FOI MEMO
Projekt/Project
Huvudstudie Ledning

Sidnr
1 (38)

Projektnummer
E12241

Uppdragsgivare
Försvarsmakten

FoT-område
Ledning och MSI

Författare
Erik Granath, Kristofer Hallgren

Datum
2022-02-09

Memo nummer
FOI Memo 7807

Artificiell intelligens och militära operationer

Innehållet är granskat och omfattar ingen information som är underställd exportkontrollagstiftningen

Innehåll

Förkortningar ... **3**
1 Inledning ... **5**
2 Om artificiell intelligens ... **6**
2.1 Historik ... 7
2.2 Tekniker för AI ... 9
2.2.1 Maskininlärning (ML) ... 9
2.2.2 Deep Learning ... 10
2.3 Angränsande trender ... 10
2.3.1 Internet of Things (IoT) ... 10
2.3.2 Cloud/Fog/Edge – paradigmen ... 11
2.3.3 Hårdvaruutveckling ... 12
3 Militär AI ... **13**
3.1 Förväntade nyttor ... 13
3.2 Utmaningar ... 14
3.2.1 Tekniska utmaningar ... 14
3.2.2 Taktiska och operativa utmaningar ... 15
3.2.3 Strategiska utmaningar ... 16
3.2.4 Etisk debatt ... 16
3.3 Viktiga säkerhetspolitiska aktörer ... 17
3.3.1 USA ... 17
3.3.2 Kina ... 20
3.3.3 Ryssland ... 21
3.3.4 Storbritannien ... 23
4 Framåtblick ... **25**
4.1 Accelererande beräkningskapacitet, singularitet och superintelligens ... 26
4.2 Utveckling av militära applikationer med AI ... 27
5 Diskussion ... **30**
Referenser ... **33**
Böcker ... 33
Rapporter ... 33
Artiklar ... 34
Websidor ... 35

Förkortningar

ACS	*Autonomy Core System* (i amerikansk kontext)
ACS	*Automated Control Systems* (i rysk kontext)
AFRL	*Air Force Research Lab*
AGI	*Artificial Generell Intelligence*
AI	*Artificial Intelligence*
ANI	*Artificial Narrow Intelligence*
ANN	*Artificial Neural Networks*
ASI	*Artificial Super Intelligence*
ASIC	*Application Specific Integrated Circuit*
ATS	*Airpower Teaming System*
AWCFT	*Algorithmic Warfare Cross-Function Team*
BDAA	*Big Data and Advanced Analytics*
CCW	*Convention on Certain Conventional Weapons*
CEW	*Cognitive Electronic Warfare*
CNN	*Convolutional Neural Network*
COP	*Common Operational Picture*
CPU	*Central Processing Unit*
DARPA	*Defense Advanced Research Projects Agency*
DL	*Deep Learning*
DNN	*Deep Neural Network*
DSTL	*The Defence Science and Technology Laboratory*[1]
DUI	*Defense Innovation Unit*
FCNN	*Fully Connected Neural Network*
FCS	*Feedback Control Systems*
GOFAI	*Good Old Fashioned AI*
GPT-3	*Generative Pre-Trained Transformer 3*
GPU	*Graphical Processing Unit*
HMT	*Human-Machine Teaming*
HPE	*Human Performance Enhancement*
IADS	*Integrated Air Defence System*
IoBT	*Internet of Battle Things*
IOC	*Integrated Operating Concept (Storbritannien)*
IoMT	*Internet of Military Things*
IoT	*Internet of Things*
ISR	*Intelligence, Surveillance, Reconnaissance*
ISTAR	*Intelligence, Surveillance, Target Acquisition, Reconnaissance*
JADC2	*Joint All-Domain Command and Control*
JAIC	*Joint Artificial Intelligence Center*

[1] Brittisk motsvarighet till FOI.

JEDI	*Joint Enterprise Defense Infrastructure*
JSOU	*Joint Special Operations University*
JTARR	*Joint Tactical Autonomous Resupply and Replenishment*
LANCA	*Lightweight Affordable Novel Combat Aircraft*
LAWS	*Lethal Autonomous Weapon Systems*
MCF	*Military-Civil Fusion Development Strategy*
ML	*Machine Learning*
MUM-T	*Manned-Unmanned Teaming*
NDMC	*National Defense Management Center (Ryssland)*
NLP	*Natural Language Processing*
PLA	*Peoples Liberation Army (Kina)*
PLAAF	*Peoples Liberation Army Air Force (Kina)*
RNN	*Recurrent Neural Network*
SCO	*Strategic Capabilities Office*
SNN	*Spiking Neural Networks*
SOAR	*Special Operations Aviation Regiment*
SSE	*Single Synthetic Environment*
UAS	*Unmanned Aircraft System*
UAV	*Unmanned Aerial Vehicle*
UCAV	*Unmanned Combat Aerial Vehicle*
UGV	*Unmanned Ground Vehicle*
USSOCOM	*United States Special Operations Command*
UUV	*Unmanned Underwater Vehicle*
WUC	*Work Unit Code*
XAI	*Explainable AI*

1 Inledning

Vad är artificiell intelligens (AI)? Vi omges idag av produkter som påstås använda AI och vara ”smarta”. Det finns smarta telefoner, självkörande bilar och autonoma drönare. Den första smarta telefonen var visserligen mer avancerad än tidigare modeller, men den var inte särdeles smart med dagens mått mätt. Hur smarta är egentligen dagens smarta system och finns det en militär nytta av dem? Flera länder bedriver intensiv forskning på området och det hörs varningar från forskare om riskerna med beväpnade autonoma system.

Föreliggande memo är ett resultat av den kunskapsuppbyggande verksamhet som projekt Huvudstudie Ledning bedriver för att förstå framtida möjligheter och utmaningar för svenska militära operationer. Syftet med memot är att ge läsaren en bakgrund till området och bidra med orienterande kunskap om AI. Utvecklingen inom området går väldigt snabbt och stora summor pengar investeras. Nästan dagligen kan vi läsa om nya forskningsrön, teknikdemonstratorer eller användningsområden för AI. Denna text utger sig inte för att vara allomfattande på området utan återger endast de stora penseldragen. Med tanke på den snabba utvecklingen som sker är det viktigt att hålla i åtanke att texten är en ögonblicksbild för hur forskningsområdet såg ut i slutet av 2021. Texten kan fungera som ett avstamp i diskussioner kring ämnet och ge inblick i vilka trender som är värda att studera närmare.

Inledningsvis försöker texten definiera AI och ge en kort historik av utvecklingen från att termen myntades 1956 till idag. I samband med detta diskuteras ett antal olika teknikområden och angränsande trender. Därefter ges en inblick i hur AI kan nyttjas i militär verksamhet och hur stormakterna (USA, Kina, Ryssland och Storbritannien) ser på AI med exempel på pågående och planerad utveckling. Slutligen sammanfattas ett flertal framåtblickande arbeten med en diskussion om framtida nyttjande av AI i militär verksamhet .

Memot behandlar inte svensk AI-utveckling, varken civil eller militär. Ekonomiska aspekter, såsom hur mycket olika länder eller organisationer investerar i AI-utveckling, eller kostnader associerade med utveckling av AI behandlas inte heller i memot.

2 Om artificiell intelligens

Det finns ingen enhetligt vedertagen definition av artificiell intelligens, men det finns många försök att beskriva vad som avses med begreppet. Exempelvis:

> ”The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.” (Encyclopædia Britannica)
>
> “The field of computer science dedicated to solving cognitive problems commonly associated with human intelligence, such as learning, problem solving, and pattern recognition.” (Amazon Web Services, u. å.)
>
> “…the study and design of intelligent agents (IAs). An IA is an autonomous entity capable of sensing, reasoning and acting in an environment.” (Luotsinen et al., 2019)

Ovanstående beskrivningar skiljer sig åt på olika sätt, men några gemensamma drag är att de nämner förmågan att *göra* något och att detta något kräver någon form av tankeverksamhet.

En aspekt av hur vi uppfattar intelligenta system är att vi konstant normaliserar de framsteg som gjorts. Således är det nuförtiden ovanligt att man häpnar över hur snabb Googles sökmotor är, eller hur en smartphone kan översätta talat språk till text. Denna förskjutning av vad vi förväntar oss av maskiner kan beskrivas med Teslers teorem:

"Intelligence is whatever machines haven't done yet“(Tesler, u. å.).

Tegmark (2017) problematiserar vad som avses med intelligens (oavsett biologisk eller artificiell) och konstaterar att det finns många konkurrerande definitioner av intelligens. Därmed blir definitionsfrågan kring artificiell intelligens än svårare att svara på. Tegmark utgår i sina resonemang utifrån följande definition:

Intelligens = förmåga att uppnå komplexa mål

Tegmark menar att med denna definition så går förmågan att jämföra mellan olika system. Till exempel så utklassar maskiner människor på alla typer av avancerade och i många fall komplexa beräkningar. Maskinerna är däremot skapade av människan som verktyg för att lösa dessa avancerade beräkningar åt oss som ett steg mot ännu mer komplexa mål. Människa har också möjlighet att lära sig nya förmågor, vilket maskiner (än så länge) är väldigt dåliga på. För att kunna differentiera mellan olika typer av artificiell intelligens (AI) använder sig Tegmark med flera av följande begrepp:

- *Snäv AI* (*Artificial Narrow Intelligence,* ANI) innebär förmågan att uppnå en begränsad uppsättning mål, exempelvis att framföra en bil eller spela ett spel.[2]
- *Artificiell generell intelligens* (*Artificial General Intelligence,* AGI) innebär förmågan att uppnå många olika kognitiva uppgifter lika väl som människor.[3]
- *Superintelligens* (*artificial superintelligence,* ASI) innebär generell intelligens på en nivå som vida överstiger mänsklig intelligens.

Enligt denna indelning är det framför allt *snäv AI* som avses i dagens diskussioner om AI i militära sammanhang.

[2] Snäv AI kallas även ofta *svag AI*, men de applikationer som avses är långt ifrån svaga, snarare tvärtom.
[3] AGI och även Superintelligens kallas ofta *stark AI*,.

2.1 Historik

Termen ”Artificiell Intelligens” myntades inför en sommarworkshop på Dartmouth College i USA 1956, men flera av de tekniker som fältet baseras på är äldre än så.

Under lång tid var det rådande paradigmet för AI-utveckling traditionellt programmerade system, med uttömmande, regelbaserade logiska instruktioner för olika händelser. För väl avgränsade och enklare problem fungerade detta bra. För öppna eller komplexa problem visade sig detta angreppssätt dock problematiskt. Generellt förlitar systemen sig på uttömmande sökning, och när mängden potentiella lösningar till ett problem blir för många, tar det helt enkelt för lång tid att hitta den bästa lösningen. Bland de system som utvecklades på detta sätt under 1960-talet finns många lyckade exempel, såsom *General Problem Solver* (ett program för matematiska bevis), *ELIZA* (en tidig chat-robot som låtsades agera terapeut) och en mängd spelprogram framtagna för att överträffa en mänsklig spelare i spel såsom dam eller schack. Detta paradigm refereras till ibland som *Good Old-Fashioned AI* (GOFAI) (Bostrom, 2014).

De initiala framgångarna skapade en stor teknikoptimism, där forskare och finansiärer fick överdrivna förhoppningar om vad AI-tekniken skulle kunna prestera, men många problem visade sig betydligt svårare att lösa än vad man trott, såsom till exempel maskinöversättning från ett språk till ett annat. Detta resulterade under 70-talet i en så kallad AI-vinter, där både Europa och USA drog ner på forskningsmedel för AI-forskning (Schuchmann, 2019). I början av 1980-talet satsade Japan på projektet ”femte generationens datorsystem”, med syftet att utveckla kraftfull hårdvara för AI. Detta drog åter igång en våg av AI-satsningar världen över, som under 1980-talet gav upphov till det som kallas *expertsystem*, särskild hårdvara och mjukvara framtagna för ett specifikt ändamål, till exempel som beslutsstöd inom medicinsk diagnostik, planering och hårdvarudesign. Mot slutet av 80-talet visade det sig att expertsystemen inte motsvarade förväntningarna, varpå en andra AI-vinter inföll (Bostrom, 2014; Schuchmann, 2019).

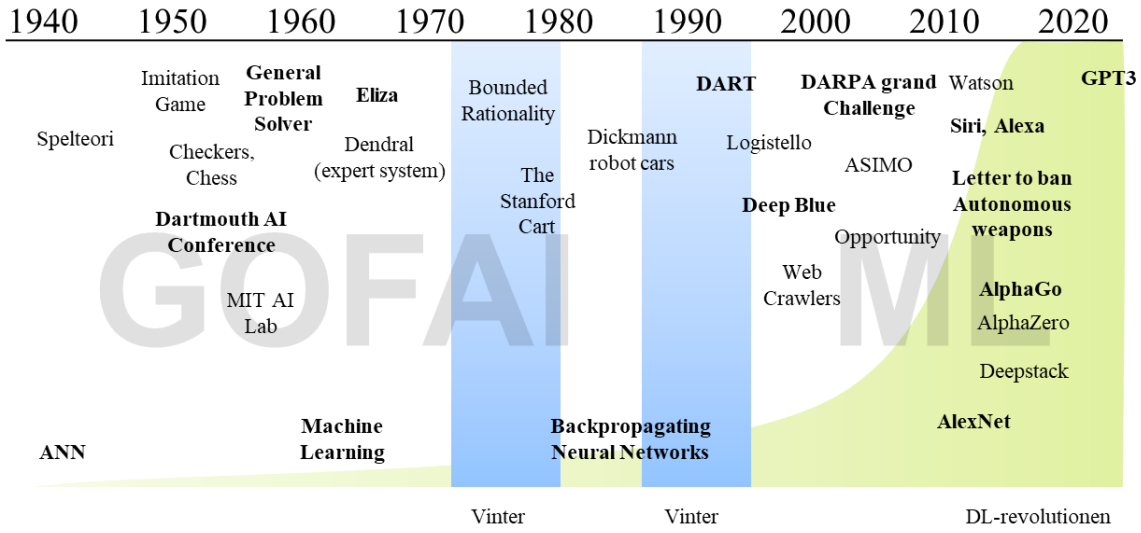
Under 1990-talet grydde i och med ett antal framgångar en ny AI-vår. Under operation *Desert Shield* och *Desert Storm* 1990-1991 använde USA ett system som kallades *Dynamic Analysis and Replanning Tool* (DART) för planering av logistik. Detta system ska enligt den dåvarande DARPA-chefen ha betalt tillbaka motsvarande 30 år av investeringar i AI (Hedberg, 2002). 1997 uppnåddes en viktig milsten för expertsystemen, då IBMs expertsystem *Deep Blue* lyckades besegra schackmästaren Kasparov.

Parallellt med utvecklingen av GOFAI-system baserade på logik och uttömmande sökning utforskades också i mindre omfattning andra paradigm för AI. På 50-talet dök enkla varianter av artificiella neurala nätverk (*Artificial Neural Networks*, ANN), så kallade perceptroner, upp. Tanken var att efterlikna neuroner i biologiska system och baserade sig på forskning från tidigt 40-tal, en rörelse som kallas konnektionism. Utvecklingen haltade, delvis på grund av begränsad hårdvara, delvis på grund av brist på effektiva algoritmer för hur nätverken skulle vara sammankopplade och arbeta. På 80-talet kom neurala nätverk med *backpropagation*, vilket kraftigt förbättrade ANN:s förmåga att lära sig av exempel (Kurzweil, 2005). ANN:er ingår i en större familj av probabilistiska metoder som vi idag kallar maskininlärning, *Machine Learning* (ML). Förutom ANN ingår här till exempel bayesiska nät, beslutsträd och olika regressionsmodeller. Under 2000-talet skedde ett paradigmskifte, där expert-system fick ge vika för olika ML-metoder.

Under 2000-talet bedrevs också flera parallella utvecklingar inom fordonsindustrin i syfte att ta fram självkörande bilar. 2004 introducerade DARPA tävlingen *Grand Challenge*, där universitet och företag fick tävla om att låta självkörande fordon ta sig genom 230 km ökenterräng. Det första året lyckades inga lag, men året därpå tog sig 5 av 20 lag i mål (Morgan et al., 2020).

Nästa stora utveckling skedde under 2010-talet. *ImageNet Challenge* var en klassificeringstävling där målet var att automatiskt klassificera bilder så korrekt som möjligt. Bidragen rankades sedan efter hur många fel de hade. *ImageNet* är en stor bilddatabas med uppmärkta klassificerade bilder föreställande allt möjligt. Till databasen finns det ett facit för vad som är en korrekt klassificering. År 2012 ställde ett lag upp med AlexNet, ett system som nyttjade ett så kallat *(deep) convolutional neural network* (CNN), där djupet avser att metoden använder ett neuralt nätverk som består av många lager. AlexNet presterade väsentligt bättre än konkurrenterna och året därpå använde samtliga toppresterande bidrag olika varianter av DNN (*Deep Neural Network)*. Sedan 2017 har det inte hållits fler tävlingar då alla deltagande system hade en lägre felprocent än 5 %. AlexNets framgång var genombrottet för *deep learning* (DL) och startskottet för den våg av AI-applikationer vi ser idag (Gershgorn, 2017).

