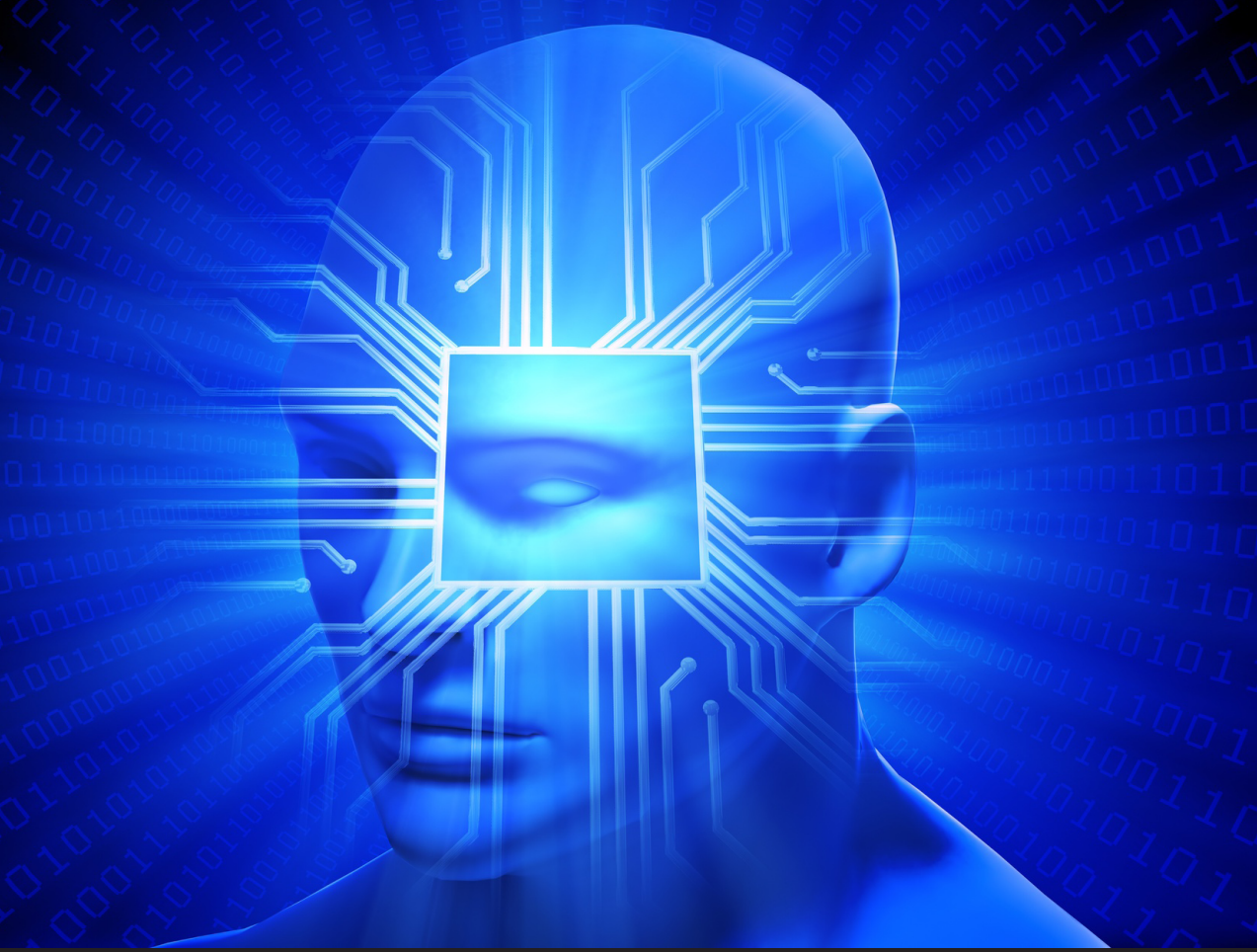

Artificiell Intelligens

Teknisk prognos

Linus Gisslén

FOI-R--3919--SE

September 2014

FOI

Linus Gisslén

Artificiell Intelligens

Teknisk prognos

Bild/Cover: Pixtal/TT.se

Titel	Artificiell intelligens – teknisk prognos
Title	Artificial intelligence – technology foresight
Rapportnr/Report no	FOI-R--3919--SE
Månad/Month	September
Utgivningsår/Year	2014
Antal sidor/Pages	44 p
ISSN	1650-1942
Kund/Customer	FMV
Forskningsområde	1. Beslutsstödssystem och informationsfusion
FoT-område	Avskanning av forskningsfronten
Projektnr/Project no	E34220
Godkänd av/Approved by	Lars Höstbeck
Ansvarig avdelning	Informations- och aerosystem

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk.
All form av kopiering, översättning eller bearbetning utan medgivande är förbjuden

This work is protected under the Act on Copyright in Literary and Artistic Works (SFS 1960:729).
Any form of reproduction, translation or modification without permission is prohibited.

Sammanfattning

Syftet med denna tekniska prognos är att beskriva aktuella och viktiga områden inom artificiell intelligens (AI) och göra en prognos om deras utveckling i framtiden.

Nyckelord: teknisk prognos, artificiell intelligens, självlärande maskiner, mönsterigenkänning, bildigenkänning, intelligenta system, autonoma farkoster, big data, cybernetik, robotik, svärmintelligens

Summary

The purpose with this technology foresight is to describe currently important areas within artificial intelligence (AI) and to do a foresight about their development in the future.

