi et nytt lys over debatten. Hvis evnen til å handle er et resultat av ulike menneske-maskin konfigureringer, må vi tenke over hvordan disse blir designet og utviklet. Vi må også tenke gjennom hvilke visjoner og forestillinger vi har om krig og krigføring og om fremtiden som ligger bak denne utviklingen, samtidig som vi tenker nøye over hvordan teknologier er med på å endre oss. Det er ikke bare slik at automatiske eller autonome teknologier erstatter mennesker, men snarere at mennesker får forskjellige roller innenfor det samme systemet. Det kan øke eller minske

menneskelig kontroll, men dette beror på at vi ikke ser på mennesket som noe autonomt. Mennesket i meningsfull menneskelig kontroll er ikke ubesudlet av de relasjonene vi allerede har med de tingene vi omgir oss med, men er formet av og samtidig former de tingene. Likeså så har vi muligheten til å forme verden med de tingene vi tenker fra og handler med, men da er det også viktig at vi tenker på alle elementene i systemer som integrerte. Designer og konstruerer vi systemer der det feilbarlige menneskelige element blir fortrent til fordel for den tilsynelatende mer ufeilbarlige maskinen, oppstår det mest sannsynlig andre uante dynamikker med andre «feil». Spørsmålet om meningsfull menneskelig kontroll er derfor ikke et ja eller nei, mer eller mindre, har eller ikke har, som så enkelt kan besvares. Kontroll er allerede den dag i dag distribuert på måter som ikke alltid er forståelig eller gir mening selv for de som sitter midt oppi det.

Det er viktig at vi i Norge har en åpen og informativ debatt rundt spørsmålet om autonomi og kontroll og hva slags systemer vi vil ha. Vi må nøye tenke gjennom balansen mellom de lovnader som kommer med kunstig intelligens og autonomi i form av bedret effektivitet og prediksjon og de etiske, legale og politiske behovene for kontroll. Systemene som amerikanerne nå utvikler vil muligens ha store konsekvenser for ikke bare den umiddelbare bruken på slagmarken, men også bredere for sikkerhetspolitikk. Siden vi er geografisk plassert der vi er og har den naboen vi har, er det viktig å tenke gjennom hva disse mer autonome systemene har å si for våre sikkerhetspolitiske grunnpillarer som avskrekking og beroligelse. Systemer som lover transparens, presisjon og slagkraft uten sidestykke er ikke nødvendigvis spesielt gunstig for Norge. Når vår avskrekking er garantert gjennom NATO og underbygget av USAs atomvåpen, så virker en lovnad om transparens, der ingenting kan gjemme seg, kombinert med fart å utfordre denne skjøre balansen. Denne balansen hviler på at både USA og Russland vil ha andreslagskapabiliteter skulle en atomkrig bryte ut. Uavhengig om disse systemene vil virke slik de lover, vil bare frykten for at disse kapabilitetene skal kunne tas ut være nok til å få situasjonene til å tippe over. Selv om det av militær nødvendighet eller rasjonalitet virker fornuftig å utvikle disse systemene, så er det også mye som tilsier at vi bør gå tilbake til Clausewitz' advarsel om at krig og krigføring er så bundet opp i tilfeldigheter at vi bør gjøre vårt ytterste for å ikke komme dit. Gitt kortere og kortere responstider gir dette også økte muligheter for

feiltolkninger og forhastede slutninger gjort av menneske-maskinsystemene, med de følger det måtte ha. Derfor er spørsmålet om kontroll ytterst viktig. Dette fordrer på sin side at Forsvaret tar en aktiv rolle i å skjønne og formulere hva dette betyr og at de utvikler systemer hvor de tar innover seg at kontroll er distribuert – mellom mennesker, mellom mennesker og maskiner, mellom maskiner og i systemene som helhet – og at denne kontrollen skjer i mange faser gjennom hele beslutningsprosessen og ikke bare i den utøvende fasen.

Det å bytte ut mennesker med maskiner har både positive og negative sider. Det kan frigjøre arbeidskraft til viktigere oppgaver, men det kan også føre til at vi ikke forstår de valgene som danner grunnlag for avgjørelser. Vi må derfor gjøre oss bevisst hva som skjer når vi utvikler teknologier. De vil forandre hvordan vi ser på verden, de beslutninger vi tar, og hvordan vi handler, og de vil gjøre det gjennom å ekskludere alternative måter å gjøre dette på. Ved å bevisstgjøre oss disse distribuerte dynamikkene kan vi få et bedre og mer produktivt bilde av hva kontroll er og hvordan vi kan «sikre» det.