DL används idag i de flesta applikationer för datorseende (*computer vision)*, till exempel i biometri-system, i självkörande bilar som kan se sin omgivning samt bildanalys och bildmanipulation. En viktig milstolpe för DL uppnåddes 2016 då Google Deepminds *AlphaGo* tävlade mot den regerande världsmästaren i Go och vann. Detta är en viktig händelse av två skäl. För det första är Go ett betydligt svårare spel att bemästra än schack (som expertsystemet Deep Blue lyckades med 1997), eftersom mängden möjliga drag är otroligt mycket större. Enligt det rådande paradigmet skulle ett expertsystem av GOFAI-snitt därför aldrig kunna bli bra på Go. För det andra överraskade detta Kina och dess militära tänkare och tvingade dem att ta AI på allvar (Huang, 2017; New York Times, 2017; Layton, 2021). I dag arbetar Deepmind med *AlphaStar*, ett AI-system som spelar realtidsstrategispelet *StarCraft II* på mästarnivå mot mänskliga spelare. *StarCraft II* är ett betydligt svårare problem än Go. Spel som *StarCraft II* tillför ytterligare en nivå av komplexitet på så vis att i Go har spelarna fullständig information om vad som händer i spelet (detta gäller en stor mängd brädspel), men i *StarCraft II* måste spelarna aktivt söka information om motståndarens drag och ta beslut på ofullständig information. Deepmind menar att den utveckling som ligger bakom *AlphaStar* är applicerbar i många domäner (Deepmind, 2021).



Figur 1: En översiktlig historik över AI. Bild: Erik Granath, FOI.

Ett forskningsområde som varit en stor fokus för AI sedan begreppet myntades har varit *Natural Language Processing* (NLP), bland annat med syftet att låta människor interagera med maskiner på ett mer människovänligt sätt genom skrivet eller talat språk. DL har möjliggjort stora framsteg inom NLP och återfinns i samtliga digitala assistenter såsom *Siri*, *Alexa*, etc. för att kunna tolka talat språk (Wikipedia, 2021; Tegmark, 2017). Utöver språkförståelse kan NLP-AI också användas till att

generera naturligt språk. Ett modernt exempel är GPT-3 (*Generative Pre-Trained Transformer 3*) som forskningsorganisationen OpenAI släppte 2020. GPT-3 är baserat på ett DNN med 175 miljarder parametrar och är tränad med ett massivt träningsset bestående av 100-tals miljarder ord. Detta möjliggör för GPT-3 att med minimal input skapa trovärdiga skönlitterära texter i samma stil som en känd författare, eller att generera programmeringskod utifrån en önskan om en viss funktion (Heaven, 2021).

2.2 Tekniker för AI

Beroende på hur inkluderande man vill vara i sin definition av AI, så finns det många olika sätt att klassificera AI. Ett övergripande sätt att se på utvecklingen är enligt DARPA (2017) som tre vågor:

Handgjord kunskap (*Handcrafted Knowledge*) – denna våg innehåller all GOFAI[4] och innebär att människor skapar kunskapsrepresentationen för ett specifikt problem samt att maskinerna därefter gör bedömningar utifrån denna kunskapsrepresentation i en snäv problemdomän. System i denna våg kan inte lära sig på egen hand och är dåliga på att hantera osäkerhet. I denna kategori finns även *Feedback Control Systems* (FCS), som nyttjar reglerteknik tillsammans med logik för att åstadkomma intelligenta beteenden, framför allt i automatiserade system såsom autopiloter, eller styrsystem för robotar (Allen, 2020).

Statistisk inlärning (*Statistical Learning*) – här återfinns all ML, inklusive artificiella neurala nätverk I denna våg skapar människor statistiska modeller för specifika problem och låter sedan maskiner träna sig på stora mängder data för att finjustera modellerna. System i denna våg kan lära sig av nya exempel, de lämpar sig väl för klassificering av olika fenomen men är dåliga på att resonera och att förstå sammanhang.

Anpassning till sammanhanget (*Contextual Adaptation*) – i denna framtida våg skapar systemen förklaringsmodeller för fenomen i olika sammanhang och det är därigenom också möjligt att spåra och förklara hur systemet har kommit fram till ett visst resultat. Detta gör systemet mindre känsligt för ändrade sammanhang. För närvarande samexisterar applikationer från den första och den andra vågen, medan den tredje vågen fortfarande befinner sig i forskningsstadiet.

2.2.1 Maskininlärning (ML)

Den för närvarande dominerande vågen av AI enligt indelningen ovan är statistisk inlärning, eller maskininlärning (ML). Det finns många olika modeller för ML, såsom till exempel Bayesiska nät, beslutsträd, stödvektormaskiner, regressionsmodeller, genetiska algoritmer och artificiella neuronnätverk.[5] Det finns några huvudsakliga inlärningsstrategier för ML:

Övervakad inlärning (*Supervised Learning)* innebär att människor skapar ett dataset som är ”korrekt” uppmärkt. Tanken är att systemet får ett stort antal exempel från vilket systemet sedan ska kunna generalisera och dra slutsatser. Detta ger stor kontroll över inlärningsprocessen och kan ge hög kvalitet i klassificeringen, men är kostsamt då det krävs manuell konstruktion av ofta stora dataset, ju större, desto bättre (Luotsinen et al., 2019).

Förstärkningsinlärning (*Reinforcement Learning*) innebär att systemet får agera fritt, det vill säga utan kunskap om vad som anses vara ”rätt”. Systemet belönas efterhand utifrån hur väl dess handlingar leder mot ett bestämt mål av en så kallad belöningsfunktion (*reward function*). Det typiska exemplet för förstärkningsinlärning är AI som får lära sig att spela olika spel, där det är enkelt att skapa en belöningsfunktion utifrån spelets regler och mål i en väl förstådd och tydlig miljö. Det är betydligt svårare att skapa belöningsfunktioner för tillämpningar i den verkliga världen och en bristande

[4] Begreppet *Good Old-Fashioned AI* (GOFAI) används ibland för alla ”smarta” system som inte tillhör ML-kategorin, men egentligen avser GOFAI det paradigm för att uppnå intelligent beteende som baseras på symbolisk kunskapsrepresentation och logik.

[5] För syftet med detta memo räcker det att känna till att ML innefattar betydligt fler tekniker än ANN och de olika typerna kommer därmed inte beskrivas närmare.

belöningsfunktion kan medföra oförutsägbart beteende. Fördelen med förstärkningsinlärning är att systemet lär sig självt genom att prova sig fram, vilket därmed kräver mindre handpåläggning (Luotsinen et al., 2019).

Oövervakad inlärning (*Unsupervised Learning*) innebär att ett system på egen hand identifierar kategorier eller mönster i ej uppmärkt data och lär sig skilja dessa åt enligt en förutbestämd metrik, såsom till exempel multidimensionellt avstånd (Luotsinen et al., 2019). Denna teknik kan ge upphov till kategorier som kan vara svåra att förstå för människor och kan därmed ge oväntade resultat. En fördel med denna teknik är att den kan hjälpa människor att hitta nya mönster i ostrukturerad data. Ett exempel på tillämpningsområde för denna teknik är bedrägeridetektion i finanssystem (Allen, 2020).

Det finns utöver de grundläggande strategierna även ett antal varianter, såsom delvis övervakad inlärning (*Semi-supervised Learning*), som innebär en hybridapproach, där ett mindre set av uppmärkt träningsdata används tillsammans med ett mycket större set av icke-uppmärkt data, i syfte att nyttja styrkorna hos både övervakad och oövervakad inlärning (Allen, 2020).

2.2.2 Deep Learning

Den mest omtalade formen av ML och den som skördat många framgångar sedan genombrottet 2012 är *Deep Learning* (DL). DL avser den specifika klass av ML-system som använder sig av ANN med många lager (därav djupet) av artificiella neuroner, ett så kallat DNN. DL är en av många olika ML-tekniker, och de ovan beskrivna träningsstrategierna gäller även för DL-system. Alla neuronnät i DL-applikationer är inte likadana – en stor del av utmaningen med att applicera DL för ett givet problem är att konstruera ett DNN som är lämpat för problemet, för att därefter träna upp det med träningsdata. Det finns några olika större klasser av DNN:

Fully Connected (FCNN): alla artificiella neuroner är sammankopplade med alla i nästa lager.

Convolutional (CNN): faltningsnätverk lämpligt för analys av spatial data såsom bilder

Recurrent (RNN): för analys av temporal eller sekventiell data, exempelvis mänskligt tal.

I verkliga applikationer sätts flera olika DNN samman i ett större system som anpassas efter problemet som ska lösas. Ett DNN kan ha miljarder parametrar som behöver justeras genom träning (Luotsinen et al., 2019).

2.3 Angränsande trender

Tillgången till öppen källkod och utvecklingsramverk för AI-applikationer bidrar till att det pågår intensiv forskning och utveckling av nya algoritmer för AI, framför allt inom ML och DNN. En annan grundläggande förutsättning för ML är data av tillräcklig mängd, kvalitet, representativitet och variation (Allen, 2020). Det är delvis på grund av detta som ML blivit så populärt sedan genombrottet 2012, samt de konkurrensfördelar företag [6] som kontrollerar dessa datamängder åtnjuter vid vidareutveckling av teknikerna. Det finns numera stora mängder av data (i mindre mån uppmärkt träningsdata) tillgänglig i och med vår ständigt ökande uppkopplade tillvaro. En annan aspekt av AI:s framgångar på den civila marknaden handlar också om innovation – att använda existerande algoritmer i nya tillämpningar och i andra domäner (Lee, 2018).

2.3.1 Internet of Things (IoT)

En annan drivkraft, eller trend synergisk med AI, är det som kallas ”sakernas internet”. Detta begrepp avser trenden att allt mer konsument- och industrielektronik är uppkopplade mot internet och att den

[6] Exempelvis Alphabet (Google), Amazon och Meta (fd Facebook).

virtuella och den fysiska världen därmed flyter ihop.[7] Det som är gemensamt för alla IoT-system är att enheterna är uppkopplade och att systemet via enskilda enheters sensorer samlar in data som sedan skickas vidare för analys på olika nivåer. Om enheten har någon form av ställdon, så kan enheten på kommando också utföra en förändring i sin miljö, ett så kallat cyberfysiskt (*cyber-physical*) system. Även om konceptet i sig är lätt att förstå, så är miljön i vilket en IoT-enhet verkar ett lapptäcke av olika typer av trådlös kommunikation, med olika krav, standarder och protokoll. För att kunna hantera detta krävs en tillräcklig nivå av intelligens hos enheten för att kunna uppfatta och analysera sin omgivning i syfte att anpassa sig därefter. Därmed är AI en nödvändig komponent i IoT-system (Johansson et al., 2019).

2020 fanns det ca 8 miljarder IoT-enheter uppkopplade mot internet. En projektion visar att det kommer att finnas ca 25 miljarder IoT-enheter uppkopplade mot internet 2030 (Jay, 2021). En annan projektion pekar mot 500 miljarder IoT-enheter 2030 (NATO STO, 2020). Oavsett den faktiska mängden enheter, så kommer alla dessa enheter att generera (och i mindre mån konsumera) en enorm mängd data.

I militära sammanhang används begreppen *Internet of Battle Things* (IoBT), *Internet of Military Things* (IoMT)[8] eller *Defence Internet of Things*[9], för att beskriva snarlika koncept för hur IoT-teknik ska kunna användas på slagfältet. Kott (2016) beskriver hur det framtida slagfältet är tätt befolkat av en stor mängd olika saker som kan förnimma, kommunicera, agera och samarbeta med varandra och med människor – det kan vara verkansdelar, sensorer, vapen, fordon och bärbara enheter (*wearables*). För detta krävs smarta kommunikationsnätverk som autonomt och dynamiskt ändrar konfiguration efter förhållanden och vilka enheter som är anslutna. Vidare krävs att IoBT hjälper människorna i systemet att skapa överblick och förståelse (*sense-making*) i syfte att inte dränka användarna i information. IoBT innebär därmed också en sårbarhet som fienden kan nyttja med följdeffekten att IoBT-system har höga säkerhetskrav.

2.3.2 Cloud/Fog/Edge – paradigmen

Under 2000-talet blev molntjänster vanliga, där en användare i stället för att köra en tillämpning på en lokal enhet kan köpa en tjänst bestående av olika resurser, såsom lagring, hårdvara och mjukvara, ofta distribuerade över flera fysiska datacentra. Användaren använder ett enhetligt gränssnitt för tjänsten och behöver inte hantera några av tjänstens interna resurser. Användaren behöver inte ens förstå eller känna till vad de består av eller hur de organiserade; de blir ett ”moln”. Många AI-applikationer använder molnberäkningar som är alldeles för tunga för en enskild dator för att kunna leverera lösningar till konsumenter – till exempel Googles sökmotor. För utvecklare av ML-applikationer som vill kunna träna sina modeller, är det ofta ohållbart att själv investera i den hårdvara som krävs. Då blir de olika molnleverantörernas AI-lösningar ett mer rimligt alternativ. IBM, Google, Amazon och Microsoft erbjuder alla olika varianter av utvecklingsmiljöer för AI.

I och med den stora ökningen av internetanslutna mobila enheter (framför allt mobiltelefoner, men i ett större perspektiv alla IoT-enheter) och ett konstant ökande krav på bandbredd och snabb respons, så har *edge*-paradigmet blivit alltmer relevant. Detta innebär att en större del av beräkning och lagring förläggs på ”kanten” till molnet, antingen i noder i det lokala nätverket som ligger närmre slutanvändaren (*Fog*), eller direkt på klientenheten (*Edge*). Detta ger några fördelar mot en ren molntjänst:

[7] Det finns en mängd olika koncept som överlappar IoT, till exempel *pervasive* eller *ubiquitous computing*, där den gemensamma nämnaren är att allt omkring oss blir allt mer ”smart” med hjälp av små datorer och uppkopplat genom olika varianter av trådlösa nätverk.
[8] Denna term används huvudsakligen i USA av amerikanska försvarsdepartementet.
[9] Denna term används av European Defence Agency.

- Fördröjningen för användaren minskar ju närmre användaren beräkningarna sker.
- Data som genereras hos användaren kan filtreras, valideras och optimeras innan den överförs till molnet, vilket innebär att nätet avlastas då användaren inte behöver kommunicera all data fram och tillbaka till molnet.
- I och med ökade beräkningsresurser närmare användaren, så kan användaren också använda tjänsten (om än på ett begränsat sätt) även i de fall kommunikationen med molnets övriga resurser är nere för tillfället. Detta är viktigt för mobila enheter som riskerar att tappa kontakten med nätet.
- Om data krypteras lokalt innan den kommuniceras vidare, minskar det risken för säkerhetsluckor (Digiteum, 2020).

För AI-applikationer innebär detta att en AI-applikation kan köras direkt på den lokala enheten. Ett exempel på detta är ansiktsigenkänning på en Iphone – om denna endast hade varit tillgänglig som en molnapplikation, så hade telefonen behövt skicka en videoström över nätet, där en molntjänst sedan hade utfört klassificeringen, för att sedan skicka tillbaka resultatet, vilket hade inneburit en ökad belastning på nätet samt väntetid för användaren. Av detta skäl utvecklas för närvarande en hel konsumenthårdvara specifikt för AI-tillämpningar (Durmaz, et al., 2020). *Cloud/Fog/Edge*-paradigmen bör ses som ett kontinuum, där olika applikationer, eller olika delar av en applikation, hanteras i molnet eller på kanten.

2.3.3 Hårdvaruutveckling

En förutsättning för den snabba utveckling av AI vi ser idag är att det sedan 2010 finns tillräckligt kraftfull hårdvara i en rimlig prisklass för att kunna köra de krävande algoritmer[10] som ligger till grund för ML. Det är framför allt utvecklingen av kraftfulla grafikkort (*graphical processing unit*, GPU) för speldatorer som öppnat upp möjligheterna för AI-utveckling. Traditionella processorer (*central processing unit*, CPU) lämpar sig bäst för snabba sekventiella beräkningar, vilket gör dem ineffektiva för applikationer som kräver parallella beräkningar såsom ANN. GPU:er är däremot massivt parallella – de är framtagna för att kunna genomföra en stor mängd beräkningar samtidigt. Av detta skäl har de börjat användas till andra syften än grafik, såsom beräkningar till ML. Exempelvis används en stor mängd GPU:er i de datacentra som hanterar beräkningarna för molnbaserade AI-applikationer.