Keywords: technology foresight, artificial intelligence, self-learning machines, pattern recognition, image recognition, intelligent systems, autonomous vehicles, big data, cybernetics, robotics, swarm intelligence

Innehållsförteckning

Innehåll

1 Introduktion 7

2 Artificiell intelligens (AI) 9

2.1.1 Definition av AI 9
2.1.2 Var tillämpas AI 10
2.1.3 Olika typer av AI-metoder 10
2.1.4 Beräkningskraft 11
2.1.5 Forskningstrender och prognos 12
2.1.6 Etiska aspekter 13

3 Tillämpningar av AI och framtida användningsområden 14

3.1 Autonoma farkoster 14

3.1.1 Autonoma robotar 14
3.1.2 Autonoma bilar 15
3.1.3 Autonoma flygmaskiner 16
3.1.4 Robotsvärmar och svärmintelligens 16
3.1.5 Sammanfattning autonoma farkoster 17

3.2 Visuell intelligens 17

3.2.1 Bildigenkänning 17
3.2.2 Händelseigenkänning 18
3.2.3 Aktiva visuella system 19
3.2.4 Situationsbedömning 19
3.2.5 Sammanfattning visuell intelligens 20

3.3 Intelligenta hjälpmedel 21

3.3.1 Exoskelett 21
3.3.2 Proteser 21
3.3.3 Sjukvård 22

3.4 Cyberhot 22

3.4.1 Intelligenta virus-attacker 23
3.4.2 Intelligent försvar 24
3.4.3 Summering cyberhot 24

3.5 Övervakning 25

3.5.1 ”Tracking” via övervakningskameror ... 25
3.5.2 Anomalidetektion ... 25
3.5.3 Sjukvård ... 27
3.5.4 Övervakning av samhällsfunktioner ... 27
3.5.5 Sammanfattning övervakning ... 28

3.6 Prognoser ... 28

3.6.1 Väderprognoser ... 29
3.6.2 Samhällsprognoser ... 29
3.6.3 Pandemi- och epidemiprognoser ... 29
3.6.4 Sammanfattning prognoser ... 30

3.7 Expertsystem ... 30

3.7.1 Kliniska expertsystem ... 31
3.7.2 Personlig assistent/rådgivare ... 31
3.7.3 Sammanfattning expertsystem ... 32

3.8 Big data och informationshantering ... 32

3.9 Språk och kommunikation ... 33

3.9.1 Sökfunktioner ... 33
3.9.2 Taligenkänning ... 34
3.9.3 Språköversättning ... 34
3.9.4 Text- och talförståelse ... 34
3.9.5 Sammanfattning språk och kommunikation ... 34

4 Bevakningsbehov 36

4.1.1 Tekniktrender att följa ... 36
4.1.2 Intressanta trender inom länder och stater ... 37
4.1.3 Intressanta trender inom industrin ... 38

5 Säkerhetsaspekter 39

6 Referenser 41

1 Introduktion

Människan har ständigt uppfunnit verktyg för att bättre utnyttja resurser och underlätta livet under de omständigheter som råder. Var det kallt tillverkades hus och kläder, var det ont om mat tillverkades verktyg för att bruka jorden. Nu när dessa faktorer påverkar oss i mindre utsträckning och samhället är fundamentalt annorlunda ställs vi inför andra problem: i dagens informationsbaserade samhälle exponeras vi potentiellt för mer information under en dag än vad vi gjorde under en hel livstid för 10000 år sedan. Frågan är nu vilka verktyg vi behöver nu för att hantera denna nya verklighet?

Att filtrera, sortera, förstå, manipulera, extrahera i det dataflöde vi omges av har blivit informationsålderns stora utmaningar. Från att fysiska tillgångar historiskt sett har lett till en nations välmående och rikedom, har ett skifte skett till att det är de mer abstrakta tillgångarna som påverkar framgång mest. Det kommer mer och mer bli förmågan att använda och förstå information som blir avgörande för utveckling. För att göra detta måste vi bygga system som kan hjälpa oss människor med de delar vi inte förmår, kan, är för långsamma för, eller helt enkelt tycker är för tråkiga att hålla på med.

Som en lösning på detta har datorn blivit ett oöverträffat verktyg. Den kan analysera, spara, visualisera och producera data på ett sätt som inte är möjlig för oss människor. Fortfarande är användningsområdet för datorer begränsat eftersom de saknar den intelligens som krävs för att verkligen kunna hantera informationen de matas med på ett bra sätt. En naturlig lösning blir att bygga datorer som kan förstå och använda information på samma sätt vi själva gör, fast snabbare och/eller bättre.

Det man bland annat försöker uppnå inom forskningsområdet *artificiell intelligens* (AI) är att förmå en maskin att lära sig av situationer som den upplever och med dessa erfarenheter lösa uppgifter. Med andra ord försöker man göra datorn bättre på att hantera information och dra nya lärdommar. Om än inte bättre än oss, så i alla fall på en nivå där den kan lösa nya typer av uppgifter. Tanken är att en AI antigen ska fungera som en assistent till människan, eller helt ersätta människan och fungera helt självständigt.

En AI kan ha olika tillvägagångssätt för att lösa en uppgift och ett stort problem är att avgöra hur pass intelligent en faktisk lösning är. Att mäta kapaciteten på en AI är svårt eftersom dess kunskapsdomän ofta är begränsad och det inte finns något lätt sätt att jämföra intelligensen hos olika algoritmer. Till exempel så är ett klassiskt ”intelligens-testande” problem inom AI att hitta utgången i en labyrint. Vad som är bästa sätt beror på situation och oftast handlar det om att lösa det på snabbast tid, kortast väg eller med minst energiåtgång. Det här har blivit ett populärt scenario att testa kapaciteten på olika AI med eftersom det har likheter

med många andra situationer som kräver lärande och planering. Det anses mäta en relativt generell förmåga även om den inte är i närheten av så generell som ett ”mänskligt” IQ-test eller liknande. Än mer domänspecifika uppgifter för AI kan vara att klassificera fingeravtryck eller spela schack vilka är än lättare kvantifierbara och därmed enklare att avgöra kvalitén på.