Bibliografi

- Bornstein, A. M. (2016, 01.09.2016). Is Artificial Intelligence Permanently Inscrutable?(40). Retrieved from <http://nautil.us/issue/40/learning/is-artificial-intelligence-permanently-inscrutable>
- Bourne, M. (2012). Guns don't kill people, cyborgs do: a Latourian provocation for transformatory arms control and disarmament. *Global Change, Peace & Security*, 24(1), 141–163. doi:10.1080/14781158.2012.641279
- Correll, J. T. (2004). Igloo White. *Air Force Magazine*, 87(11). Retrieved from <https://www.airforcemag.com/article/1104igloo/>.
- Cummings, M. (2004). Automation Bias in Intelligent Time Critical Decision Support Systems. *AIAA 1st Intelligent Systems Technical Conference*. doi:10.2514/MISC04
- Echevarria, A. J. (2004). *Toward an American Way of War*. Retrieved from www.jstor.org/stable/resrep11199
- Edwards, P. N. (1996). *The closed world: Computers and the politics of discourse in Cold War America*. Cambridge, Mass.: MIT press.
- Ferry, M. (2013). F3EA - A Targeting Paradigm for Contemporary Warfare. *Australian Army Journal*, 10(1).
- Forsvarsstaben. (2014). Forsvarets fellesoperative doktrine (FFOD). Retrieved from <https://fhs.brage.unit.no/fhs-xmlui/handle/11250/224031>
- Hayden, M. V. (2016). *Playing to the edge: American intelligence in the age of terror*. New York: Penguin Press.
- Healey, J. (2019). The implications of persistent (and permanent) engagement in cyberspace. *Journal of Cybersecurity*, 5(1), 4. doi:10.1093/cybsec/tyz008
- Kania, E. B. (2017). Battlefield Singularity. *Artificial Intelligence, Military Revolution, and China's Future Military Power*, CNAS.
- NATO. (2016). Allied Joint Doctrine for Joint Targeting (AJP-3.9). (1).
- Nordin, A. H. M., & Öberg, D. (2015). Targeting the Ontology of War: From Clausewitz to Baudrillard. *Millennium: Journal*

- of International Studies*, 43(2), 392–410.
doi:10.1177/0305829814552435
- Roff, H. M., & Moyes, R. (2016). Meaningful human control, artificial intelligence and autonomous weapons.
- Scharre, P. (2016). Centaur warfighting: the false choice of humans vs. automation. *Temp. Int'l & Comp. LJ*, 30, 151.
- Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situated actions* (2. ed. ed.). Cambridge: Cambridge University Press.
- Tucker, P. (2017). The Future the US Military is Constructing: a Giant, Armed Nervous System. Retrieved from <https://www.defenseone.com/technology/2017/09/future-us-military-constructing-giant-armed-nervous-system/141303/>
- U.S.AirForce. (2019). *Annex 3-60 - Targeting*. Montgomery, AL: LeMay Center for Doctrine
- U.S.DoD. (2012). *Department of Defense Directive 3000.09: Autonomy in Weapon Systems*.
- van Creveld, M. (1989). *Technology and war: From 2000 BC to the present*: Simon and Schuster.
- Williams, A. P., & Scharre, P. D. (Eds.). (2016). *Autonomous systems: Issues for defence policymakers*. Norfolk, Va.: HQ Sact.



Norsk Utenrikspolitisk Institutt

Norsk Utenrikspolitisk Institutt [NUPI] ble etablert i 1959, og er et ledende, uavhengig forskningsinstitutt på internasjonal politikk og områder av relevans for norsk utenrikspolitikk. Formelt er NUPI underlagt Kunnskapsdepartementet, men opererer likevel som en uavhengig, ikke-politisk virksomhet i alle sine faglige aktiviteter. Forskning utført ved NUPI spenner fra kortsiktig anvendt forskning til mer langsiktig, grunnforskning.

Om forfatteren:

Erik Reichborn-Kjennerud er forsker i forskningsgruppen for sikkerhet og forsvar ved Norsk Utenrikspolitisk Institutt. I tillegg til forskning ved NUPI er Erik doktorgradsstipendiat ved Department of War Studies, King's College London, der han studerer moderne militære målretningsmetoder og praksis. Hans forskningsinteresser omfatter militærteori og strategisk tenkning, moderne vestlig krigføring, teknologi, kritiske sikkerhetsstudier, IR-teori og Science and Technology Studies.

NUPI

Norsk Utenrikspolitisk Institutt
C.J. Hambros plass 2D
Postboks 8159 Dep. 0033 Oslo
www.nupi.no | post@nupi.no