GPU:er passar väl i datacentra, där energiförsörjning och kylning går att effektivisera, men är inte lämpade för mobila applikationer där storlek, vikt, energieffektivitet och termiska egenskaper är avgörande. På grund av detta sker en utveckling där mobila enheter (exempelvis mobiltelefoner) utrustas med hårdvara som är särskilt optimerad för att kunna implementera neuronnätverk, så kallade AI-acceleratorer. En AI-accelerator är en hjälpprocessor, liknande de integrerade grafikprocessorer som finns i många moderna processorer. AI-acceleratorer finns i samtliga moderna smartphones och används där för att till exempel identifiera ansikten, försköna bilder, etc.

En mer energieffektiv variant av en processor är en ASIC (*Application Specific Integrated Circuit)*. Här är hårdvaran helt optimerad för endast en specifik tillämpning. Så kallad *neuromorfisk* hårdvara är en utvecklingstrend där en ASIC används för att efterlikna biologiska neuroner. En neuromorfisk ASIC är därmed optimerad för tillämpningar som använder neurala nätverk (ANN/DNN). Detta innebär att tillämpningen kan bli extremt energieffektiv i jämförelse med traditionell hårdvara. Neuromorfisk hårdvara kan också implementera *spiking neural networks* (SNN), som är en vidareutveckling av ANN. Den stora fördelen med SNN är att de artificiella neuronerna i nätet ger en utsignal så fort villkoren för insignaler är uppfyllda, helt parallellt och helt oberoende av systemets klockfrekvens. Detta innebär att SNN:er är väldigt energieffektiva och samtidigt snabbare än ANN implementerade på en CPU eller en GPU. Av dessa skäl tros SNN och neuromorfisk hårdvara kunna lösa hårdvarubegränsningarna som idag hindrar utvecklingen för att träna ett system på ”kanten” (*edge*) (Durmaz et al., 2020).

[10] Till exempel GPT-3 (se 0 Historik), där 175 miljarder parametrar ska justeras för varje nytt träningsexempel.

3 Militär AI

Till skillnad från många tidigare tekniska utvecklingar som drivits av militära behov (exempelvis flyg- eller ubåtsteknologi) drivs AI-utveckling av kommersiella intressen på den civila marknaden och de centrala forskningsresultaten och mjukvaruteknologierna är allmänt tillgängliga. Därmed är det rimligt att förvänta sig att utvecklingen av militära applikationer för AI kommer att innebära en anpassning av tidigare civilt utvecklade lösningar (Morgan et al., 2020). I närtid kommer AI att användas för att förbättra existerande operationskoncept, men AI kan också komma att helt förändra hur militära operationer genomförs (Layton, 2021).

3.1 Förväntade nyttor

Det finns många förväntningar på vad AI ska kunna bidra med i en militär kontext. Här följer ett försök till klustring av de nyttor som presenterats i ett antal olika rapporter:

Analys av sensordata. Mängden sensorer som producerar data förväntas öka kraftigt. Det finns idag många applikationer där automatisk analys och klassificering av vad som syns i bild- och videodata ger mycket bra resultat (Morgan et al., 2021). På liknande sätt är det sannolikt att analys av även andra sensorer såsom radar, sonar, akustiska sensorer också kan automatiseras (Andersson et al., 2019). Aktörer med de bästa sensorerna, och tillgång till relevant data i rätt tid kommer att med hög säkerhet kunna positionera motståndare och icke-kombattanter vilket ger ett informationsöverläge (DCDC, 2018).

Att hantera stora mängder data. Mängden sensordata och information som kommer att flöda i framtidens ledningsstödssystem är alldeles för stor för människor att hantera. Det kommer därmed att krävas AI för att analysera och prioritera data och dataflöde i syfte att uppnå informationsöverlägsenhet (Gustavi et al., 2019). Vidare presterar AI generellt bättre ju mer information som finns tillgänglig.

Beslutsstöd. Det finns stor potential att förbättra existerande beslutsstödsystem med hjälp av AI, exempelvis genom att stödja arbetet med att ta fram tidskritiskt beslutsunderlag. Här ingår beslutsstöd i form av gemensam lägesbild, förutsägelser om fiendens agerande, samt generering av handlingsalternativ (Brynielsson et al., 2018). Även planeringsstöd till operativ och taktisk nivå samt administrativt stöd till staber har möjlighet att bidra till en snabbare beslutsprocess (Gustavi et al., 2019). Utöver detta kan AI användas även till automatiserad planering och resursallokering samt till analys av alternativ vid förmågeutveckling (Kindvall & Tarras-Wahlberg, 2021).

Ökad hastighet i beslutsfattande. Detta avser till del möjligheten att accelerera beslutsfattande genom att snabbare kunna hantera flödet av information. Men nyttjande av AI i tidskritiska situationer kan bli mer intressant efterhand som vapen i form av hypersoniska robotar och liknande medför krav på en väldigt snabb OODA-loop. Detta kommer sannolikt att driva på utvecklingen av AI-kontrollerade försvarssystem. (DCDC, 2018; Morgan et al., 2021).

Autonoma rörliga system kan användas till uppgifter som är *dull, dirty and dangerous* och därmed minska risken för egen personal samt frigöra denna till andra uppgifter. Därtill kommer system med en ökande autonomi att kunna användas till mer och mer kvalificerade uppgifter i alla domäner. Exempel på detta är underhåll, logistik, minröjning, ISTAR[11], telekrig samt som vapenplattformar. En uppgift för autonoma system som flera aktörer utforskar är att operera i motståndarens A2/AD-områden (och andra högriskområden) där det bedöms för farligt för människor att verka i syfte att degradera eller förstöra motståndarens luftvärnssystem (Morgan et al., 2021; Kindvall & Lindberg, 2020).

[11] Intelligence, Surveillance, Target Acquisition, Reconnaissance

ISR. De möjligheter AI ger avseende analys av sensordata innebär en stor potential inom ISR-förmågan. Automatisk inhämtning, bearbetning och analys av data från sensorer i alla arenor ökar tillgången till underrättelseinformation av hög kvalitet (Morgan et al., 2021). Automatiserad avvikelsedetektion kan användas för att outtröttligt upptäcka det som avviker från normalbilden, till exempel för att upptäcka undervattensintrång. ”Smarta” spaningssensorer kan ersätta soldater för spaningsuppdrag över lång tid i motståndarkontrollerade områden (Gustavi et al., 2019).

Cyberförsvar. AI skulle kunna användas dels defensivt, till automatiserad upptäckt och analys av intrångsförsök (Gustavi et al., 2018) dels offensivt till automatiserade påverkans- och cyberoperationer (Kindvall & Tarras-Wahlberg, 2021).

Informations/påverkansoperationer och vilseledning. På den civila sidan har vi de senaste åren kunnat se resultatet av ML-applikationer som analyserar inlägg och användarbeteende på sociala medier i syfte att rikta information till specifika mottagargrupper. Företaget *Cambridge Analytica* och deras påverkan på det amerikanska valet är ett uppenbart exempel på detta (Lee, 2018). Det har också blivit populärt med applikationer som givet en viss mängd träningsdata kan generera nya konstverk i samma stil som exempelvis van Gogh. Ett användningsområde som fått stor uppmärksamhet är spridning av falsk information med *deep fake*-videor där politiker och andra auktoriteter trovärdigt syns framträda i prekära situationer eller – i bästa fall – säga nonsens (Kindvall & Tarras-Wahlberg, 2021). Används denna teknik på ett subtilt sätt så är det numera i princip omöjligt för en människa att upptäcka att något inte står rätt till.

Effektivisering och optimering. Militära organisationer har mycket att vinna på AI-tillämpningar för att prioritera och optimera begränsade resurser, effektivisera underhåll och logistik (Kindvall & Lindberg, 2020), eller för att automatisera administrativa uppgifter (Andersson et al., 2019). Vidare kan AI frigöra personal, eller mildra brist på densamma, inom vissa kategorier såsom bild- och textanalys samt översättning (Morgan et al., 2021).

3.2 Utmaningar

I detta avsnitt presenteras ett antal utmaningar för utveckling och användning av AI på olika nivåer.

3.2.1 Tekniska utmaningar

ML, som för närvarande är det AI-paradigm som utvecklas och implementeras i en stor mängd olika applikationer, har redan en bevisad nytta och potentialen i att implementera ML i nya tillämpningar är enorm. Men ML har ett antal svagheter:

Systemets beteende styrs av träningsdata. Om träningsdata inte representerar den riktiga uppgiften på ett tillräckligt bra sätt så kommer systemet att göra (ur mänskligt perspektiv) överraskande fel, eller helt enkelt inte prestera som förväntat. Vidare måste träningsdata innehålla en tillräckligt stor variation av situationer och fenomen som det tränade systemet kan tänkas användas till. Om ett system till exempel tränas på att upptäcka fordon i en ökenmiljö, så är risken stor att systemet inte lyckas med detta i skog eller en arktisk miljö (Allen, 2020). När nya fenomen dyker upp som också måste kunna klassificeras, eller existerande fenomen ändrar karaktär, måste systemet tränas om för att inte hamna i konceptdrift; att klassificeringen blir sämre när omvärlden förändras (Layton, 2021).

Denna utmaning ställs på sin spets i en militär situation. Till skillnad från majoriteten av alla civila applikationer, så förväntas militära AI-system att agera i en miljö där motståndaren kommer att göra sitt yttersta för att *inte* upptäckas, *inte* agera förutsägbart och för att vilseleda. Detta ställer ytterligare krav på träningsdata. Med kunskap om hur ett system är tränat och vilken träningsdata som använts så går det att lura ett upptränat ML-system (Layton, 2021). Att få fram uppmärkt träningsdata av tillräckligt mängd och kvalitet är (och kommer troligen fortsätta vara) den främsta utmaningen på både kort och lång sikt för militära ML-tillämpningar (Gustafsson et al., 2020).

I det nuvarande sättet att träna en AI med DNN krävs utöver en stor mängd träningsdata också väldigt kraftfull (och energikrävande) hårdvara[12] (Gustafsson et al., 2020). Resultatet – det färdigtränade systemet – kan sedan köras på enklare hårdvara för tillämpningen. Men då kan systemet inte tränas om. Detta är ett av skälen till varför det är i dag svårt att låta DNN träna sig i fält, eller på ”kanten” *(edge)*. Tillräckligt kraftfull hårdvara med låg vikt/storlek/energiförbrukning finns helt enkelt inte tillgänglig (Durmaz et al., 2020).

Andra utmaningar är att hitta sätt att träna AI-system med betydligt mindre träningsdata, att dra generella slutsatser från en begränsad mängd information samt hur man ska kunna föra över lärdomar från en domän till en annan (Layton, 2021).

Till skillnad från vissa andra ML-tekniker såsom beslutsträd är DNN inte transparenta. Det går inte att förklara hur systemet kom fram till ett visst resultat. För en militär tillämpning blir detta särskilt problematiskt – en beslutsfattare måste kunna motivera sina beslut och en blind tillit till ett system vars sätt att arbeta inte går att förklara ingjuter inget förtroende. En utmaning är således hur system ska utformas så att deras beslut går att förstå och härleda för en människa. Av detta skäl är *explainable AI* (XAI) ett aktivt forskningsområde, sprunget ur DARPA:s projekt med samma namn. Syftet med XAI är att utveckla ”förklarbara modeller” för maskininlärning.

De svagheter som en AI baserad på dagens ML-paradigm har kan innebära att de ännu är olämpliga för en krigssituation. De kan vara alltför bräckliga för att uppnå en tillräcklig nivå av säkerhet i sina bedömningar, de uppvisar ett oförutsägbart beteende, eller så kan de vara alltför lätta att vilseleda (Morgan et al., 2021).

3.2.2 Taktiska och operativa utmaningar

Militära tillämpningar innebär hårdare krav än civila. Militära system verkar i ostrukturerade och antagonistiska miljöer, vilket innebär att civila system inte alltid går att anpassa för militära tillämpningar, eller att det medför alltför stora kostnader (Kindvall & Lindberg, 2020) Även för relativt konkreta uppgifter som inte är direkt kopplade till stridsuppgifter är bedömningen att fullt autonoma system ligger flera år framåt i tiden. (Ackerman, 2021)

AI kommer att innebära autonoma vapensystem med ett alltmer komplext och avancerat beteende. Att med övermänsklig precision och snabbhet kunna identifiera och verka mot mål på stora avstånd ger uppenbara fördelar i militära operationer, men krav på robusthet och tillförlitlighet gör att det kommer att dröja lång tid innan dessa vapen blir vanliga. Å andra sidan kan en aktör som inte ställer höga krav på tillförlitlighet snabbt utveckla avancerade autonoma vapen till en relativt låg kostnad redan idag (Gustavi et al., 2019). Eventuellt har den första helt autonoma attacken redan hänt, då *Kargu*-drönare utvecklade i Turkiet användes i en attack mot en motståndarfaktion i Libyen. Det är dock inte klarlagt huruvida drönaren var i fullständigt autonomt läge eller ej (Cramer, 2021).

Ett allt större inslag av obemannade och autonoma system ställer höga krav på framtida kommunikationslösningar, både avseende robusthet och signaturhantering, men också på bandbredd. Detta blir särskilt viktigt för svärmande system, där hela poängen med svärmen är koordinering och samarbete mellan de ingående enheterna. En trend för att lösa kommunikation på taktisk nivå är att använda ett dynamiskt nätverk där små autonoma drönare agerar länkar (Kindvall & Tarras-Wahlberg, 2021).

Autonoma system är kanske inte mer mottagliga än bemannade system för cyberangrepp, men konsekvenserna om angreppet lyckas kan vara betydligt större – därmed är det viktigt att vara medveten om detta i utformning och användning av autonoma system (Kindvall & Lindberg, 2020). Mindre autonoma system är också mer sårbara för motmedel såsom telekrigsvapen, mikrovågsvapen, laser eller automatkanoner (Kindvall & Tarras-Wahlberg, 2021).

[12] Se 2.3.3 Hårdvaruutveckling.

Beslutsfattare kan komma att lita i för hög utsträckning på AI-system och oreflekterat följa systemets rekommendationer. Denna risk ökar med system som har låg transparens och förklarbarhet (Morgan et al., 2021). Framför allt ML-baserade AI-system kan komma fram till väldigt annorlunda slutsatser mot vad en människa skulle göra, beroende på att de ”tänker” på ett fundamentalt annorlunda sätt (Gustavi et al., 2019).

3.2.3 Strategiska utmaningar

Att inte inse vilken potential och vilka förmågor som AI kan möjliggöra innebär att lämna fältet öppet för de aktörer som väljer att nyttja AI. På sikt kommer människa och maskiner att integreras allt tätare. De aktörer som är villiga att exploatera möjligheterna som AI ger kommer att få ett försprång. För att inte hamna på efterkälken krävs integration av industrisektorn, försvarsforskning och teknologiföretag. Det krävs också starkt ledarskap och genomtänkt investering i militär FoU. Att attrahera personal som kan utveckla AI för militära ändamål är (och kommer att vara) en strategisk utmaning för många stater, dels på grund av löneläget, dels på grund av att många utvecklare kommer att vägra arbeta med militära tillämpningar. Det är dock troligt att både Kina och Ryssland kommer tvinga inhemska bolag att utveckla militära applikationer, eftersom detta är en prioritet för dessa stater. Avtal och rustningskontroll behöver utvecklas för att även inkludera cybervapen, informationskrigföring och AI. För närvarande saknas internationella ramverk för teknik, etik och juridik för AI och behöver utvecklas. (DCDC, 2018)

Det finns flera risker med AI-vapen. För en makthavare kan tröskeln sänkas för ta till vapen om man har tillgång till autonoma system då risken för att spilla egna soldaters liv minskar. Detta kan leda till en större riskbenägenhet vilket i sin tur kan resultera i en eskalation. Utöver detta kan autonoma system som får agera på egen hand utan medvetenhet om politiska aspekter också öka risken för eskalation. (Morgan et al., 2021) Vidare är AI-teknik öppet tillgänglig och lätt att anpassa efter nya syften, vilket innebär att även icke-statliga aktörer kan få tillgång till AI för militära ändamål. (Andersson et al., 2019)

3.2.4 Etisk debatt

I och med den pågående utvecklingen av autonoma militära system så är det många som varnat för effekterna av och som arbetar för ett internationellt förbud mot LAWS – *lethal autonomous weapon systems*. Inför konferensen *International Joint Conferences on Artificial Intelligence Organization* (IJCAI) 2015 publicerades ett öppet brev riktat till FN, *Autonomous Weapons: An Open Letter From Ai & Robotics Researchers*, där författarna bland annat konstaterade att autonoma vapen kommer att vara morgondagens kalashnikovs och att autonoma vapen utom mänsklig kontroll måste hindras genom ett förbud mot offensiva autonoma vapen. Brevet är påskrivet av en mängd kända namn från forskarvärlden och industrin såsom Hawking, Musk, Dennet, Chomsky, Tegmark och Tallin (Future of Life Institute, u. å.). Ett annat debattinlägg som fått stort genomslag är filmen ”Slaughterbots” från 2017, producerad av *Future of Life*-institutet som en varning om vad som kan hända utan ett förbud mot LAWS. I filmen används svärmar av små bestyckade UAV:er av en ljusskygg organisation för att mörda politiska meningsmotståndare (Future of Life Institute, 2017).