Även om inte maskinerna har blivit fullt lika analytiskt begåvade som människor har de ofta snabbheten som fördel. Att kunna behandla miljoner bilder och texter per sekund kan kompensera för en sämre förståelse. Som exempel så skulle inte Googles internet-sökfunktion göras lika bra av en människa eftersom datamängden helt enkelt inte är hanterbar för en person, trots att algoritmen inte har samma kognitiva nivå. Å andra sidan kan inte algoritmen göra det sista urvalet människan gör för att hitta precis rätt sida som söktes. Det här kan exemplifiera hur maskin-människa interaktionen kan ske på ett effektivt sätt: maskinen för-filtrerar informationen och presenterar en lista på möjliga svar för användaren, som i sin tur då behöver bara söka i en väldigt begränsad mängd data. För att läsa mer om människa-maskin-interaktion se FOI rapporten [Johansson 2006].

Redan idag används AI inom vitt skilda domäner såsom bilkörning, spel, ansiktsigenkänning, övervakning, diagnosticering, epidemiprognoser, expertsystem och sökfunktioner. Användningsområden blir fler och fler allt eftersom tiden och utvecklingen går framåt. Området artificiell intelligens har bara funnits i ungefär 60 år vilket är väldigt kort jämfört med många andra forskningsområden. Vi kan vänta oss stora framsteg framöver eftersom området till stor del är outforskat. Dessutom är AI till del beroende av beräkningskraft och kommer att få ta del av den exponentiella utvecklingen som sker där. Även om man fortfarande är långt ifrån en ”människolik” artificiell intelligens har man kommit långt inom många specifika områden. Det är dessa områden den här rapporten ska handla om för att ge läsaren en inblick i den aktuella forskningsfronten och möjliga framtida tillämpningar. Denna rapport sammanfattar kortfattat kunskapsläget, lyfter fram färska genombrott och identifierar trender som kan leda till omvälvande nya tillämpningar samtidigt som vi diskuterar fundamentala begränsningar. Fokus är på teknik grundad i AI som kan påverka samhälle samt försvar och säkerhet. Speciellt viktiga tekniktrender markeras med symbolerna TRL1, TRL2 osv. där TRL står för Technology Readiness Level. TRL1 är lägsta nivå och TRL9 är högsta nivå på teknisk mognad.

2 Artificiell intelligens (AI)

2.1.1 Definition av AI

Att definiera artificiell intelligens är svårt men en vanlig och populär definition är *människolik intelligens*. I detta begrepp ingår det man brukar definiera som specifikt för vår intelligens: självlärande, kreativt, reaktivt och logiskt tänkande, för att bara nämna några aspekter. AI är för närvarande i sin linda och även om många tillämpningar kan ses som intelligenta är de inte självlärande i dess rätta bemärkelse. Till exempel har schack-datorer eller autopiloter inte möjlighet att på egen hand lära sig av situationen de befinner sig i och anpassa sitt beteende även om de på ytan beter sig intelligent. Ett exempel där en schackdator faktiskt skulle kunna kallas ”människolikt intelligent” vore att den skulle kunna sättas framför andra spel eller uppgifter och bemästra dem i samma utsträckning som schack. Den skulle dra nytta av erfarenheter den lärt sig i schack och tillämpa taktiker som passar i den nya domänen.

I dagsläget skulle dock den icke-lärande schack-datorns taktik försämras även om minimala ändringar skulle göras i schack-reglerna. På samma sätt har en autopilot problem med att flyga om icke-förprogrammerade scenarier uppstår. I båda dessa scenarier handlar det om en oförmåga att anpassa sig efter en situation som maskinen inte har sett förut genom att använda lärdommar ifrån tidigare erfarenheter, vilket människor fortfarande excellerar i. Med det sagt finns det självklart schackdatorer som är självlärande och flygplan som kan lära sig hantera oförutsedda situationer – men de är nästan alltid sämre än den mindre intelligenta motparten i en given, förutsedd situation. Tekniken är inte mogen än för att fullt ut använda självlärande metoder istället för deras förprogrammerade motsvarigheter, men inom många områden börjar man närma sig. Till dess gäller en mer pragmatisk inställning för kommersiella tillämpningar. Ofta benämner man dessa system som s*vag AI* där intelligent teknik används men det finns inte så mycket lärande eller tänkande i algoritmen. Fortfarande kan dock en svag AI vara bättre än människor på specifika uppgifter även om de saknar dessa element. Att världens bästa schackspelare inte har en chans mot de bästa schackdatorerna visar att man inte alltid behöver den självlärande aspekten utan inom många områden har man nått långt ändå.

I andra ändan av spektrumet har vi lösningar som kräver just lärande och tänkande för att lösa en uppgift. Dessa problem antas inte komma att lösas innan det centrala problemet inom AI har blivit löst: att få maskiner som fungerar på samma nivå som den mänskliga hjärnan. Man benämner ofta algoritmer som har nära mänsklig intelligens som *stark AI*. Exempel på egenskaper en stark AI förväntas ha är abstrakt resonerande och språk-förståelse (Se kapitel 3).

Gränsen mellan traditionella metoder och AI kommer att suddas ut med tiden då tekniken fortsätter att utvecklas. Eftersom AI på många områden bara är en del av helheten kommer vi i denna rapport även att beröra angränsande områden. Typiska exempel på det är robotik där den fysiska roboten i sig inte har något med AI att göra, men ska den fungera i praktiken måste AI-tekniker implementeras i mjukvaran. Precis som i många andra områden har AI inte slagit igenom full ut där men när den väl gör det kommer den troligtvis ha en stor påverkan hur problemen inom robotik löses. I den här rapporten kommer vi att använda en relativt bred definition på vad AI är för att överlåta åt läsaren att avgöra vad som är väsentligt utifrån dennes intresse.

