

→ Offentlig etik: Ska teknikjättarna stötta det nationella försvaret?¹

Isaac Taylor (översättning Peter Samuelsson)

I SITT AVSKEDSTAL 1961 varnade president Dwight D. Eisenhower för faran med det militärindustriella komplexets »otillbörliga inflytande».² Detta var en referens till det framväxande samarbetet mellan privata företag och de offentliga amerikanska försvars- och säkerhetssektorerna. I den händelse vi behöver en påminnelse om hur avgörande privata företag har kommit att bli i detta partnerskap räcker det att tänka på hur den amerikanska regeringen nyligen valde att klassa Anthropic som en »risk i leveranskedjan» efter att bolaget vägrat ge det amerikanska försvarsdepartementet obegränsad tillgång till dess AI-system.

I ett uttalande beskrev Anthropics vd Dario Amodei företagets »röda linjer», det vill säga vilka användningar av dess produkter som man inte var beredda att bidra till. En sådan användning var autonoma vapensystem.³ Även om begreppet »autonomt vapensystem» har karaktäriserats på olika sätt refererar det grovt sagt till vapensystem som utan direkt mänsklig inblandning förmår välja ut mål som det sätts in mot.⁴

Visserligen existerar helt autonoma vapensystem inte enligt vissa definitioner, men AI har redan integrerats i vapen på ett antal sätt. Till exempel har Israel i sina operationer i Gaza använt sig av system såsom »The Gospel» och »Lavender»,⁵ och USA har inlemmat AI-verktyg som utvecklats i det så kallade Project Maven⁶ för att hjälpa till att identifiera militära mål i Iran.

Privata aktörers ovilja att ge sitt stöd till militära syften är inget nytt. Till exempel kom Albert Einsteins forskning att bidra till utvecklingen av atombomben. Men senare, efter att USA använt den på de japanska städerna Hiroshima och Nagasaki, uttryckte han ånger över att han manat på president Roosevelt att bygga den.⁷

Anthropics ovilja att främja autonoma vapensystem kan på ett liknande sätt ses som en rädsla för utvecklingen av tvivelaktiga vapen. Men den här händelsen väcker också ännu mer oroväckande frågor än om AI kan följa krigets lagar. Även om autonoma vapen som betar sig etiskt skulle kunna konstrueras finns risken när privata företag utvecklar dem att kontrollen över kriget förskjuts från demokratiska institutioner till Silicon Valley.

Etisk krigföring

AMODEIS TVEKAN HANDLAR framförallt om teknikens nuvarande nivå. Han hävdar att »de mest avancerade AI-systemen helt enkelt inte är tillräckligt pålitliga för att styra helt autonoma vapen».⁸ Dessa farhågor handlar inte bara om hur effektiva dessa vapen kan bli, utan huruvida de är i stånd att hålla sig inom de etiska gränser som vi anser gälla för soldater i väpnade konflikter.

I krig betraktas civila mål vanligen som förbjudet område: personer som gör civila och civil infrastruktur till mål brukar kallas »terrorister». Men skulle autonoma vapensystem på ett säkert sätt kunna skilja på soldater och civila? Kanske inte. En annan etisk gräns handlar om proportionalitet: krigets lagar specificerar att den förstörelse som orsakas av en krigshandling inte får vara större än som motiveras av den militära nyttan. Men hur ska vi väga det otydliga begreppet militär nytta mot de kvantifierbara dödsfall som orsakas av en viss militäroperation? Det antas allmänt att existerande AI-system inte har förmåga till den sortens kontextkänsliga omdömen som krävs för att tillämpa denna princip.⁹ Naturligtvis kan man även fråga sig om vi kan lita på att människor alltid följer sådana principer, även om de som argumenterar så skulle hävda det. Maskiner, i alla fall som de ser ut i dag, saknar fullständigt den relevanta förmågan.