Debatten om etiken kring AI och utvecklingen av autonoma vapen har troligen gett effekt. Ett exempel på detta kan ses i att Google drog sig ur utvecklingen av övervakningssystemet *Maven* 2018 efter protester bland de anställda. Kort därefter publicerade Google ett antal vägledande principer för utvecklingen av AI. Även andra viktiga aktörer inom industrin såsom DeepMind har tydligt tagit avstånd från militära applikationer av deras teknologi (Morgan et al., 2021).

FN:s *Office for Disarmament Affairs* (UNODA) har sedan 2013 organiserat multilaterala diskussioner om LAWS inom ramen för CCW, *the convention on certain conventional weapons*, där flera länder och organisationer vill att LAWS ska förbjudas i CCW. I diskussionerna har konceptet *meningsfull mänsklig kontroll* debatterats, där ett av argumenten som framförts är att när det gäller autonoma

vapen, så räcker det inte att garantera att en människa är i eller på beslutsloopen för systemet (*human-in-the-loop* eller *human-on-the-loop*). För att en mänsklig kontroll ska uppfattas som meningsfull bör den även gälla för hur systemet tränats och kvalitetssäkrats. Storbritannien presenterade 2018 i detta syfte ett ramverk för hur mänsklig kontroll kan utövas igenom alla faser i ett vapensystems fulla livscykel (Morgan et al., 2021). Hittills har USA, Storbritannien och Ryssland motsatt sig ett tillägg om LAWS till CCW, men 2019 publicerade länderna bakom CCW 11 vägledande principer för LAWS, där man bland annat kan läsa följande:

> Ansvar för att använda ett vapensystem kan inte läggas på en maskin, en människa är alltid ansvarig genom hela livscykeln för systemet. Riskhantering bör vara en del av hela utvecklingskedjan. Vid forskning kring och utveckling av nya vapen måste hänsyn tas till om användandet av vapnet kan komma att bryta mot internationell lag. Framför allt måste mänskliga rättigheter värnas. Vidare måste risken att systemet kan hackas eller vilseledas, samt risken att systemet kan spridas vidare och användas av terrorister beaktas. Vid utformande av policy bör teknologier inom LAWS-området inte förmänskligas (anthropomorphisized). (UN, 2019)

3.3 Viktiga säkerhetspolitiska aktörer

Forskning, utveckling och tillämpning av AI i militära syften sker för närvarande i en snabb takt världen över. De flesta stora militärmakter har dedikerade program för att utveckla och operationalisera olika fördelar som AI kan ge. I detta avsnitt beskrivs översiktligt utvecklingen i de fyra länder som just nu ligger längst fram inom AI, med avseende på ambitioner, övergripande strategi, pågående doktrin- och konceptutveckling, samt exempel på pågående utveckling.

3.3.1 USA

På senare år har USA etablerat flera nya organisationer för forskning och utveckling av AI. År 2015 startade USA:s försvarsdepartement upp en *Defense Innovation Unit* (DIU) i Silicon Valley, en organisation som syftar till att accelerera överföring och användandet av civil teknologi i den militära sektorn. AI, Autonomi och Cyber är tre av sex fokusområden för DIU (Defense Innovation Unit, 2021). 2018 startades *Joint Artificial Intelligence Center* (JAIC), med syftet att accelerera införandet av AI i alla organisationer under försvarsdepartementet, samt att koordinera olika AI-initiativ (JAIC, 2019).

USA har sedan 2018 en strategi för AI, där det bland annat fastställs en gemensam grund för data, ramverk, standarder samt moln/edge-tjänster för hela försvarsdepartementet. Ett annat spår gäller att utveckla implementering av och experimentering med AI genom utbildning och rekrytering av personal. USA ser också behovet av att etablera samarbeten med industri och akademi och uttrycker en vision av att vara världsledande gällande etik och säkerhet för militär AI (US DoD, 2018).

I och med skiftet från expeditionära insatser mot underlägsna motståndare till krig mot likvärdiga eller nästan likvärdiga motståndare (*near peer adversary*), håller USA på att utveckla nya koncept för krigföring. US Army har till exempel tagit fram ett koncept för multidomänoperationer, *The Army in Multi-Domain Operations*. I konceptet tas AI upp som en revolutionerande teknologi (tillsammans med robotik, överljudsvapen och nanoteknologi) som innebär fundamentala förändringar för krigföring.

DARPA har tagit fram ett koncept kallat *Mosaic Warfare*, i vilket AI och autonoma system nyttjas för att implementera så kallad beslutscentrisk krigföring (*decision-centric warfare*). Konceptet fokuserar på att en befälhavare ska ta snabbare och bättre beslut än motståndaren och därigenom hindra motståndaren från att nå sina mål. I konceptet ses AI och autonoma system som nyckelteknologier. Genom att nyttja dessa kan man åstadkomma ett mer distribuerat och dynamiskt styrke- och förmågepaket, där ledningsförhållanden och styrkesammansättning kan ändras och anpassas efter hand i syfte att agera snabbare än motståndaren. AI och autonomi ger en möjlighet att förbättra kommunikation och ledning i störda miljöer och som ett sätt att stötta uppdragstaktik (Clark et al., 2020).

Joint All-Domain Command and Control (JADC2) är ett ledningskoncept som innebär att kunna koppla ihop sensorer från alla grenar av USA militär. Det amerikanska försvarsdepartementets (US DoD, 2018) vision är en molnbaserad tjänst där all ISR-data ska flöda genom en stor mängd olika kommunikationsnätverk, bearbetas och analyseras av AI för att identifiera hot och generera lösningsförslag. Behovet av JADC2 kommer av att amerikanska försvarsdepartementet insett att framtida konfrontationer med nästan likvärdiga motståndare kommer att innebära multidomänoperationer i A2/AD-miljö. Därmed krävs en delad och uppdaterad situationsuppfattning för att kunna ta snabba beslut. Man har provat konceptet i övningar, en i USA 2019 där endast amerikanska enheter från flygvapen, armé och flotta ingick, samt en i Svarta havet 2020 där flottan, specialförband och åtta andra NATO-länder deltog (CRS 2021). Försvarsdepartementet gav 2019 uppdraget att ta fram en molntjänst för militära syften, *Joint Enterprise Defense Infrastructure* (JEDI), till Microsoft till ett värde av 10 miljarder USD. Beställningen avbröts dock 2021 och en ersättare, kallad *Joint Warfighter Cloud Capability,* ska tas fram med andra leverantörer (Burns, 2021).

DARPA har sedan 2018 arbetat med *AI Next Campaign,* ett initiativ som syftar till att utveckla ”tredje vågens” AI.[13] *AI Next Campaign* innehåller en mängd olika projekt indelade i fem huvudområden:

- Nya förmågor – här ges som exempel dynamiska verkanskedjor (*dynamic kill chain)* för multidomänoperationer, realtidsanalys av cyberattacker och detektion av vilseledande bildmaterial.
- Robust AI – inom detta område sker forskning för att förstå och förebygga feltillstånd i AI-system.
- AI som motståndare – i detta område undersöks bland annat hur AI kan vilseledas samt hur vilseledning kan motverkas.
- Hög prestanda – i detta område undersöks energieffektiviseringar, hårdvaruoptimering samt hur AI kan tränas med betydligt mindre dataset.
- Nästa generations AI – här ingår forskning på nya algoritmer för AI; hur AI ska gå från att vara ett verktyg till att vara en kollega samt hur AI-system ska kunna förklara sina resultat (DARPA, u. å.).

Av särskilt intresse är XAI[14], ett forskningsprogram för att komma tillrätta med att ML-system i dag i princip är svarta lådor, där det är i princip omöjligt att förklara hur systemet kommit fram till ett visst resultat. Tanken med projektet är att tilliten till AI-system som analytiker och beslutsfattare ökar om det är möjligt att förklara hur systemet har resonerat (Turek. u. å.).

Vid *Special Operations Forces Industry Conference* 2020 talade representanter för USSOCOM (*United States Special Operations Command*) om vikten av AI, framför allt till stöd för informationsoperationer. Andra exempel som togs upp var hur AI kan användas till fusion av information från en mängd olika sensorer och informationsflöden till stöd för analytiker samt hur UGV och UAV kan få en högre grad av autonomitet i syfte att avlasta operatörer. Vidare nämndes AI-baserad avvikelsedetektion för telekrig samt AI för anti-UAV-system (Harper, 2020).

I det senaste (2021) inriktningsdokumentet för forskning från Joint Special Operations University (JSOU) är ett av fokusområdena AI och *Big Data.* Området är i sin tur indelat i delämnen:

- Tillämpning för AI och *Big Data*. Här nämns bland annat funderingar kring vilka problem som lämpar sig eller ej för AI-tillämpningar samt vilka möjliga sårbarheter specialförband skapar genom att förlita sig alltför mycket på AI.

[13] DARPA räknar expertsystem som våg 1 och ML som våg 2.
[14] Detta är ett av forskningsprogrammen i DARPA:s ”tredje våg”. Se 2.2 Tekniker för AI.

- AI och *Big Data* för administrativa och logistiska uppgifter. Här nämns att specialförband gärna vill använda AI operativt, men att det sällan finns tillräckligt bra träningsdata. Av detta skäl kan det vara lämpligt att fokusera på administration och logistik, där stora mängder strukturerad data redan existerar.
- Gränssnitt och visualisering. Hur kan AI användas för att minska kognitiv belastning? I detta tema nämns också neural förbättring (*enhancement*) och etiken kring detta.
- Hur ska specialförbandsoperatörer utbildas och vilka strukturella förändringar krävs i syfte att åstadkomma en datasmart (*data savvy*) kultur inom specialförbanden? (JSOU, 2021).

3.3.1.1 Exempel på pågående utveckling

Project Maven är ett DNN för bearbetning och analys av videoströmmar från luftburna sensorer med automatisk klassificering av en mängd objekt samt möjlighet att följa fordon eller personer. Projektet inleddes 2017 av det då nyinrättade *Algorithmic Warfare Cross-Function Team* (AWCFT) i samarbete med Google. Syftet med AWCFT var att accelerera införandet av ML-teknologi i operativa system och sedan 2018 är AWCFT en del av JAIC (Gustavi et al., 2019). Projektet var en stor framgång, men på grund av etiska betänkligheter gjorde anställda på Google uppror, varpå Google hoppade av utvecklingen av *Maven.* Sedan 2019 utvecklas *Project Maven* av Palantir Technologies (Business Insider, 2019).

I det program som kallas *Mission Command System/Common Operational Picture* (COP) har USSOCOM valt att förnya ett kontrakt med Palantir Technologies som leverantör för AI-stödd ledning och *enterprise data management.* USSOCOM har använt Palantirs plattform sedan 2016 för att sammanställa information från en mängd disparata datakällor i en gemensam, global lägesbild. Systemet kan användas både under planering och under genomförande av uppdrag (Army Technology, 2021). Detta avser antagligen Palantirs *Gotham*, eller en variant därav. *Gotham* är en kommersiell produkt[15] för att hantera stora mängder data i ”nära realtid” från olika källor i en samlad lägesbild. I *Gotham* ingår olika ML/AI-baserade analysmodeller för beslutsstöd i olika situationer (Palantir, 2021).

USSOCOM och relaterade organisationer i USA har under de senaste åren lagt mycket energi på att ta fram eller anpassa AI-lösningar till stöd för specialförbanden. Inom ramen för *Joint Logistics mission initiative*, har JAIC tillsammans med *160th Special Operations Aviation Regiment* (SOAR) tagit fram ett AI-verktyg för prediktivt underhåll av regementets Blackhawk-helikoptrar. Verktyget, *Work Unit Code (*WUC) *Corrector* använder NLP, dels för att läsa existerande ostrukturerade underhållsrapporter i syfte att identifiera felaktiga koder, dels för att hjälpa tekniker med att mata in rätt koder i nya rapporter. Förutom att höja kvaliteten på underhållsinformationen, kan verktyget också ge underlag för att bedöma underhållsteknikernas duglighet (JAIC Public Affairs, 2020).

USA har länge experimenterat med obemannat stridsflyg, *Unmanned Combat Aerial Vehicles* (UCAV) och *Manned-Unmanned Teaming* (MUM-T), då amerikanska flygvapnet ser UCAV som ett sätt att minska kostnaderna och riskerna associerade med att använda modernt stridsflyg som F-35 eller F-22 i en konflikt mot en mer likvärdig motståndare. Ett av de senaste initiativen, *Skyborg*, är ett flygvapenprojekt lett av AFRL (Air Force Research Lab). Visionen för *Skyborg* är system av autonoma system i flera lager som kan agera tillsammans med mänskliga beslutsfattare i bemannade plattformar. (AFRL, u. å.). I *Skyborg* utvecklas en kärna av hårdvara och mjukvara, *Autonomy Core System* (ACS), som ska kunna användas för att flyga en mängd olika plattformar, med olika egenskaper och utrustning (Insinna, 2021). Exempel på dessa är Kratos *XQ-58 Valkyrie* och *UTAP-22 Mako*, General Atomics *MQ-20* samt Boeings *Loyal Wingman*. *Loyal Wingman* utvecklas av Boeing för Australiens flygvapen (RAAF) som *Airpower Teaming System* (ATS) (Newdick & Trevithick, 2021).

USA har flera program för utveckling av drönarsvärmar, såsom *Perdix* och *Gremlins*. *Perdix* är ursprungligen ett projekt från MIT där ingenjörsstudenter utvecklade idén om en svärm av små, billiga

[15] Palantir själv kallar det för ett operativsystem

drönare baserade på kommersiella komponenter. Projektet adopterades av *Strategic Capabilities Office* (SCO) och har sedan genomgått flera generationer av utveckling. 2016 genomfördes ett stort test där 104 *Perdix*-drönare släpptes av tre F/A-18 för att sedan själva organisera sig och genomföra vissa navigationsmoment på egen hand. Svärmens intelligens är distribuerad och självläkande, vilket innebär att även om ett antal enheter går förlorade, kan svärmen omorganisera sig för att fortsätta lösa sina uppgifter (Kallenborn, 2021).

Gremlins är ett DARPA-program som syftar till att ta fram återanvändbara drönare som kan avfyras och bärgas i luften. *Gremlins* är tänkta att bära olika typer av sensor- eller vapenlaster, till och med mindre drönare. Ett sannolikt scenario för hur *Gremlins* ska användas är att en taktisk transport rör sig utanför motståndarens A2/AD-bubbla och släpper en mindre svärm (ca 20 enheter) av *Gremlins* som sedan får genomföra uppdrag autonomt i motståndarens område (Hambling, 2021a).

3.3.2 Kina

Kina upplever att de hamnat på efterkälken gentemot Ryssland och framför allt USA vad gäller AI. Då Kina vill kunna dra nytta av utvecklingen inom AI satsar man stort på att hinna ikapp och helst ta över ledarpositionen. Kina avser att ha en fullt moderniserad militärmakt (Folkets Befrielséarmé, PLA) år 2035 samt att vara jämbördig med USA år 2050 (Dahm, 2020). Nicolas Chaillan, högste ansvarige för mjukvara i amerikanska flygvapnet avgick i oktober 2021 då han ansåg att den teknologiska förändringen av USA:s militär går för långsamt. Chaillans åsikt var att AI-kapplöpningen mot Kina redan var förlorad och att Kina kommer att vara ledande inom AI, syntetisk biologi och genetik inom ett decennium samt att USA inte kommer att ha en chans mot Kina om 15-20 år (Reuters, 2021).

2015 anammade Kina sin *Military-Civil Fusion Development Strategy* (MCF), som innebär att den civila sektorn involveras i allt högre grad i militär utveckling för att främja teknologiutveckling för dubbel nytta inom både civil och militär sektor. I detta syfte har två AI-centra skapats, vid Akademien för militärvetenskap och vid det nationella universitetet för militär teknologi (Office Of The Secretary Of Defense, 2020; Gustavi et al., 2019).

2017 lanserade Kina *Next Generation Artificial Intelligence Development Plan*, i syfte att se till att Kina är världsledande inom AI vid 2030. I strategin står bland annat att Kina måste fånga det strategiska initiativet i den internationella kapplöpningen för att skapa affärsmässiga fördelar och för att effektivt stötta nationell säkerhet (Chinese State Council, 2017).