2.1.2 Var tillämpas AI

Som nämnt tidigare finns de flesta tillämpningar än så länge inom svag AI. AI i bemärkelsen självlärande är en ganska ovanlig företeelse för tillfället men trenden pekar mot att det blir fler och fler. Bild- och röstigenkänning är till exempel områden som kommer starkt och som redan används kommersiellt av företag som Google, Apple och Facebook. De flesta personer kommer i kontakt med lärande AI-tekniker på en daglig basis utan att veta om det: riktad reklam i appar och föreslagna låtar på musiktjänster bygger till exempel ofta på vad datorn har lärt sig om användarna.

2.1.3 Olika typer av AI-metoder

Det finns ett stort antal sätt att simulera intelligens. Här följer en lista på några av de mest populära metoderna för att ge läsaren en känsla för hur olika metoder kan se ut. Rapporten kräver dock inte teknisk förståelse för dessa metoder.

- Neurala nätverk – Inspirerade av den mänskliga hjärnan med neuroner som beräkningsenheter och kopplingar mellan dessa. Har fått en renässans på senare år och används flitigt i bland annat bildigenkänning med extremt bra resultat. Ofta pratar man om djupa neurala nätverk vilket innebär att man har flera sammankopplade lager av neuroner.
- Bayesianska nätverk – Bygger på grafstrukturer där kopplingar beskriver relationer mellan stokastiska variabler. T.ex. kan sannolikheten att ha en specifik sjukdom beräknas med hjälp av symptomen. På samma sätt kan även olika faktorers (t.ex. sanitet, kontakt med virus) sannolikhet att orsaka en specifik sjukdom beräknas.
- Svärmintelligens – Bygger på hur naturen löser uppgifter med hjälp av en grupp av individer som myror eller bin genom att utnyttja samarbetet för att uppnå synergieffekter.

- Support Vector Machine (SVM) – En matematisk metod för att klassificera olika data i kategorier. En metod ifrån mitten av 90-talet som fortfarande är väldigt effektiv och kraftfull i många användningsområden.
- Evolutionära metoder – Metoden bygger på hur organismer som löser uppgifter utvecklats i naturen. Av en grupp slumpvis genererade lösningar på ett problem tillåts de bästa ”överleva” och mutera för att sen korsa sig med andra lösningar och ge ”avkomma” i form av nya lösningar. Efter flera generationer av detta hittas en lösning som (förhoppningsvis) är bra nog.

2.1.4 Beräkningskraft

Man kan inte göra bra prognoser om hur AI utvecklas utan att titta på beräkningskraften eftersom den i hög grad ligger till grund för vad AI kan åstadkomma. Som exempel så har artificiella neurala nätverk (ANN) fått lite av en renässans av främst två anledningar: dels har man hittat en lösning på hur man kan lösa s.k. icke-linjära problem (till skillnad på linjära problem som är lättare men mindre generella), och dels har beräkningskraften 100 000-faldigats på 20 år vilket har bidragit extremt mycket till utvecklingen: 1993 arbetade den snabbaste superdatorn i 60 GFLOPs (se sidfot[1]), idag ligger den över 10 PFLOPs. Estimat av datorkraft är att vi vid 2020 har EFLOPs kapacitet och ZFLOPs kapacitet vid 2030. För att sätta det i perspektiv så är ungefärlig beräkningskraft för att göra korrekt global väderprognos i två veckor framöver 1 ZFLOPs [DeBenedictis 2005]. Denna utveckling förutsätter att Moores lag (dvs. att beräkningskraft dubblas var 18:e månad) fortsätter. Enligt en RAND studie finns det inga tecken att den Moores lag inte fortsätter gälla fram till 2020 och därefter [Silberglitt 2006].

En vanlig stationär dator ligger idag i området kring TFLOPs. Beräkningskraften har alltså nått en mänsklig hjärna där uppskattningen av beräkningskraften ligger på runt 1 TFLOPs-1 PFLOPs. Uppskattningen bygger på analys av antalet neuroner, antalet kopplingar mellan dessa, och med frekvensen som hjärnan skickar signaler. De här siffrorna är dock mycket spekulativa eftersom man inte vet exakt hur hjärnan fungerar men författarna stödjer det här estimatet som rimligt. Accepterar man detta estimat finns alltså ingen begränsning i dagsläget vad beräkningskraft anbelangar utan nu finns den teoretiska möjligheten att göra AI i nivå med mänsklig beräkningskapacitet. Frågan är dock om vi kan göra algoritmer som är i nivå med vad som har utvecklats under hundratals miljoner år av evolution.

---

[1] Beräkningar per sekund: 1 GFLOPs = 10^9, 1 TFLOPs = 10^{12}, 1 PFLOPs = 10^{15}, 1 EFLOPs = 10^{18}, 1 ZFLOPs = 10^{21}

2.1.5 Forskningstrender och prognos

Generellt kan man säga att AI alltid kommer att ha lättare att hantera väldefinierade uppgifts- och problemområden. Svårare uppgifter för AI är situationer där resonemang och situationsförståelse. Exempel på ett relativt väldefinierat problem är autonoma bilar som opererar i en typ av rutnät med tydliga (trafik)-regler och begränsade möjliga scenarier. Motsvarigheten där problemet kräver mer situationsförståelse skulle vara en robot som måste navigera i en okänd miljö t.ex. en delvis kollapsad byggnad för att söka efter överlevande. Där måste den bland annat kunna skilja på människor och material på marken innan den kliver, prioritera sökområden som är mer sannolika att innehålla överlevande, och prioritera behoven hos dessa. Samtidigt som den ska navigera och röja upp en väg fram utan att skada eventuella begravda. Av dessa skäl kommer troligtvis den förstnämnda tekniken (bilkörning) att slå igenom tidigare om man inte använder metoder för att underlätta för räddningsroboten. Det kan man göra till exempel genom att förenkla uppgifterna eller låta en människa ta de taktiska besluten.