Sådana farhågor innebär naturligtvis inte ett absolut hinder för en etisk användning av autonoma vapensystem. Vi kanske bara behöver vänta på att bättre AI-system ska dyka upp. Det tycks vara Amodeis ståndpunkt: han hävdar att helt autonoma vapensystem »kan visa sig vara avgörande för vår nationella försvarsförmåga».¹⁰ Om vi en vacker dag lyckas bygga AI-system som har förmåga att följa krigets lagar försvinner den invändningen.

Demokratiska stater och krig

ANTA ATT ANTHROPIC skulle kunna bygga autonoma vapensystem som är jämförbara med eller överträffar en genomsnittlig amerikansk soldats förmåga att agera etiskt. Ska de då användas? Mer principiella invändningar har gjorts mot att dessa vapen används med tanke på vad som ser ut som en »ansvarslucka» som då skulle uppstå. När en maskin fattar ett beslut om vem som ska dödas i krig tänker man sig ibland att någon borde kunna ställas till svars för detta beslut. Men eftersom ingen kan förutsäga eller kontrollera hur autonoma vapensystem beter sig och eftersom systemen själva inte är den sorts aktörer som är föremål för ansvarsutkrävande, hamnar vi i en situation där ingen kan ställas till svars.¹¹

Praktiker och etiker har bemött dessa farhågor genom att argumentera för att organisatoriska eller tekniska interventioner kan garantera att mänskligt ansvar upprätthålls. Sådana skulle göra det möjligt för människor att på ett bättre sätt förutsäga och kontrollera hur de autonoma vapensystemen uppträder och på så sätt bli ansvariga. Jag ska här inte ta ställning till om sådana lösningar kan fungera.¹² Det jag vill fokusera på är något som man ofta bortser ifrån: att det kanske spelar roll *vem* som har ansvar.

I demokratier tänker man ofta att politiska handlingar utförs »i folkets namn». Lagar stiftas visserligen av politiker och verkställs av dem som arbetar i den offentliga sektorn men i välfungerande demokratiska samhällen är de inte framtvungade av en liten grupp individer som försvarar sina egna privata intressen, utan de är i stället något som tas fram kollektivt.

Samma sak kanske kan sägas om krig. När soldater utkämpar ett krig gör de det i vårt namn. Om deras handlingar ska legitimeras måste de ytterst sett kunna föras tillbaka på oss: det är den centrala tanken bakom principen om medborgarnas kontroll över militären. Vi må ha många institutioner som är till för att säkerställa att vi behåller kontrollen: militärutbildning, politiskt bestämda uppförandekoder och befälsordningar. Men vad händer när AI-system som tagits fram av privata företag får uppgifter som tidigare har getts till soldater?

Det har länge varit vedertaget att teknologier som AI inte är värde-
neutrala verktyg utan något som är beroende av implicita moraliska
omdömen.¹³ När ett företag tillverkar ett AI-system bör det vara med-

vetet om de värderingar som ska realiseras i dessa beteenden. Värderingar som är kopplade till etisk krigföring bör naturligtvis inkluderas. Men om vi vill behålla den demokratiska kontrollen över krig bör även värderingar i den bredare politiska gemenskapen beaktas. När ett privat företag verkar utanför normala demokratiska institutioner i utvecklingen av autonoma vapensystem löper vi risk att folkets vilja ersätts av det privata företagens värderingar.

Om företag som Anthropic så småningom bygger de här systemen kommer de inte bara att leverera verktyg. De kommer att bygga in sina egna värdeomdömen i krigets infrastruktur. Föreställ er till exempel ett autonomt målsystem byggt av ett företag i Silicon Valley. Dess konstruktörer måste bestämma hur försiktigt det ska vara när det skiljer mellan civila och stridande. Ska systemet slå till när det är till 60 procent säkert? 80 procent? 95 procent? Att svara på detta är omöjligt utan att ta ställning till vissa moralfrågor.