I CCW-forumet så har Kina efterfrågat ett bindande protokoll för stark kontroll eller förbud mot LAWS. Men den definition som Kina föreslår på LAWS är alltför snäv för att vara meningsfull – det finns idag inga pågående initiativ till autonoma vapen som skulle kunna förbjudas med denna definition. Eventuellt handlar Kinas agerande i CCW om att de till synes vill vara för ett förbud, men i praktiken kunna fortsätta att utveckla autonoma vapen (Morgan et al., 2021).

Kina ser krig som en konfrontation mellan motsatta system, där kriget vinns genom att slå ut kritiska funktioner i motståndarsystemet genom informationsdominans, precisionsattacker och gemensamma operationer, tills motståndaren förlorar viljan till motstånd. För att uppnå detta ämnar PLA använda AI, människa/maskin-teaming och autonoma system (Dahm, 2020).

För PLA innebär den nuvarande teknologiska utvecklingen en utveckling från informationsdriven till intelligensdriven krigföring, där AI, *Big Data* och autonoma system blir tongivande. I intelligensdriven krigföring är AI tänkt att användas till bland annat automatisering av ledningssystem, prediktiv operationsplanering, ISR-fusion, autonoma farkoster, svärmintelligens och informationsoperationer (Office Of The Secretary Of Defense, 2020).

PLA konceptualiserar sig själv och motståndaren i framtidens konflikter som allestädes närvarande informationscentriska system-av-system, vilket möjliggör snabbare OODA-loopar. Detta påminner till del om DARPA:s *Mosaic Warfare*-koncept (Dahm, 2020).

PLA undersöker också ett antal nya operationskoncept som baserar sig på AI. Svärmkrigföring innebär decentraliserade operationer, där tillfällig kraftansamling med en stor mängd svärmande system ska övervinna motståndarens försvarsmekanismer. Vilande autonom överfallsstrid (*autonomous dormant assault warfare*) innebär överfallsstrid utförd av vilande autonoma system som kan ligga och vänta länge på rätt förhållanden eller aktiveras vid signal. Autonom rörlig tvärdomänskrigföring (*autonomous cross-domain mobile warfare*) innebär storskalig och långräckviddig autonom krigföring. Kognitiv kontroll innebär i princip att med hjälp av AI snabba upp sin egen OODA-loop och samtidigt försvåra för motståndaren att ta beslut (Dahm, 2020).

För regimen i Kina är AI ett verktyg inte bara för militär verksamhet, utan också för inre säkerhet. Regimen använder AI-beslutsstöd för övervakning och för att motverka social oro (Gustavi et al., 2019).

3.3.2.1 Exempel på pågående utveckling

Kina experimenterar för närvarande med en stor mängd autonoma system, såsom autonoma logistikfordon, ytfartyg och ubåtar. Men den för närvarande mest intressanta utvecklingen sker i utvecklingen av UAV och UCAV. Kina är den största exportören av UAV:er i världen, mycket beroende på att de är mindre nogräknade än USA i vilka de väljer att sälja till. Flera av de drönare Kina har som exportprodukter idag är delvis autonoma och kan troligen göras helt autonoma. Ett exempel är Blowfish A2, en beväpnad, mindre drönare som använder AI för att identifiera mål. Ett annat är GJ1/2 (Wing Loong), en UCAV som kan jämföras med USA:s MQ-1/9 (Kania, 2020).

Kina håller också på att utveckla olika varianter av smyg-UCAV i utvecklingsprogrammen *Dark Sword*, *Star Shadow* och *Sharp Sword*, som troligen kommer att kunna agera helt autonomt (Morgan et al., 2021). Kinas flygvapen (PLAAF) experimenterar antagligen likt USA och Ryssland också med MUM-T, där GJ11 *Sharp Sword* (en avancerad smyg-UCAV) antas vara experimentplattformen (Joe, 2019).

Kina utvecklar också olika typer av svärmande helikoptrar, quadkoptrar och flygplan, för nya förmågor inom ISR, kommunikation och verkan. Framöver planeras drönarsvärmar som kan avfyras och bärgas i luften (motsvarande USA:s *Gremlins*) som ska kunna genomföra spaning, telekrig och bekämpning av till exempel en hangarfartygsgrupp. Även mindre, burna system såsom den intelligenta ISR/attack-drönaren CH-901 är under utveckling (Morgan et al., 2021).

Även Kina ser AI som ett kraftfullt verktyg för fusionering av information från många olika källor till en gemensam lägesbild. Den kinesiska flottan har initierat flera projekt där ML-tekniker är tänkta att kunna användas för att analysera data från en mängd olika sensorer i olika miljöer (rymd, luft, sjö, telekrig) för att kunna autonomt kunna övervaka stora områden. Därtill har Kinas ministerium för nationellt försvar uttryckt att AI kan fungera som den perfekta digitala stabsofficeren till stöd för operationsplanering (Morgan et al., 2021).

Med tanke på Kinas fokus på att kunna slå ut noder i motståndarens system, så är sannolikt att de också satsar på AI i cyberdomänen (Morgan et al., 2021).

3.3.3 Ryssland

Ryssland ser motståndaren som ett system med flera lager vars funktion måste försämras så att motståndaren inte längre kan strida. I detta system inkluderas utöver militära styrkor även motståndarens samhälle. Därmed blir hybridkrigföring mot hela samhället ett sätt att degradera motståndarens system. Ryssland lägger större vikt vid att påverka motståndarens kognition än USA och Kina, vilket märks i begreppet reflexiv kontroll – att påverka motståndarens omvärldsuppfattning i syfte att få denne att agera i linje med rysk intention (Layton, 2021). I detta syfte är cyberkrigföring, telekrig, påverkansoperationer och vilseledning olika typer av tillgängliga verktyg som passar väl in i Rysslands agerande. Det är därmed sannolikt att Ryssland kommer att nyttja (och redan har nyttjat) AI i dessa verksamheter (Morgan et al., 2021).

Ryssland tog ett första övergripande nationellt grepp om AI i och med en konferens organiserad av Ryska försvarsdepartementet 2018. De 10 förslag som konferensen genererade resulterade sedan i fyra större program; Programmet för digital ekonomi, en färdplan för AI, en nationell AI-strategi samt ett federalt AI-projekt. Ryssland har de senaste åren lagt stor vikt vid utvecklingen av AI för militära syften, med en särskild tonvikt på informationshantering för beslutsstöd och autonomi. Skälet till detta är att den enorma mängden information som genereras av olika sensorkedjor och i cybermiljön kräver snabb AI-stödd analys i syfte att uppnå informationsdominans i militära operationer. Begreppen digitalisering och intellektualisering används för att beskriva utvecklingen mot en initialt informationsfokuserad, därefter AI-förstärkt militärmakt genom införandet av informationsteknologier på bred front (Edmonds et al., 2021).

I CCW-forumet är Rysslands inställning till LAWS att det förvisso är viktigt att det finns meningsfull mänsklig kontroll över autonoma vapen, men att det bör vara upp till varje enskilt land att bestämma hur den kontrollen ska ske. Ryssland motsätter sig ett förbud mot LAWS med motivationen att LAWS är illa definierat, samt att internationell humanitär rätt redan är tillräcklig för reglering av nya vapen (Morgan et al., 2021).

De primära riktningarna för rysk AI-förmåga är (Edmonds et al., 2021):

- Informationshantering, ledning och beslutsstöd. Exempel på detta är det ryska nationella försvarsledningscentralen (*National Defense Management Center*, NDMC)
- Autonomi. Här ingår alla UCAV, UAV, UGV, UUV, etc.
- Träning, logistik, underhåll och upphandling
- Varningssystem och luftvärn.

Den militär-industriella kommittén har satt målet att 30% av rysk militärmateriel ska vara automatiserad vid 2025 (Morgan et al., 2021).

3.3.3.1 Exempel på pågående utveckling

På strategisk nivå är det sannolikt att Ryssland har introducerat AI för informationsbearbetning och prediktion i NDMC. NDMC startades 2014 med syftet att kunna hantera all information gällande nationens säkerhet i fred och i krig (Morgan et al., 2021). Den prediktiva förmågan är tänkt att kunna agera varningsklocka för gryende väpnade konflikter, baserat på en mängd data kring truppansamlingar och -förflyttningar (Edmonds et al., 2021).

På operativ nivå strävar Ryssland mot att kunna integrera information från plattformar och sensorer från samtliga försvarsgrenar. Informationen hanteras i automatiserade kontrollsystem (*automated control systems*, ACS) på olika organisationsnivåer och samlas i ett gemensamt övergripande ACS. Detta kan troligen utgöras av Akatsiya-M, ett system för realtidslägesbild på operativ nivå. Enligt Ryska försvarsministeriet utvecklar man en gemensam plattform för informationstjänster. Både denna och ACS-konceptet bygger troligen på *cloud/fog/edge*-paradigmet. ACS med vissa AI-komponenter har redan testats i övningar (Edmonds et al., 2021).

På taktisk nivå finns AI i integrerade luftvärnssystem (*Integrated Air Defence System*, IADS) såsom Pantsir och S-500 i syfte att korta ner tiden i OODA-loopen och prioritera mellan hot med olika hastighet och banor (Edmonds et al., 2021). 2018 genomfördes ett test där S-300, S-400 och Pantsir tillsammans med andra radarsystem hanterades i ett gemensamt ACS (Morgan et al., 2021).

AI används också i ett system för terränganalys i syfte att förbättra försvaret av militärbaser mot drönare. Systemet identifierar den lämpligaste attackvägen för drönare så att motmedel kan planeras och är ett resultat av erfarenheterna från Syrien (Tass, 2020).

Status-6 är en autonom reaktordriven UUV under utveckling som är tänkt att kunna avfyras från en ubåt och därefter autonomt tyst och långsamt ta sig fram till sitt målområde. Väl framme vid området ska den ligga i vänteläge tills den aktiveras med en yttre signal. Vid aktivering ska den kunna gå över

i höghastighetsläge för att snabbt ta sig till målet och detonera en kärnladdning (Weintz, 2018). En särskild risk med detta system är att det är svårt att återkalla, med tanke på problemen associerade med undervattenssignalering.

Ryssland experimenterar också med svärmteknologi, men ligger sannolikt långt efter USA och Kina på detta område (Morgan et al., 2021). Det existerar flera koncept för hur autonoma system i svärm ska kunna användas i olika arenor, bland annat UUV för maritim patrullering, spaning och UCAV (Altius-U) för ubåtsjakt. I markarenan tänker sig Ryssland kombinerade svärmar av UGV, UAV och UCAV som särskilt lämpade för urbana miljöer.

I luftarenan finns det ett pågående MUM-T-projekt där en SU-57 kan agera tillsammans med en S-70 Okhotnik-drönare, samt utveckling av drönarsvärmar för en kombination av telekrig, ISR och vapenverkan (Edmonds et al., 2021).

3.3.4 Storbritannien

Storbritannien är det land i Europa som ligger längst fram avseende AI. Dels har de en lång tradition av ledande forskning inom datavetenskap, dels har de idag världsledande industriaktörer såsom ARM (Hårdvara) och Google Deepmind (se Historik) (Gustavi et al, 2019). Storbritannien har uttryckt ambitionen att vara en vetenskaplig och teknologisk stormakt (*…a Science and Tech Superpower…*) och att hålla minst en tredjeplats i världen vad gäller vetenskap och innovation. Vidare ämnar de befinna sig i framkanten på AI-forskning (DSTL, 2021a). Som ett led i denna ambition ämnar Storbritannien etablera ett AI-center med syfte att accelerera nyttjandet av AI över hela förmågebredden och i samtliga verksamheter (HM Government, 2021).

2021 publicerade Storbritanniens försvarsdepartement sitt *Integrated Operating Concept (IOC)*. Konceptet som presenteras innebär den största förändringen av hur man tänker kring användningen av militär makt på flera generationer. Konceptet beskriver hur Storbritanniens konkurrenter använder informationsmiljön till stor effekt under tröskelnivån för väpnade konflikt (*sub-threshold operations*) och är framtaget som ett svar på hur andra stater suddar ut gränsen mellan militära operationer och andra maktmedel i syfte att uppnå sina mål. I konceptet görs skillnad på operationer och krigföring, där det militära maktmedlet måste kunna användas även för operationer som ligger under gränsen för regelrätt krig. För att uppnå detta krävs integrering mellan flera domäner och genom samtliga organisationsnivåer i försvaret. Vidare krävs en övergripande integrering av hela nationen samt starka internationella samarbeten. (UK MoD, 2021c).

Till stöd för IOC har Storbritannien nyligen även tagit fram en ny digital strategi, där fokus initialt ligger på att ta fram en digital ryggrad (*a digital backbone*) som ska kunna koppla ihop sensorer, beslutsfattare och utförare på samtliga nivåer i syfte att åstadkomma integrering över alla domäner. Denna ryggrad kommer att använda en skalbar molnarkitektur och nyttja ”next generation networks” för att kunna verka globalt. Utöver den digitala ryggraden nämns en ”digital smedja” (*digital foundry*), ett ekosystem av akademier, utvecklare och innovatörer, där bland annat försvarets AI-center ingår. Det är denna ”smedja” som är tänkt att realisera potentialen i försvarets samlade data, bland annat genom att nyttja AI.

DSTL anordnade en virtuell konferens i juli 2021 - *Supercharging Science*, som syftade till att utveckla framtida försvarsförmågor tillsammans med akademi och industri. I sessionen om AI beskrev de tre primära områden inom vilka AI kommer göra stor skillnad för försvarsdepartementet:

- Datavetenskap – att extrahera insikter från data
- Intelligenta digitala system som verkar i informations- och cyberarenan
- Robotik och autonoma fysiska system.

Vidare presenterades en vision om hur AI kan användas för att uppnå ett multidomäns-ledningssystem för att kunna uppnå de behov av informationsdelning som IOC innebär.

3.3.4.1 Exempel på pågående utveckling

Under den marina övningen *Formidable Shield 21* testades DSTL-utvecklad AI för första gången under en övning med realistiska förhållanden. Systemet bidrog till att överljudsrobotar kunde upptäckas tidigare samt med ett beslutsstöd för hur hotet skulle mötas (DSTL, 2021a).

Under sommaren 2021 använde Storbritannien framgångsrikt ett AI-system i ett övningsmoment under operation *Cabrit* i Estland. Systemet ska kunna ge planeringsstöd i markarenan baserat på information om omgivningen och terrängdata. Detta ska ha enligt brittiska försvarsdepartementet ha bidragit till en förbättrad ledning (UK MOD, 2021).

Single Synthetic Environment (SSE) är ett beslutsstödsystem som *Improbable* håller på att utveckla för UK Strategic Command. SSE kan fusionera information från samtliga domäner och miljöer och ger beslutsfattare från taktisk till politisk nivå möjlighet att testa olika alternativ och planer i den virtuella miljön. SSE kan använda sig av en mängd olika analytiska verktyg, varav vissa använder AI/ML. (Improbable, 2020).

Storbritannien har på senare år satsat på att föra in obemannade system i armén, framför allt riktade mot logistik. Projekt *Theseus* syftar till att ta fram ett till stor del autonomt system för logistik på frontlinjen, JTARR (*Joint Tactical Autonomous Resupply and Replenishment*), bestående av ett logistikinformationssystem och ett antal olika UGV:er och UAV:er. För närvarande håller projektet på att utvärdera ett par olika UGV-varianter, Titan (bandgående) och Viking (Hjulgående) (White, 2021).

Storbritannien håller på att utveckla ett 6:e generationens stridsflygplan, Tempest, som är tänkt att kunna ersätta Typhoon och vara operativt 2035. Tempest ska kunna fungera både med en flygförare och autonomt. I den tänkta arsenalen för Tempest nämns bland annat svärmande system (Royal Air Force, 2021). Storbritannien har också en pågående UCAS-utveckling i form av *Lightweight Affordable Novel Combat Aircraft* (LANCA) som är tänkt fungera enligt MUM-T-konceptet tillsammans med Typhoon, F-35 och det framtida Tempest (UK MoD, 2021b).

4 Framåtblick

Det finns flera olika perspektiv på vad som kommer att ske i utvecklingen av AI framöver och i detta memo har endast en liten delmängd presenteras. Att uttala sig om framtiden är svårt och många drar sig för att tidsbestämma en viss utveckling, då risken är att förutsägelsen visar sig fullständigt felaktig. I ett arbete för amerikanska flygvapnet intervjuade tankesmedjan RAND 29 AI-experter om hur och när olika AI-teknologier kommer att introduceras i militär verksamhet och när det i så fall kommer att ske någon revolution av krigsföringen. Svaren var så spridda att författarna inte ville försöka beskriva en tidslinje i rapporten, men samtliga var överens om att AI kommer att i allt högre grad integreras i militära system (Morgan et al., 2020). De flesta är också överens om att den civila marknaden driver utvecklingen av AI, att militära applikationer (ofta) baserar sig på civila innovationer och introduceras med viss fördröjning mot den civila marknaden.