Att göra en teknisk prognos om hur tekniken ser ut i framtiden är svårt av flera anledningar. För att hjälpa läsaren att tolka prognoserna i den här rapporten är det viktigt att tänka på en rad aspekter:

- Befintlig/obefintlig teknik: Att kvantifiera när en nästan icke-existerande teknik finns är svårt eftersom det finns för många okända faktorer. Där en teknik redan delvis finns (till exempel autonoma bilar) är det lättare att se vad som behövs för att den ska bli praktiskt användbar, och när ungefär det kan tänkas komma.
- Mognadsgrad: Hur bra måste tekniken vara för att vara användbar? Som exempel är dagens automatöversättning användbar i vissa begränsade områden, men långt ifrån tillfredställande inom andra mer krävande tillämpningar. Det kommer alltid vara en flytande gräns mellan när en teknik blir användbar (för någon tillämpning) och när den är fullt utvecklad (för alla tillämpningar).
- Användbarhet/Kommersialisering: Hur bred kommer användningen av tekniken vara? Det var till exempel svårt att på 80-talet uppskatta det enorma genombrottet internet har fått medan många andra geniala, men inte så användbara, uppfinningar har fallit i glömska. Ett skäl är att inte bara tekniken i sig spelar roll, utan även har samhället tar den till sig – den så kallade Metcalfes lag säger att värdet på ett telenät växer med kvadraten på antalet användare.
- Trender: I som mycket annat går det även trender inom forskningsområden. Ett exempel inom AI är området ”neurala nätverk” där en bok i slutet av 60-talet [Minsky 1969] delvis tog död på mycket

av forskningsanslagen. Den drog slutsatsen att neurala nätverk inte är användbara på grund av en rad uppräknade tillkortakommaden. Idag vet vi att det här var i mångt och mycket en felaktig slutsats men det visar hur enkelt en trend kan försvinna utan egentlig anledning.

- Synergieffekter: Många områden inom AI överlappar till stor del, t.ex. använder sig övervakningssystem och drönare av bildigenkänningsalgoritmer (som beskrivs i en separat del). Den här effekten gör att när det sker en stark utveckling inom ett område möjliggör det i sin tur användning inom andra områden.

I den här rapporten kommer vi därför inte att försöka dra för stora slutsatser om tidshorisonten och de praktiska användningarna i framtiden utan mer beskriva tekniken i sig. Vi gör ibland en försiktig prognos där det kan vara motiverat men vi vill reservera oss för felaktigheter just på grund av ovan nämnda anledningar.

2.1.6 Etiska aspekter

Precis som med andra tekniker så har mycket av det som nämns här potentiella etiska problem. Speciellt gäller detta AI-tekniker som används för beslut där människan är frånkopplad och övervakning. Det kan även finnas indirekta effekter som gör att människor blir negativt påverkade. I denna rapport bortser vi helt från dessa aspekter eftersom inriktningen är att beskriva befintliga och möjliga tekniker så brett som möjligt. Vi bjuder dock in läsaren att fundera på dessa aspekter på egen hand.

3 Tillämpningar av AI och framtida användningsområden

I detta kapitel försöker vi ge en bild av vilken nivå AI ligger på i olika användningsområden, och en uppskattning hur de utvecklar sig i framtiden. Många områden överlappar dock, så valet av rubrik har inte alltid varit självklart.

3.1 Autonoma farkoster

C:a 100 miljoner kilometer ifrån jorden kör en robot omkring på en röd planet och tar bilder, analyserar materia och utforskar sin omgivning. Den styr sig själv till stor del eftersom avståndet gör det praktiskt omöjligt att detaljstyra den ifrån jorden eftersom styrsignalen tar flera minuter att nå fram. Därför har man till stor del fått överlåta ansvaret till roboten själv [Estlin 2007]. Människor kontrollerar fortfarande det övergripande uppdraget men hur den tar sig från punkt A till punkt B sköter den själv. *Curiosity*, som roboten heter, är kanske det mest kända exemplet på en autonom robot som utför uppgifter som förut bara en människa kunde klara av. Tillsammans med sina kusiner Opportunity och Spirit gör de Mars till den första planeten i vårt solsystem bebodd av endast robotar. Dessa robotar är bara ett exempel på möjliga tillämpningar som har kommit till på senare år. Vi går här igenom olika typer av autonoma farkoster och deras användningsområden.

3.1.1 Autonoma robotar

DARPA challenge[2] syftar till att stimulera utveckling av självgående robotar som kan klättra, flytta skräp, navigera och utföra andra uppgifter för att kunna assistera och även ersätta mänskliga räddningsarbetare vid svåra katastrofer. Det kräver att robotarna ska kunna greppa och lyfta bort tunga föremål, balansera i trappor, undvika hinder, och andra uppgifter en människa kan göra. Dagens generation robotar är dock långt efter människor i många av de här uppgifterna. Det beror dels på sämre utvecklad motorrik, dels på sämre mekaniska möjligheter och tålighet. Till exempel kan en robot inte lyfta och bära en människa på egen hand samtidigt som den utför uppgifter som att öppna dörrar och liknande. Snarare råder det specialisering, till exempel att en robot kan gå över ett hinder, en kan öppna dörrar, en kan bära en människa, en kan skruva på ett reglage, en kan klättra i en stege och ytterligare en kan lyfta hundratals kg. Detta gör att den praktiska tillämpningen inte är där än men inom en snar framtid

[2] http://www.theroboticschallenge.org/

hoppas man paketera dessa egenskaper i en enda robot. Utvecklingen går snabbt framåt just nu och det finns det många forskningsprojekt inom detta område även om DARPA challenge är det kanske mest kända. Inom några år kommer vi med stor sannolikhet att se autonoma robotar utföra uppgifter som förut var explicit för människor, och även komplementera människor med olika färdigheter. **TRL: 4-6**