Det här kanske alla inte oroar sig för. Man kanske skulle kunna hävda att så länge det inte är AI som fattar beslut om att starta krig (och det beslutet ligger åtminstone för tillfället kvar hos människor) kan de beslut som fattas under kriget vad gäller beskjutning och liknande på ett passande sätt delegeras till AI-system. När allt kommer omkring delegeras sådana beslut till militär personal, som med tanke på den moderna krigföringens verklighet inte alltid kan hållas under strikt uppsikt. Likväl ställs den sortens beslut på vissa sätt ibland under politisk kontroll. Tänk till exempel på den strategiska vägledning som utfärdades under Obamas presidentskap, då dödligt våld i terrorismbekämpning endast tilläts när man »nära nog var fullständig säker» på att ha att göra med en terrorist.¹⁴ Jag skulle hävda att en tilltagande användning av AI utgör ett hot genom att vissa former av sådan politisk kontroll undermineras. Det innebär en betydande överföring av makt i fråga om att fatta beslut om liv och död.

Debatten om autonoma vapen formuleras ofta som en teknisk fråga: kan man få AI att följa krigets lagar? Men den djupare frågan är politisk. Om de system som bestämmer hur krig utkämpas konstrueras av privata företag så kan kontrollen över kriget långsamt glida ifrån de demokratiska institutionerna och hamna i företagens händer. Utmaningen är inte bara att bygga mer etiska maskiner utan att säkerställa att bestämmanderätten över att inleda krigshandlingar stannar

där den hör hemma: hos de politiska institutionerna och ytterst sett den allmänhet i vars namn kriget förs. →

Isaac Taylor är universitetslektor i praktisk filosofi vid Stockholms universitet. Hans forskning är inriktad på etiska frågor i relation till användningen av militärmakt och införandet av AI. Han är författare till The Ethics of Counterterrorism (Routledge, 2018).

Noter

1. Denna notis publicerades först som ett inlägg på *Public Ethics*, en blogg som ges ut av Stockholm Centre for the Ethics of War and Peace: https://www.publicethics.org/post/should-big-tech-support-national-defense#_ednref1. Översättningen har möjliggjorts tack vare finansiellt stöd från Stiftelsen Erik och Gurli Hultengrens fond för filosofi.
2. <https://www.archives.gov/milestone-documents/president-dwight-d-eisenhowers-farewell-address>.
3. <https://www.anthropic.com/news/statement-department-of-war>.
4. I en ofta citerad definition beskriver det amerikanska försvarsdepartementet autonoma vapensystem som ett »vapensystem som när det en gång aktiverats kan välja ut och slå till mot mål utan ytterligare inblandning av en [mänsklig] operatör». US Department of Defence, »Autonomy in weapon systems», Directive 3000.09, 2012, s. 13–14.
5. https://en.wikipedia.org/wiki/AI-assisted_targeting_in_the_Gaza_Strip.
6. https://en.wikipedia.org/wiki/Project_Maven.
7. https://archive.org/details/sim_newsweek-us_1947-03-10_29_10.
8. Se fotnot 3.
9. Dessa farhågor formuleras mycket tydligt i Noel Sharkey (2012), »The Evitability of Robot Warfare», *International Review of the Red Cross* 94(886), s. 787–799.
10. Se fotnot 3.
11. En klassisk redogörelse för den risken hittar man i Robert Sparrow (2007), »Killer Robots», *Journal of Applied Philosophy* 24(1), s. 62–77.
12. För en bedömning av sådana möjligheter se Isaac Taylor (2025), »Is Explainable AI Responsible AI?» *AI & Society* 40 (3), s. 1695–1704; Isaac Taylor (2024), »Collective Responsibility and Artificial Intelligence», *Philosophy & Technology* 37: 27, s. 1–18.
13. <https://faculty.cc.gatech.edu/~beki/cs4001/Winner.pdf>.
14. <https://obamawhitehouse.archives.gov/the-press-office/2013/05/23/fact-sheet-us-policy-standards-and-procedures-use-force-counterterrorism>.