Det är svårt att sia om *när* olika innovationer och tillämpningar av AI kommer att dyka upp, men det är kanske lättare att resonera kring i vilken *ordning* det kommer att ske. Investeraren Kai-Fu Lee (Lee, 2018) menar att AI håller på att introduceras i den civila sektorn i fyra vågor[16], som skiljer sig åt genom vilka datakällor och inom olika sektorer de verkar.

Den första vågen kallas för Internet-AI (*Internet AI*) och startade redan i början av 2000-talet, men blev mainstream efter 2012. Den typiska applikationen för Internet-AI är rekommendationssystem, som analyserar konsumenters beteenden och åsikter baserat på den enorma mängd data som finns i sociala medier, analys av vilka sidor människor besöker och så vidare. I denna våg återfinns också tillämpningar i likhet med hur Cambridge Analytica påverkade val i USA och England med hjälp av data från Facebook.

Den andra vågen kallas affärs-AI (*Business AI*). Denna våg påverkar affärsvärlden och drar nytta av data som finns i företag såsom ekonomisystem, databaser, kundinformation, information om logistikkedjor och leverantörer, etc. i syfte att optimera företagens operation. Lee nämner applikationer inom finansvärlden och försäkringsbranschen, men också inom medicinsk diagnostik och juridik.

Den tredje vågen kallas perceptions-AI (*Perception AI*), i vilken Lee menar att AI ger maskiner förmågan att faktiskt förstå innehållet i data och möjligheten att uppfatta och förstå sin omvärld. I denna våg finns alla tal-till-text-system, smarta assistenter och övervakningssystem som till exempel Alibabas *City Brain*. Denna våg innebär att gränsen mellan den digitala och konkreta världen blir mindre skarp.

De första tre vågorna möjliggör den fjärde vågen – autonom AI (*Autonomous AI*), där AI inte bara kan förnimma, förstå och lära sig av verkligheten, utan även påverka och forma den för att uppnå komplexa mål. Lee menar att autonom AI initialt kommer att användas i strukturerade miljöer och situationer (såsom fabriker och lager), men att i den i takt med att tekniken mognar kommer att kunna användas i mer och mer ostrukturerade miljöer och situationer (Lee, 2018).

Det finns också röster som varnar för att AI-utvecklingen kan vara på väg att bromsa upp. I en sammanfattning av de största artiklarna inom AI som gjordes av den internationella ingenjörs och vetenskapssammanslutningen IEEE under hösten 2021 noteras att det finns flera indikationer på att AI inte lever upp till alla förväntningar och att det därmed finns risk för en ny vinter (Strickland, 2021). Detta uttalande vilar på en sammanställning av flera artiklar om AI av IEEE under 2021. De talar bland annat om att problemet med att träna AI-system med konventionell ML är att det är enormt beräknings- och energikrävande, AI-system är fortfarande sköra i betydelsen att de endast fungerar inom ett väldefinierat område och systemen är känsliga för indata under inlärningen – inkorrekt indata kan ge icke-önskvärt beteende hos det operativa systemet.

[16] Notera att dessa vågor inte har någon koppling till de vågor som DARPA ser inom AI-utveckling (kapitel 2.2).

Användningen av AI i form av ML och prediktiv analys inom vissa civila applikationer har börjat ifrågasättas då de ibland kan förstärka negativa trender, exempelvis ojämlikhet och missinformation. Det råder även en viss begreppsförvirring då AI och ML blandas samman med prediktiv analys. Båda har liknande användningsområden, och inget av dem är felfritt. Båda bygger på en beslutsstödsalgoritm som skapas av existerande data, ofta historisk sådan. Utifrån dessa data kan systemet användas för prediktioner eller beslutsunderlag om förhållanden idag eller framåt i tiden. Hur väl dessa utsagor faktiskt stämmer bygger dels på indata men också om utfallet ligger inom ett ”förväntat” utfallsrum. Detta kan innebära problem när det appliceras på människor som inte agerar rationellt eller på ett eller annat sätt agerar utanför vad samhället förväntar sig (Wired, 2021). Ytterligare en problematik kan komma av att man idag använder känsliga algoritmer och sensorer men gränsvärdena som modellen bygger på togs fram med andra sensorer eller för andra algoritmer. Det är med andra ord inte nödvändigtvis så att algoritmen inte är tillämpbar, men gränsvärdena som prediktionen bygger på har inte uppdaterats i takt med tekniken (Jordan, 2019). Det har också visat sig att mänskliga erfarna meteorologer är bättre på att förutspå väder under dåliga förhållanden än de prognosverktygen. Till del bygger det på att meteorologen ytterligare förfinar och balanserar det första utkastet till väderprognos som tagits fram av prognosverktyget. Detta innebär en arbetsfördelning där systemet tar fram enkla ”vackert väder”-prognoser medan meteorologerna kan spendera mer tid på att förfina och nyansera prognoser med mer dynamik (Wired, 2022).

4.1 Accelererande beräkningskapacitet, singularitet och superintelligens

Ett av skälen till det senaste decenniets framgångar för AI (mer specifikt ML) är som tidigare nämnts en ökad beräkningskapacitet. Robotforskaren Hans Moravec jämförde i slutet av 1990-talet ögats förmåga att upptäcka rörelser i en scen med motsvarande förmåga hos datorer. Utifrån denna jämförelse kunde han uppskatta ”beräkningskapaciteten” i näthinnan som sedan extrapolerades till en approximation av beräkningskapaciteten för den mänskliga hjärnan (Moravec, 1998). På så sätt skapades en möjlighet att bedöma hur stor maskinell beräkningskapacitet som skulle motsvara en mänsklig hjärna, och när det skulle kunna ske. Moravec konstaterade att beräkningskapaciteten hos datorer har ökat mer än exponentiellt vid en jämförelse av beräkningsmaskiner och datorer från 1900-talets början fram till slutet på 1990-talet.[17] Futuristen Ray Kurzweil har vidare extrapolerat Moravecs resultat för att resonera kring hur utvecklingen kommer att se ut i framtiden. Kurzweil konstaterade att utöver att beräkningskapaciteten ökar exponentiellt, så ökar även *takten* på utvecklingen. Han kallar detta för lagen om accelererande utkomster (*the Law of Accelerating Returns*) och menar att denna dubbelt exponentiella ökning av beräkningskapacitet (och utveckling av metoder för AI) förr eller senare kommer att ge upphov till AGI, *Artificiell generell intelligens* på mänsklig nivå och kort därefter på övermänsklig nivå (ASI). Kurzweil menar att när detta sker, kommer det att fundamentalt förändra vårt samhälle. Denna händelse kallar han för en *singularitet* och kommer enligt Kurzweils bedömning att inträffa när $1000 kan köpa beräkningskraft som motsvarar samtliga människors hjärnor på planeten. Med tanke på den accelererande exponentiella ökningen av beräkningskraft blir detta ca 2045 (Kurzweil, 2005).

I sin bok *Superintelligence* beskrev filosofen Niklas Boström (oftast refererad till som Nick Bostrom) bland annat resultaten från ett antal olika undersökningar där forskare och andra kunniga inom AI-området fått frågan: *när kommer maskinintelligens på mänsklig nivå att uppnås?* En sammanvägning av resultaten gav att sannolikheten för AGI på mänsklig nivå bedömdes till 10% år 2022, 50% år 2040 och 90% år 2075. Boström konstaterade också att resultaten varierade kraftigt, inte bara i bedömningen om när utan också om huruvida det var möjligt överhuvudtaget. Boström manar till försiktighet i tolkningen av resultatet, men konstaterar att det finns en rimlig chans att AGI på mänsklig nivå uppnås i mitten på detta århundrade, att det finns en icke försumbar möjlighet att det kan ske tidigare eller

[17] Moravec använde skalan beräkningar/sekund/1000 $ för att få ett rimligt mått att kunna jämföra olika system med varandra.

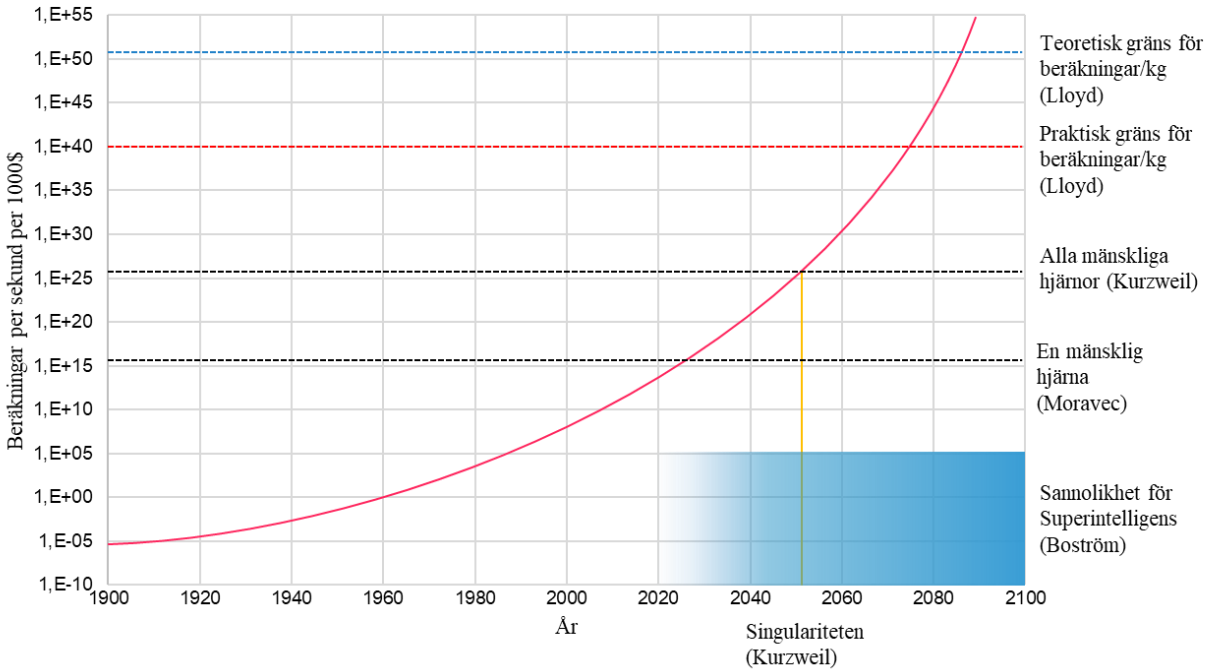
långt senare, att superintelligens kommer att utvecklas kort därpå och att detta innebär en existentiell risk för människan (Bostrom, 2014).

Det är flera som vill varna för de risker (och möjligheter) som en artificiell superintelligens kan komma att innebära för mänskligheten. Ett första steg att mitigera detta hot är att förbereda sig och utreda frågan djupare. Med denna bakgrund har ett flertal forskare och entreprenörer grundat *The Future of Life Institute*, ett institut vars syfte är att få upp frågan på den vetenskapliga, ekonomiska och politiska agendan världen över (Tegmark, 2017).

Hur länge är det rimligt att en exponentiellt ökande beräkningskapacitet fortsätter? MIT-fysikern Seth Lloyd har försökt fastslå ett antal gränser utifrån ett fysikaliskt perspektiv kring beräkningskapacitet, och tog fram följande formel:

$E = \frac{\pi\hbar}{2\Delta t}$ där ħ=Plancks konstant och Δt=tiden för en operation.

Då $E = mc^2$ innebär detta att 1 kg materia skulle kunna nyttjas för att kunna ge ca 5×10^{50} operationer/s. Detta förutsätter fullständig omvandling av materia till energi och bör därför ses som en teoretisk gräns. En mer rimlig praktisk gräns, som inte innebär fullständig konvertering materia till energi skulle vara ca 5×10^{40} op/s (Lloyd, 2000). Om Kurzweils prognos för utvecklingstakten av beräkningskapacitet stämmer, så kommer vi att nå denna gräns redan 2080.



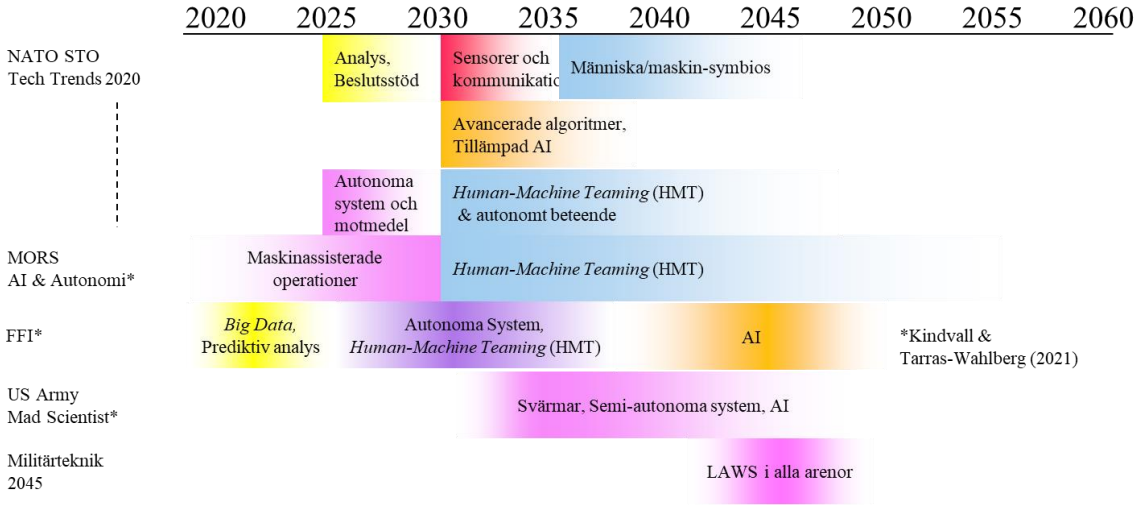
Figur 2: Moravec och Kurzweils projektion av ökande beräkningskapacitet, med tillägg av Seth Lloyds teoretiska och praktiska gränser för beräkning samt Boströms resonemang om sannolikhet för superintelligens.
Bild: Erik Granath, FOI.

4.2 Utveckling av militära applikationer med AI

I FOI-rapporten *Det framtida tekniklandskapet* konstaterar Kindvall och Tarras-Wahlberg (2021) att en rimlig ordning för hur AI kommer att påverka militära frågor kan vara att AI först introduceras för övervakning och anomalidetektion, därefter för informationsfusion, lägesbild och beslutsstöd i ledningsstödssystem och därefter för autonoma farkoster samt samverkan mellan maskin och människa. I ett kapitel går författarna även genom vad ett antal olika militära forskningsorganisationer

fokuserar på. Här dyker AI upp i några mer eller mindre tidsbestämda framtidsspaningar från US Army, *Forsvarets Forskningsinstitutt (FFI)*[18] samt en konferens om AI och autonomi (se Figur 3).

I rapporten *Science & Technology Trends Report 2020-20240* presenterar NATO *Science & Technology Organisation* (STO) möjliga utvecklingar för ett antal teknologier, däribland AI, *Big Data and Advanced Analytics* (BDAA) och autonomi, med några tidsbestämda förutsägelser (se Figur 3). I rapporten konstateras bland annat att AI-utvecklingen kommer att accelerera i princip all forskning och utveckling, i takt med att fler och fler tillämpningar av AI börjar användas. Därutöver kommer konvergensen mellan tillgången till enorma mängder data, AI och autonomi innebära den största kullkastande effekten för NATO de närmaste 10 åren.



Figur 3: Sammanställning av förutsägelser från olika källor. Visualiseringen är en tolkning av källornas beskrivningar och högst översiktlig. Bild: Erik Granath, FOI.

Tredje vågens AI[19], där inlärning kan kombineras med resonerande, kommer att innebära nästa stora genombrott för AI efter ML. System i denna våg förväntas förklara bakomliggande resonemang, vilket kommer att lösa problemet med dagens ”svarta lådor”. En annan utveckling är att systemen kommer att behöva mindre träningsdata och bli bättre på att hantera osäkerheter, tvetydigheter och motsägelser. Detta kommer att göra tredje vågens AI-system mer robusta samt ge dem möjlighet att i högre mån kunna generalisera och föra över lärdomar från en domän till en annan. AI kommer också att breda ut sig, från molnet ut på kanten (*Edge*-paradigmet) i och med att utvecklingen av neuromorfisk hårdvara möjliggör kraftfulla AI-tillämpningar med låg energiåtgång och låg vikt i autonoma system (NATO STO, 2020). En kommande hot/motmedels-kapplöpning är sannolik där AI kommer att användas för att upptäcka och motverka fiendens AI-system.