Industrirobotar har också utvecklats starkt på senare år, dels börjar tekniken bli mogen, dels av ekonomiska skäl. I många år har Kina haft en ledande ställning i tillverkningsindustrin men i västvärlden har företag har börjat hämta hem sin tillverkning och använder då i större utsträckning automatiserade fabriker. Ett bra exempel är bil-industrin som har helautomatiserat stora delar av sin tillverkning. Istället för löpande band körs bilarna fram på små robotar som följer magnetband för mer flexibilitet. Många av dagens industri-robotar är även multi-funktionella och utför flera uppgifter med olika verktyg[3]. Dagens robotar kan även ha inbyggda intelligenta säkerhetssystem vilket innebär att människor kan vistas runtom utan att bli träffade av rörliga delar vilket underlättar för reparation och robot-människa interaktion. IFR (International Federation of Robotics) uppskattade 2012 att totala värdet av industrirobotar är $26 miljarder och ungefär 160000 enheter skeppas varje år med en tillväxttakt på 6 % per år[4]. Vi förutsätter att den här trenden inte bara håller i sig utan även accelererar. **TRL: 6-9**

3.1.2 Autonoma bilar

En av den vanligaste orsaken till dödliga trafikolyckor är fortfarande den mänskliga faktorn. Lägger man på kostnaden för tiden som ägnas åt att styra miljontals bilar varje dag inser man att det är samhällsekonomiskt dyrt med mänskliga bilförare. Ett genombrott inom autonoma farkoster skulle potentiellt kunna öka produktiviteten och frigöra resurser som annars inte skulle kunna användas. Idag har flera företag, bl.a. Google, självkörande bilar som kan ta sig fram effektivt i stadstrafik utan att kollidera med omgivningen. Dessa bilar kör redan i områden med specialtillstånd för förarlösa bilar i Kalifornien och Nevada [Fagnant 2013]. Även om utvecklingen fortfarande är i sin linda så påstår Google att deras bilar har en lägre olycksfrekvens än den genomsnittlige mänskliga föraren. Volvo går ännu längre och förutspår att de har en olycksfri (crash-proof) bil 2020. De flesta stora biltillverkare har nu någon sorts forskning relaterat till bilar som kan agera på egen hand. Vilka som kommer först återstå att se men sannolikt kommer en kommersiell teknologi för säkra självkörande bilar finns redan inom några år. **TRL: 4-6**

[3] http://www.businessinsider.com/inside-the-tesla-factory-2013-8
[4] http://www.ifr.org/industrial-robots/statistics/

3.1.3 Autonoma flygmaskiner

Unmanned Aerial Vehicle (UAV) och Unmanned Combat Air Vehicle (UCAV) är flygmaskiner som antingen kontrolleras ifrån marken av en mänsklig pilot eller kör autonomt via en dator ombord. Även om en mänsklig pilot har varit den vanligare formen tidigare har autonoma UAV:er tagit över mer och mer. Fördelen med en AI som styr systemet är att reaktionstiderna blir kortare och att manövreringsförmågan inte begränsas i samma utsträckning av pilotens begränsande biologiska faktorer såsom vikt, sömnbehov och g-kraft-känslighet. Att göra UAV:er mer autonoma är således ett stort forskningsområde med både militära och civila användningsområden. Redan idag kan de starta, landa och flyga enklare uppdrag på egen hand och forskningen inriktas på mer autonom funktionalitet. Det senaste året har hundratals vetenskapliga artiklar publicerats med nyckelordet UAV. Enligt FAA (Federal Aviation Administration) så finns enbart i USA över 50 företag, universitet och myndigheter som utvecklar mer än 100 olika UAV-modeller[5]. Marknaden från i år (2014) fram till 2023 uppskattas vara värd c:a $90 miljarder (c:a 600 miljarder SEK) där stor del av tillväxten kommer ifrån Ryssland och Kina[6]. Självklart finns det många hinder att övervinna inom området men vi ser redan idag att autonoma flygmaskiner är en högst verklig och använd teknik. Ett projekt som står ut är Stanfords universitets program för en autonom helikopter. I dagsläget används RC-helikoptrar men maskinen som styr uppvisar en extraordinär kontroll över fordonet där en mängd komplicerade manövrar kan utföras[7]. För en analys av obemannade luftfarkostsystem se FOI-rapporten [Forssell 2013]. **TRL: 6-9**

3.1.4 Robotsvärmar och svärmintelligens

Ett område inom AI försöker dra nytta av ett bekant fenomen från djurriket: arter som tillsammans kan bygga komplexa system som till synes kräver mer intelligens än de enskilda individerna har. Välbekanta exempel är myror och bin. Genom att följa enkla regler och kommunicera med exempelvis doftspår kan de kollektivt uppnå en ”högre” intelligens än vad de enskilda individerna har. Regler för hur varje agent ska reagera i en situation gör att svärmen kan interagera i ett komplext beteende för att nå sitt mål. Dagens forskningsområden sträcker sig från planering av telekommunikationsnät [Ducatelle 2010] och satellit-fördelning [Fleischer 2005] till genomsökning av katastrofområden och till att få agenter (t.ex. Quadcopters) att samarbeta för att flytta föremål de inte enskilt skulle förmå [Flushing 2013]. Även inom distribution av varor används svärm-intelligens av företag såsom AntOptima. Så här finns det redan

[5] https://www.asme.org/engineering-topics/articles/robotics/unmanned-aerial-vehicles-soar-high
[6] http://www.bbc.com/news/business-28294239
[7] http://heli.stanford.edu/

användningsområden, men svärmar av räddningsrobotar har fortfarande en bit kvar innan realisation. **TRL: 4-6**

Det har även gjorts forskning om nanorobotar ska kunna fungera inom en mänsklig kropp för att t.ex. distribuera läkemedel eller reparera skador inifrån [Casal 2003]. Svärmar av nanorobotar är dock fortfarande orealiserat eftersom tekniken att tillverka nanorobotar inte finns och om inget genombrott görs så kommer det att förbli så flera år framöver. **TRL: 1**