En intressant aspekt är ett amerikanskt försök att med snäv AI återskapa hur trollsländor jagar insekter. Istället för att försöka skapa enormt komplexa nätverk har forskarna tittat åt andra hållet – vilket är det minsta neurala nätverk som fortfarande kan lösa en väl definierad, men komplex, uppgift. Trollsländan lever av insekter och är en fantastisk jägare i luften. Genom att försöka återskapa ett neuralt nätverk som motsvarar en trollsländas komplexitet är målet att ta fram ett litet och snabbt system som kan användas i luftförsvarssituationer (Chance, 2021).

I FOI-rapporten *Militärteknik 2045* beskrivs bland annat hur IoT-trenden förväntas fortsätta för militära tillämpningar – i framtiden kommer system och soldater på grupp, pluton och kompaninivå att vara hopkopplade i nätverk där information från en stor mängd källor som varierar i tid och rum,

[18] Norges motsvarighet till FOI.
[19] Se 2.2 Tekniker för AI.

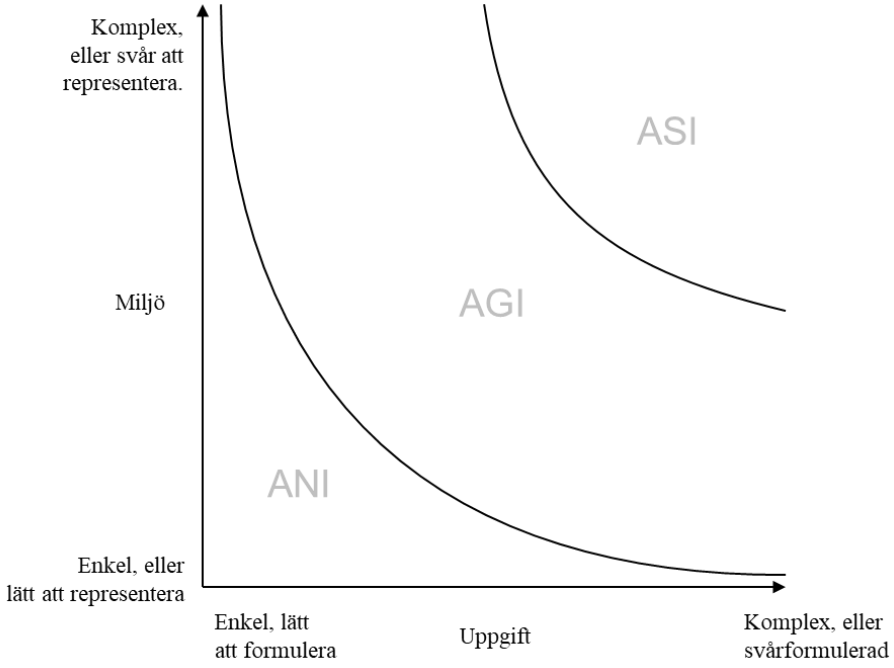
inklusive tillfälliga komponenter såsom olika UAS, kommer att behöva hanteras (Kindvall & Lindberg, 2020). För att kunna lösa kommunikationen krävs adaptiva radiosystem som på ett intelligent sätt kan analysera lokala förhållanden och anpassa sig därefter i självorganiserande och självläkande nät. Soldaten blir då likt andra plattformar en nod som kan ta emot eller generera information i ett eller flera informationsnätverk. AI kommer även att vara en nödvändig komponent i *Human Performance Enhancement* (HPE) – initiativ framöver. I och med att mängden information som förmedlas i realtid kommer att öka så kommer också den mentala belastningen att öka på soldater och officerare. AI-stödda beslutsstödsystem på soldatnivå kan därmed komma att krävas för att hjälpa soldaten att prioritera information. På lång sikt kan det bli möjligt att integrera människa och maskin, genom kognitiv och sensorisk förstärkning i form av hjärnimplantat. En sådan integration förutsätter någon form av AI-system för att tolka hjärnans signaler och för att kunna presentera information på ett sätt som är lämpligt för människan. I förlängningen kan detta ses som synergiskt med trenden Human-Machine Teaming (HMT), där det är rimligt att vänta sig ett ökat inslag av maskiner som agerar tillsammans med människor på slagfältet. Utöver HMT så kommer flera stater sannolikt att ha obemannade beväpnade system med autonoma funktioner i alla arenor till 2045, se Figur 3 (Kindvall & Lindberg, 2020).

Fler och fler militära applikationer använder sig dock av elektromagnetiska spektrumet vilket innebär att telekrig i olika varianter bedöms vara en väsentlig förmåga. I många avseende kan AI utgöra en viktig komponent i denna förmåga då många av uppgifterna lämpar sig väl för nuvarande tekniska mognadsnivå (Sharma, 2020). Amerikanska flygvapnet har nyligen påbörjat ett program kallat *Project Kaiju* i syfte att kunna utnyttja AI för att mitigera hotet från fientliga luftförsvarssystem. Konceptet kallas ofta för *cognitive electronic warfare* (CEW) (Keller, 2021). Även Ryssland lägger stort fokus vid AI-stöd för telekrigsfunktioner och man har minst ett operativt system som utnyttjar AI i någon form, *Bylina*. Systemet är en autonom C2-plattform för telekrig på brigadnivå som ska kunna analysera slagfältet, identifiera mål och välja hur de ska neutraliseras samt kunna ge uppgifter till telekrigsenheter i fält (Morgan et al., 2021; Bendett, 2021).

Svärmande system, framför allt i komplexa miljöer som stadsbebyggelse, kommer enligt NATO att innebära en *force multiplier* (NATO STO, 2020). På sikt kommer svärmarna troligen bestå av många olika komponenter i syfte att öka flexibiliteten och att för kunna genomföra olika typer av autonoma operationer. I likhet med flera andra områden kommer det sannolikt att ske en snabb utveckling av medel/motmedel för svärmar av små, billiga och korträckviddiga drönare. Svärmteknologi med flera enkla och relativt sett billigare system har potential att nyttjas i flera olika militära applikationer (Hambling, 2021b).

5 Diskussion

Ett sätt att resonera kring i vilken ordning AI kommer att påverka militära operationer i framtiden är utifrån två dimensioner, uppgift och miljö. I Figur 4 presenteras ett försök till att resonera kring dessa svårigheter. I diagrammet presenteras två dimensioner: Miljö och Uppgift, som löper från enkel till komplex. Vad som avses med en enkel eller komplex miljö i detta sammanhang är egentligen hur lätt eller svårt det är att representera miljön på ett sätt som gör det möjligt för ett AI-system att ”förstå” miljön och agera i den. På liknande vis kan man även tänka sig representera uppgiftsdimensionen – är uppgiften enkel eller svår att formulera på ett sätt som gör att systemet kan utföra uppgiften fullt ut?



Figur 4: Konceptuell illustration av vilka uppgifter och i vilka miljöer olika nivåer av AI kan användas.
Bild: Erik Granath, FOI.

De fyra hörnen av diagrammet beskriver extrempunkterna i utfallsrummet som kan representeras enligt nedan:

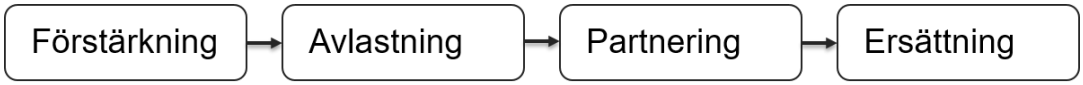
- **Enkel miljö/enkel uppgift.** I denna punkt krävs väldigt lite intelligens - här finns en mängd system som reagerar regelbaserat (om si, gör så) på en förändring i miljön. Ett extremt exempel är en termostat.
- **Enkel miljö/komplex uppgift.** När det gäller till exempel system (såsom AlphaGo) som tränas för att slå människor i olika brädspel, så är miljön mycket enkel, ett spelbräde med enkla regler där alla drag är synliga. Systemet har således fullständig kunskap om miljön. Men uppgiften är komplex och systemet måste kunna utveckla strategier och anpassa sig till olika motståndares spelstil i ett enormt utfallsrum för att kunna vinna.
- **Komplex miljö/enkel uppgift.** När det gäller självkörande fordon, så är miljön väldigt komplex. Vägar som ska navigeras, trafikregler som ska följas och andra trafikanter som måste undvikas. Den övergripande uppgiften är relativt enkel att formulera - navigera från punkt A till punkt B utan att orsaka trafikolyckor - men kräver att en stor mängd olika datakällor måste analyseras i syfte att lösa ett stort antal deluppgifter.
- **Komplex miljö/komplex uppgift** I denna punkt finns teoretiskt omnipotenta system som kan utföra alla uppgifter i alla miljöer.

Exemplen ovan är hämtade från den civila marknaden och civila tillämpningar kommer antagligen att vara drivande för utvecklingen av AI även i framtiden. Men många av de civila applikationer som dyker upp kan vara väldigt svåra eller olämpliga att direkt föra över till ett militärt sammanhang, då civila tillämpningar inte behöver ta hänsyn till den ökade komplexiteten i miljö och uppgift som en militär applikation innebär. Krigets friktioner och en aktiv motståndare som medvetet agerar mer eller mindre dold i syfte att förvirra, överlista eller vilseleda. Det civila vägnätet är en mer strukturerad miljö än exempelvis ett svenskt skogslandskap. Däremot måste ett autonomt fordon som agerar i vägnätet hantera samspelet med andra trafikanter vars beteende och samspel är svårt att förutsäga. Ett autonomt system måste kunna tolka den omgivande terrängen för att kunna navigera i och utanför vägnätet, utan tillgång till GPS. Det måste kunna kommunicera med andra enheter i en störd miljö för att bidra till en gemensam lägesbild. Det måste också kunna agera mot motståndaren med en portfölj av olika verkanstyper, och så vidare.

Den grad av autonomitet som civila konsumenter har börjat förvänta sig av exempelvis fordon är med dagens AI-teknik svår att uppnå med robusthet i den militära domänen. Det som DARPA kallar för *contextual adaptation*, nästa våg av AI-utveckling, kommer att spela stor roll i den militära domänen, då flera av begränsningarna med dagens maskininlärning försvinner. System som utvecklas i denna våg kommer att kunna tränas under tiden det används, anpassas efter nya sammanhang och kommer att vara transparenta. System som utvecklas i detta paradigm kommer i mycket högre omfattning att kunna vara autonoma även i militära applikationer, då de potentiellt kommer kunna vara mycket mer robusta.

Diagrammet innehåller också tre nivåindelningar: ANI, AGI och ASI. Dessa är menade att illustrera de ökande kraven på ett AI-system efterhand som miljö och uppgift blir alltmer komplex. Den första nivån indikerar att kraven på systemet innebär att det måste kunna hantera antingen enkla uppgifter i en komplex miljö, alternativt komplexa uppgifter i en enkel miljö vilket motsvaras av det vi idag kallar snäv AI. Den andra nivån indikerar de uppgifter som kräver en flexibilitet som i dagsläget endast människor kan uppnå, med delvis komplexa uppgifter i delvis komplexa miljöer, något vi förväntar oss en AGI kan lösa. Den tredje nivån indikerar de uppgifter som det idag krävs organisationer av människor för att lösa samt alla uppgifter som kräver övermänsklig förmåga, en nivå som skulle motsvaras av en ASI. För närvarande är det endast så kallad snäv AI som har nått tillräcklig mognadsnivå för att kunna användas operativt och då framförallt inom olika typer av stöd i en stab eller beslutskedja samt tolkning och bearbetning av sensordata.

Baserat på resonemanget kring vilka komplexa uppgifter som ska genomföras i olika miljöer, så är en rimlig ordning för hur AI kommer att förändra militär verksamhet enligt Figur 5.



Figur 5 Hypotes för i vilken ordning AI kommer förändra militär verksamhet.

Maskininlärning, det paradigm som ligger till grund för dagens explosion av AI-tillämpningar, har ett antal inneboende begränsningar som gör att det blir viktigt att identifiera tillämpningar där rätt förutsättningar för maskininlärning finns. Detta gäller framför allt tillgång till träningsdata, riskerna med felklassificering samt bristen på transparens. Av dessa skäl är det rimligt att tänka sig att ML-system framför allt kommer att användas i snäva, väl definierade uppgifter i syfte att förstärka eller avlasta människor i olika roller. I de visioner om framtida militära operationer som olika forskningsorganisationer och försvarsmakter målar upp, så kommer behovet att minska den kognitiva belastningen hos personalen att vara stort. När, var och i vilken takt denna progression inträffar och för vilka roller beror på komplexiteten i uppgiften och miljön som systemet verkar i. Först kommer denna progression att inledas för enkla uppgifter i enkla miljöer, därefter för komplexa uppgifter i enkla miljöer, sedan för enkla uppgifter i komplexa miljöer och slutligen för komplexa uppgifter i komplexa miljöer.

Det initiala värdet av AI-system blir förstärkning och/eller effektivisering redan existerande förmågor, genom intelligenta sensorer, analys av stora mängder information och beslutsstöd. På grund av att tekniken fortfarande är relativt omogen, kommer människor fortsatt att behöva värdera de rekommendationer som systemen producerar. Givet att militär AI-utveckling följer på den civila utvecklingen, går det att i Lees olika ”vågor” av AI att se likheter i den militära domänen. ”Internet AI” är direkt applicerbar för omvärldsanalys och underrättelsetjänst framförallt inom ramen för förstärkning och effektivisering.

I takt med att robustheten hos militära AI-system ökar kommer också tilliten till systemen att öka, varpå systemen kommer att kunna avlasta personal genom att autonomt värdera, kondensera, prioritera och presentera information som måste värderas av en människa. På så vis kan mänsklig kognition fokuseras på de aspekter av operationer som kräver mänsklig flexibilitet.

Nästa steg, partnering, innebär att AI-system kan tilldelas enklare uppgifter av sin operatör och förväntas lösa dessa på egen hand. I detta steg kan ett AI-system ingå som en komponent i en stab eller grupp, om än i en begränsad roll, där en nära kontakt med en mänsklig operatör krävs.

I det sista steget kan AI-system ta en roll som helt ersätter en människa. I detta steg kan en helt autonom pluton av stridsfordon få en stridsuppgift att lösa, med samma förväntningar som på en pluton med mänskliga besättningar. En effekt som kan komma att växa i betydelse efterhand som systemen blir mer självgående är det blir möjligt att öka avståndet mellan operatör och system vilket kan få omvälvande betydelse då risken för (egna) mänskliga förluster minskar.

När AI-systemen når ”ersätta”-stadiet och autonoma och intelligenta agenter av olika slag nyttjas på ett fysiskt eller digitalt slagfält kommer frågan om end-state bli allt viktigare. Kommer systemen ”släppas lösa” på slagfältet att agera fritt tills ett väldefinierat mål uppnåtts och hur väl kan man definiera eftersträvade mål? Då Ryssland och Kina sannolikt inte kommer begränsa nyttjandet av AI utifrån etiska eller moraliska perspektiv kan man ställa sig frågan om det blir en kapprustningslogik i vilken typ av mål som definieras? Inledningsvis kommer USA sannolikt inte tillåta fritt nyttjande av AI. Frågan som följer är hur USA kommer att hantera risken som det innebär att inte längre ha ett teknologiskt övertag.

Utvecklingen av mer intelligenta sensorer, sensorfusion och AI-förstärkt lägesbildhantering innebär att det kommer att bli lättare att övervaka stora områden även med en liten stab. Detta gör att det kommer att bli betydligt svårare att gömma sig för fienden. Denna trend har nära kopplingar till pågående utveckling inom rymdområdet med flera nya sensorer och plattformar. Det finns en hypotes om att dessa trender tillsammans pekar mot en GEOINT-singularitet, definierad som den tidpunkt när det finns en kontinuerlig övervakning av jorden tillgänglig för allmänheten (Koller, 2019). Ett sätt att mitigera denna utveckling är motsvarande konceptutveckling i syfte att vilseleda AI-förstärkta sensorer och system. Ur ett mer generellt perspektiv har vilseledning också identifierats som en möjlig kullkastande funktion för AI-system.

Avslutningsvis är det viktigt att ha i åtanke att AI i sig inte är ett vapen. AI bör snarare ses som ett verktyg och en möjliggörare. Hur väl man kan dra nytta av verktyget är nära knutet till vilka applikationer man tillämpar verktyget på och hur denna tillämpning sker. Det krävs också att man kan balansera en (teknisk) mognad hos både verktyg och organisation. Är inte verktyget moget för uppgiften kommer sannolikt inte resultatet motsvara förväntningen, motsvarande så krävs en tillräckligt mogen organisation för att kunna dra nytta av ett avancerat och komplext verktyg. Exempelvis är en tidskritiskt korrekt analys endast värdefull om det finns någon som kan ta emot och dra nytta av den i sitt beslutsfattande. Ju snabbare teknikutvecklingen blir desto viktigare är det att den matchas av en flexibel organisation som har processer och självförtroende att utveckla doktrin och taktik i syfte att dra nytta av de nya verktyg som står organisationen till hands.

Referenser

Böcker

Bostrom, N. (2014). *Superintelligence: The coming machine intelligence revolution.* Oxford: Oxford University Press.

Kurzweil, R. (2005). *The Singularity is Near.* London: Penguin.

Lee, K-F. (2018). *AI Superpowers: China, Silicon Valley and the New World Order.* New York: Houghton Mifflin Harcourt.

Tegmark, M. (2017). *Liv 3.0.* Stockholm: Volante.

Rapporter

Andersson, C., Gustavi, T., Karasalo, M. (2019). *Artificiell intelligens – möjligheter och utmaningar för Sveriges nationella säkerhet.* Stockholm: FOI.

Bendett, S., et al. (2021). *Advanced military technology in Russia – Capabilities and implications,* London: Chatham House.

Clark, B., Patt, D., Schram, H. (2020). *Mosaic Warfare: exploiting artificial intelligence and autonomous systems to implement decision-centric operations.* Washington: Center for Strategic and Budgetary Assessments.

DCDC (2018). *Global Strategic Trends: The Future Starts Today.* London: Ministry of Defence.

Durmaz, T., Velasco, G., León, G. (2020). *NEUROMORPHIC COMPUTING Concepts, actors, applications, market and future trends.* Madrid: Universidad Politécnica de Madrid.

Edmonds, J., Bendett, S., Fink, A., Chesnut, M., Gorenburg, D., Kofman, M., Stricklin, K., Waller, J. (2021). *Artificial Intelligence and Autonomy in Russia.* Arlington: CNA

Gustafsson, D., Ahlinder, L., Bergström, D., Björnham O., Brännlund, C., Brännström, P., Brorsson, A., Granström, N., Ingemarsdotter, J., Nordberg, M., Persson, L., Svenmarck, P., Svensson, C. (2020). *Maskininlärning inom försvarsorienterade tillämpningar – Nya ansatser och experimentella resultat.* Stockholm: FOI.

Gustavi, T., Karlholm, J., Oskarsson, D., Beran, T., Björnham, O., Gudmundson, E., Heden, H., Karlzén, H., Lindblom, J., Nordlöf, J., Rodhe, I., Sommestad, T., Svenmarck, P., Svensson, M. (2019). *Försvarsnära tillämpningar av artificiell intelligens.* Stockholm: FOI.

HM Government (2021). *Global Britain in a competitive age: The Integrated Review of Security, Defence, Development and Foreign Policy.* London: Her Majesty's Stationery Office.

Johansson, R. (red.), Andersson, M., Hörling, P., Kamrani, F. (2019). *Samverkande IoT-system och databearbetning.* Stockholm: FOI.

Joint Special Operations University (JSOU) (2021). *Special Operations Research Topics 2022.* Tampa: JSOU Press.

Jordan, M. I. (2019). *Artificial Intelligence—The Revolution Hasn't Happened Yet*. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.f06c6e61

Kania. E. (2020). *"AI Weapons" In China's Military Innovation*. Washington: Brookings.

Kindvall, G., Lindberg, A. (red.) (2020). *Militärteknik 2045: Ett underlag till Försvarsmaktens perspektivstudie.* Stockholm: FOI.

Kindvall, G., Tarras-Wahlberg, B. (2021). *Det framtida tekniklandskapet: En översikt.* Stockholm: FOI.

Koller, J. (2019). *The Future of Ubiquitous, Real-Time Intelligence – A GEOINT Singularity*, The Aerospace Corporation.

Layton, P. (2021). *Fighting Artificial Intelligence Battles: Operational Concepts for Future Ai-Enabled Wars.* Canberra: Australian Defence College.

Luotsinen, L., Oskarsson, D., Svenmarck, P., Wickenberg Bolin, U. (2019). *Explainable Artificial Intelligence: Exploring XAI Techniques in Military Deep Learning Applications.* Stockholm: FOI.

Morgan, F., Boudreaux, B., Lohn, A., Ashby, M., Curriden, C., Klima, K., Grossman, D. (2020). *Military Applications of Artificial Intelligence.* Santa Monica: RAND.

NATO STO (2020). *Science & Technology Trends 2020-2040*. Bryssel: NATO Science & Technology Organisation.

Office Of The Secretary Of Defense. (2020). *Military and Security Developments Involving the People's Republic of China 2020.* Washington: Office Of The Secretary Of Defense

Schuchmann, S. (2019). *Analyzing the Prospect of an Approaching AI Winter*. 10.13140/RG.2.2.10932.91524.

UK MoD (2021c). *Integrated Operating Concept.* London: UK Ministry of Defence.

UK MoD (2021d). *Digital Strategy for Defence: Delivering the Digital Backbone and unleashing the power of Defence's data.* London: UK Ministry of Defence.

Artiklar

Hedberg, S. (2002). *DART: Revolutionizing Logistics Planning*. IEEE Intelligent Systems, Vol. 17, 2002, pp. 81–83. DOI: 10.1109/MIS.2002.1005635.

Kott, A., Ananthram, S., West, B. (2016) *The Internet of Battle Things.* Computer, 49.12, pp. 70-75.

Lloyd, S. (2000). *Ultimate physical limits to computation.* Nature, Vol. 406, pp. 1047–1054. DOI: 10.1038/35023282.

Moravec, H. (1998), *When will computer hardware match the human brain?*, Journal of Evolution and Technology, Vol. 1

Sharma, P., Sarma, K. K., Mastorakis, N. E. (2020). *Artificial Intelligence Aided Electronic Warfare Systems- Recent Trends and Evolving Applications*, IEEE Access, Vol. 8, doi: 10.1109/ACCESS.2020.3044453.

Websidor

Ackerman, E., (2021), IEEE Spectrum, *How the U.S Army is turning robots into team players*. Hämtad 220105 från; https://spectrum.ieee.org/ai-army-robots

AFRL. (u. å.). *Skyborg*. Hämtad 210917 från: https://afresearchlab.com/technology/vanguards/successstories/skyborg

Allen, G. (2020). *Understanding AI Technology*. Hämtad 211001 från: https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf

Allison, G. (2014). UK Defence Journal, *Taranis stealth drone test flights successful*. Hämtad 210917 från: https://ukdefencejournal.org.uk/taranis-stealth-drone-test-flights-successful/

Amazon. (u. å.). *What is Artificial Intelligence?* Hämtad 211001 från: https://aws.amazon.com/machine-learning/what-is-ai/

Army Technology. (2021). *USSOCOM contracts Palantir for enterprise data management software.* Hämtad 210913 från: https://www.army-technology.com/news/ussocom-contracts-palantir-for-enterprise-data-management-software/

Burns, R. (2021). AP, *Pentagon cancels disputed JEDI cloud contract with Microsoft.* Hämtad 211014 från: https://finance.yahoo.com/news/pentagon-cancels-disputed-jedi-cloud-163913312.html

Business Insider. (2019). Business Insider, *Palantir grabbed Project Maven defense contract after Google left the program: sources.* Hämtad 210913 från: https://www.businessinsider.com/palantir-took-over-from-google-on-project-maven-2019-12?r=US&IR=T

Chance, F. (2021), IEEE Spectrum, *Fast, efficient neural networks copy dragonfly brains*. Hämtad 220105 från: https://spectrum.ieee.org/fast-efficient-neural-networks-copy-dragonfly-brains

Chinese State Council. (2017). Foundation for Law and International Affairs, *Notice of the State Council Issuing the New Generation of Artificial Intelligence Development Plan.* Hämtad 210913 från: https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan/

Congressional Research Service (CRS). (2021). *Joint All-Domain Command and Control (JADC2)*. Hämtad 210916 från: https://sgp.fas.org/crs/natsec/IF11493.pdf

Cramer, M. (2021). New York Times, *A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says.* Hämtad 210917 från: https://www.nytimes.com/2021/06/03/world/africa/libya-drone.html

Dahm, M. (2020). War on the Rocks, *Chinese Debates On The Military Utility Of Artificial Intelligence.* Hämtad 210916 från: https://warontherocks.com/2020/06/chinese-debates-on-the-military-utility-of-artificial-intelligence/

DARPA (2017). *A DARPA Perspective on Artificial Intelligence*. Hämtad 210929 från: https://www.darpa.mil/about-us/darpa-perspective-on-ai

DARPA. (u. å.). *AI Next Campaign.* Hämtad 210913 från: https://www.darpa.mil/work-with-us/ai-next-campaign

Deepmind. (2021). *AlphaStar: Grandmaster level in StarCraft II using multi-agent reinforcement learning.* Hämtad 210909 från: https://deepmind.com/blog/article/AlphaStar-Grandmaster-level-in-StarCraft-II-using-multi-agent-reinforcement-learning

Defense Innovation Board. (u. å.). *AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense*. Hämtad 210927 från: https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF

Defense Innovation Unit (2021). *Who We Are / Our Mission*. Hämtad 211014 från: https://www.diu.mil/about

Desjardins, J. (2019). World Economic Forum, *How much data is generated each day?* Hämtad 211016 från: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

Digiteum. (2020). *Difference between Cloud, Fog and Edge Computing in IoT*. Hämtad 211018 från: https://www.digiteum.com/cloud-fog-edge-computing-iot/

DSTL. (2021a). *AI and data science: defence science and technology capability*. Hämtad 210910 från: https://www.gov.uk/guidance/ai-and-data-science-defence-science-and-technology-capability

DSTL. (2021b). *Dstl research and funding priorities*. Hämtad 210910 från: https://www.gov.uk/guidance/dstl-research-and-funding-priorities#artificial-intelligence-autonomy-and-robotics

Future of Life Institute (u. å.). *Autonomous Weapons: An Open Letter From Ai & Robotics Researchers.* Hämtad 211013 från: https://futureoflife.org/open-letter-autonomous-weapons/

Future of Life Institute. (2017). *Slaughterbots.* Hämtad 211013 från: https://www.youtube.com/watch?v=9CO6M2HsoIA

Gershgorn, D. (2017). Quartz, *The data that transformed AI research—and possibly the world*. Hämtad 210909 från: https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world/

Geist, E. and Blumenthal, M. (2019). War on the rocks, M*ilitary deception: AI's killer app?* Hämtad 220105 från: https://warontherocks.com/2019/10/military-deception-ais-killer-app/

Hambling, D. (2021a). Forbes, *DARPA Gremlin Swarm Will Carry Weapons Or Sub-Drones And Re-Arm Mid-Air*. Hämtad 210924 från: https://www.forbes.com/sites/davidhambling/2021/06/17/darpa-gremlin-swarm-drones-to-carry-weapon-and-re-arm-mid-air

Hambling, D. (2021b). Forbes, *What Are Drone Swarms And Why Does Every Military Suddenly Want One?* Hämtad 220105 från: https://www.forbes.com/sites/davidhambling/2021/03/01/what-are-drone-swarms-and-why-does-everyone-suddenly-want-one

Harper, J. (2020). National Defence Magazine, *VSOFIC NEWS: SOCOM All-in on Artificial Intelligence*. Hämtad 210909 från: https://www.nationaldefensemagazine.org/articles/2020/5/12/socom-all-in-on-artificial-intelligence

Heaven, W. (2020). Technology Review, *OpenAI's new language generator GPT-3 is shockingly good—and completely mindless*. Hämtad 211023 från: https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/

Huang, E., (2017), Quartz, *China got a wake-up call on AI when a Western machine beat the best Asian players at Go.* Hämtad 220105 från: https://qz.com/1035313/china-got-a-wake-up-call-on-ai-when-a-western-machine-beat-the-best-asian-players-at-go/

Improbable. (2020). *Improbable's Defence business awarded contract for a second year.* Hämtad 211015 från: https://www.improbable.io/blog/improbable-uk-strategic-command-sse

Insinna, V. (2021). Defence News, *The Air Force's first Skyborg autonomous drone prototype made its first flight.* Hämtad 210917 från: https://www.defensenews.com/air/2021/05/05/the-air-forces-first-skyborg-autonomous-drone-prototype-made-its-first-flight/

JAIC Public Affairs. (2020). *JAIC partners with USSOCOM to deliver AI-enabled predictive maintenance capabilities.* Hämtad 210913 från: https://www.ai.mil/news_12_17_20-jaic_partners_with_ussocom_to_deliver_ai-enabled_predictive_maintenance_capabilities.html

JAIC. (2019). *About The Jaic: The Jaic Story*. Hämtad 211014 från: https://www.ai.mil/about.html

Jay, A. (2021). Finances Online, *Number of Internet of Things (IoT) Connected Devices Worldwide 2021/2022: Breakdowns, Growth & Predictions.* Hämtad 211016 från: https://financesonline.com/number-of-internet-of-things-connected-devices/

Joe, R. (2019). The Diplomat, *China's Growing High-End Military Drone Force.* Hämtad 210916 från: https://thediplomat.com/2019/11/chinas-growing-high-end-military-drone-force/

Kallenborn, Z. (2021). Bulletin of the Atomic Scientists, *Meet the future weapon of mass destruction, the drone swarm*. Hämtad 210924 från: https://thebulletin.org/2021/04/meet-the-future-weapon-of-mass-destruction-the-drone-swarm/

Keller, J., (2021), Military and Aerospace Electronics, *Air Force eyes artificial intelligence (AI) and machine learning for cognitive electronic warfare (EW)*. Hämtad 220105 från: https://www.militaryaerospace.com/computers/article/14210232/artificial-intelligence-ai-machine-learning-electronic-warfare-ew

NBCNews, (2021), Kamikaze drones: A new weapon brings power and peril to the U.S. military, Hämtad 220105 från: https://www.nbcnews.com/news/military/kamikaze-drones-new-weapon-brings-power-peril-u-s-military-n1285415

Newdick, T. Trevithick, J. (2021). The Drive, *Boeing Is Adapting Its Australian Combat Drone For The U.S. Air Force's Skyborg Program.* Hämtad 210917 från: https://www.thedrive.com/the-war-zone/39560/boeing-is-adapting-its-australian-combat-drone-for-the-u-s-air-forces-skyborg-program

New York Times. (2017). *Beijing Wants A.I. to Be Made in China by 2030*. Hämtad 220105 från: https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html

Palantir. (2021). *Gotham: The Operating System for Global Decision Making.* Hämtad 211015 från: https://www.palantir.com/platforms/gotham/

Reuters. (2021). *China has won AI battle with U.S., Pentagon's ex-software chief says.* Hämtad 211014 från: https://www.reuters.com/article/ctech-us-usa-china-ai-idCAKBN2H10C8-OCATC?

Royal Air Force. (2021). *Team Tempest.* Hämtad 210917 från: https://www.raf.mod.uk/what-we-do/team-tempest/tempest/

Strickland, E., (2021), IEEE Spectrum, *2021's Top Stories About AI*, Hämtad 220105 från: https://spectrum.ieee.org/artificial-intelligence-2021

Tass. (2020). *В России разработали систему прогнозирования маршрутов движения дронов противника.* Hämtad 210915 från: https://tass.ru/armiya-i-opk/9232665

Tesler, L. (u. å.). *Adages & Coinages*. Hämtad 210927 från: http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html

Turek, M. (u. å.). DARPA, *Explainable Artificial Intelligence* (XAI) Hämtad 210913 från: https://www.darpa.mil/program/explainable-artificial-intelligence

UK MoD. (2021). *Artificial Intelligence used on Army operation for the first time.* Hämtad 210910 från: https://www.gov.uk/government/news/artificial-intelligence-used-on-army-operation-for-the-first-time

UK MoD. (2021b). *£30-million injection for UK's first uncrewed fighter aircraft.* Hämtad 210917 från: https://www.gov.uk/government/news/30m-injection-for-uks-first-uncrewed-fighter-aircraft

United Nations (2019). *Background on LAWS in the CCW*. Hämtad 211013 från: https://www.un.org/disarmament/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/

US DoD. (2018). *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity.* Hämtad 211014 från: https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Weintz, S. (2018). National Interest, *Why Russia's Status-6 Torpedo Is Really a 100-Megaton Cruise Missile*. Hämtad 211014 från: https://nationalinterest.org/blog/buzz/why-russias-status-6-torpedo-really-100-megaton-cruise-missile-25272

White, A. (2021). Janes, *Dstl advances Project Theseus*. Hämtad 211015 från: https://www.janes.com/defence-news/news-detail/dstl-advances-project-theseus

Wikipedia. (2021). *Deep Learning.* Hämtad 210909 från: https://en.wikipedia.org/wiki/Deep_learning#Applications

Wired, (2021). *If AI Is Predicting Your Future, Are You Still Free?* Hämtad 220105 från; https://www.wired.com/story/algorithmic-prophecies-undermine-free-will/

Wired (2022). *The Danger of Leaving Weather Prediction to AI*, Hämtad 220105 från: https://www.wired.com/story/weather-forecasting-artifical-intelligence/