3.1.5 Sammanfattning autonoma farkoster

Många områden inom intelligenta autonoma farkoster är starkt sammankopplade eftersom problemen som måste lösas ofta är likartade. Dock är det så att vissa problem i civila tillämpningar kan vara enklare att lösa än militära: En civil bil ska undvika krock till nästan vilket pris som helst och stannar då för alla föremål den stöter på. Ett autonomt stridsfordon behöver däremot ofta ha en bättre situationsförståelse: i vissa fall är det korrekt att köra över objekt som är i vägen, speciellt om det tillhör motståndarsidan, medan den också måste ta säkerheten för egen trupp och civilbefolkning i beaktande. Därtill måste en militär farkost lösa uppgifter en civil motsvarighet normalt inte behöver, som verkansinsats och spaning. Å andra sidan är användningen av helt automatiserade fordon i civiltrafik förknippat med etiska och moraliska problem som var det juridiska ansvaret ligger och viljan att flyga med i ett flygplan som helt saknar mänsklig pilot.

Detta sammantaget gör det svårt att säga när tekniken är helt mogen för användning eftersom den beror mycket på politik och samhälleliga strömmar. Den grundläggande tekniken finns dock redan på plats.

3.2 Visuell intelligens

Ett av de största genombrotten inom AI de senaste åren har varit inom bildigenkänning, främst i form av *djupa neurala nätverk* som är inspirerade av den mänskliga hjärnan. Inom många domäner har den nått nivåer som har en lägre felfrekvens och framförallt en mycket högre kapacitet än människan. Detta har lett till ett stort intresse i området som handlar om visuell intelligens. Här listas några av de mest intressanta tillämpningsområden just nu.

3.2.1 Bildigenkänning

Bildigenkänning är förmågan att tolka en bild som presenteras. Det kan vara klassificering, segmentering, relationer mellan objekt, eller annan typ av bedömning. De tänkbara användningsområdena är många: navigering, detektion,

modellering, inspektion, övervakning och människa-robotinteraktion för att nämna några. Det är med andra ord även ett kritiskt område för andra tillämpningar av AI. Typiskt militära användningsområden kan vara detektion av nedgrävda minor eller hemmagjorda bomber vid vägkanten (det var och är alltjämt ett stort problem för de internationella insatserna i Irak och Afghanistan), analys av spaningsbilder eller visuell identifikation av flygplan.

Eftersom den kanske mest intressanta tekniken idag är djupa neurala nätverk så får den i den här rapporten mycket representera var tekniken ligger idag inom bildigenkänning. Inom specifika domäner som trafikmärken [Ciresan 2011] och handskriven text är artificiella neurala nätverk redan bättre på igenkänning än människan.

Ansiktsigenkänning har varit ett svårlöst problem länge på grund av den komplexitet som uppstår när ett ansikte ses från en dålig vinkel, belyses olika och/eller är delvis skymt. Idag används ansiktsigenkänning vid passkontroller men där råder kontrollerade former med avseende på ljus och vinkel. Dessa förhållanden är ganska unika och i vanliga fall har man inte så tillrättalagda bilder så där fallerar ofta ”dagens” teknik. Det finns dock olika företag som har börjat komma långt, ett exempel är Facebooks *AI Group* har utvecklat en algoritm baserad på djupa neurala nätverk (97.2% korrekt klassificering på ett datamaterial på 4000 individer) som börjar närma sig nivån för människor (97.5% korrekt klassificering) [Taigman 2014]. Håller trenden i sig finns det inget som talar mot att ansiktsigenkännings-algoritmer kan användas vid säkerhetskontroller och övervakningskameror för att snabbt identifiera efterlysta personer [Zhao 2003] inom några år. För en mer djuplodande bild se FOI rapporten [Ahlberg 2011]. **TRL: 7-9**

3.2.2 Händelseigenkänning

En teknik som av naturliga skäl bygger mycket på bildigenkänning är händelseigenkänning (Activity recognition på engelska). Eftersom händelser består av en sekvens med bilder som kan tolkas som en helhetsbild så är bildigenkänningen en förutsättning för att förstå vad som utspelar sig i ett scenario. Skillnaden består i att bildigenkänningen känner igen *en människa* medan händelseigenkänningen ska förstå om människan i fråga *springer, hoppar*, eller vad nu denne kan tänkas göra.

Det finns forskning som har kommit så långt att algoritmerna kan avgöra vilken typ av rörelser testpersonen gör. Detta sker dock bland ett begränsat antal tänkbara händelser om kanske ett tiotal olika typer av rörelser där algoritmerna kan känna igen en klar majoritet [Ni 2013]. Dagens teknik kan alltså avgöra ganska enkla beteenden där den kan skilja på folk som hoppar eller går utan kontext. I förlängningen är tanken att algoritmerna ska bli så avancerade så att de

kan avgöra ifall en person rånar, gör inbrott, eller någon annan typ av mer abstrakt aktivitet.

Det finns ännu inga metoder för att på tillfredställande nivå direkt i en övervakningskamera avgöra vad som utspelar sig. Det finns forskning som inriktar sig på den frågan [Ryoo 2011] men det är svårt att avgöra hur lång tid det tar innan det blir kommersiellt eller militärt användbar. Innan det får ett brett användningsområde behöver man även lösa situationsbedömningen: Springer en person för att den är jagad eller försöker hon bara hinna med bussen? Även om det är svårt att bedöma exakt när så finns det inget som talar mot att tekniken kommer att mogna inom en relativt snar framtid. **TRL: 4-6**

3.2.3 Aktiva visuella system

REDOWL (Robotic Enhanced Detection Outpost With Lasers) är ett system utvecklat för att lokalisera krypskyttars position. Med hjälp av laser, avlyssnare, värmekamera, GPS, och fyra kameror kan den aktivt finna ett mål och till och med avgöra vilket vapen som används. Företaget iRobot som utvecklar REDOWL påstår sig ha en 94-procentig träffsäkerhet på positionsbestämningen.

Samsungs *SGR-A1* och *DoDAMMs Super aEgis* är sydkoreanska militärrobotar som är automatiserade stationära kulsprutor med effektiv räckvidd på flera kilometer. De används för att detektera och skjuta mot objekt i den demilitariserade zonen mellan Sydkorea och Nordkorea[8].

De här systemen kan dock idag liknas mest vid personminor med extra lång räckvidd och riktad verkanseld eftersom de inte gör skillnad på vem de skjuter på om inte speciella markörer bärs. Det gör användningsområdet ganska begränsat för sådana här system både på grund av etiska och praktiska skäl. För att göra en bedömning när aktiva visuella system blir användbara måste man titta på området visuell intelligens i sin helhet. **TRL: 3-5**

3.2.4 Situationsbedömning

Idag används ofta autonoma drönare för att göra situationsbedömning i områden som av olika anledningar är oåtkomliga. Situationsbedömningen i sig görs dock oftast av en mänsklig användare i andra ändan av kommunikationslänken. Naturligtvis är det här ett system som låser upp mycket mänskliga resurser. Vad man vill uppnå med AI är att ersätta människan med en maskin som autonomt kan arbeta med analys av området. DARPAs projekt ”Combat zones that see” (CTS) använder AI-tekniker för att övervaka och få överblick över ett tredimensionellt operationsområde [Lyon 2006]. Målet är att kunna följa alla agenter i ett givet system och bedöma deras hotbild på ett adekvat sätt. Det här

[8] http://www.globalsecurity.org/military/world/rok/sgr-a1.htm

bygger självklart på bildigenkänning och händelseigenkänning, men det tillför ännu en dimension i problemet eftersom en förståelse i det stora hela måste till. Som nämnt tidigare så kan en ensam löpare betyda ingenting medan en större grupp kan vara både ett upplopp eller en löpartävling.

Ett annat DARPA projekt är *Mind's Eye*[9] vars mål är att delvis ta ur människan ”ur bilden” för att frigöra mänskliga resurser, och även upptäcka saker som människor inte är kapabla till att göra. Hela 12 stycken forskningslag jobbar med den här uppgiften sedan 2011 så det kan vara intressant att följa just detta projekt för att se hur området situationsbedömning utvecklar sig.

Inga av de här projekten inom situationsbedömning har till dags dato gett betydande resultat vad författarnen känner till men det visar att forskning pågår och att investerare har gott hopp om att det leder till resultat. Situationsbedömning är ett notoriskt svårt problem och kommer det inte något oväntat genombrott så kan det dröja ett tag innan det blir ett användbart verktyg.
TRL: 1-3

3.2.5 Sammanfattning visuell intelligens

Sammanfattningsvis så ser vi att bildigenkänning börjar bli ett riktigt intressant område att följa där tillämpningarna för AI blir fler och fler. Vad som kommer härnäst, men inte riktigt är där än, är händelseigenkänning där vi ändå börjar se några potentiella användningsområden. Aktiva visuella system och situationsbedömning är fortfarande i forskningsstadiet så det är svårt att avgöra när tekniken kommer att slå igenom utan att spekulera i metoder som inte finns än. När tekniken väl finns på plats kommer det att möjliggöra helt automatiserade bevakningssystem för bland annat:

- Säkerhetskontroller: helautomatiserade pass- och gränskontroller. Förutom ansiktsigenkänning skulle även stressnivåer och nervositet kunna utläsas visuellt för att hitta smugglare och liknande.
- Brottsutredning och gatusäkerhet: på samma sätt som fortkörningskameror fungerar skulle ett system utrustat med ansiktsigenkänning direkt kunna larma när ett brott sker tillsammans med information om förövarens namn, adress och kända medbrottslingar och annan information.
- Militära operationsområden där tidig varning för anfall är kritiska. Systemen skulle kunna snabbt ge information om bland annat omfattning och utrustning.

[9] http://www.darpa.mil/Our_Work/I2O/Programs/Minds_Eye.aspx

3.3 Intelligenta hjälpmedel

Under intelligenta hjälpmedel har vi samlat några system som med hjälp av AI tillhandahåller *stöd* till användaren. Dessa är således oftast inte autonoma även om de i princip kan vara det när det behövs.

3.3.1 Exoskelett

Ett exoskelett är en utrustning som bärs som ett slags yttre skelett och kan tillföra kraft till användaren [Science daily 2013]. Möjliga användningsområden är kraft-utväxling, som kan användas för att utföra tunga moment som inte kan/bör göras av robotar, och rehabilitering/handikapstöd för att ge förlamade människor möjligheten att gå [Huo 2009]. Det sker relativt mycket på den här fronten för tillfället: EU-projektet Robo-mates (www.robo-mate.eu) mål är att utveckla ett intelligent, lättstyrt, och bärbart exo-skelett. Exempel på kommersiella företag inom området är Lockheed Martin, Raytheon och Cyberdyne. DARPA har även ett projekt som syftar till att göra det möjligt för soldater i fält att bära tung utrustning flera gånger sin egen vikt. Utmaningen med dessa system är att förmå det att samarbeta med bäraren så att maximal kraft och flexibilitet uppnås utan att skada användaren. För att göra detta måste en förståelse finnas för vilka rörelser bäraren vill utföra så att samarbetet blir smidigt och effektivt. I förlängningen kan det även finnas behov att operera autonomt om t.ex. bäraren i strid skulle förlora medvetandet eller förlora funktionen i vissa kroppsdelar och behöva transporteras ur området. Det finns redan exo-skelett som hjälper förlamade att gå, dock med hjälp av kryckor som stöd så tekniken finns där men det fattas funktionalitet innan vi ser en bredare användning. **TRL: 3-6**

3.3.2 Proteser

Inom medicin har artificiella kroppsdelar funnits i årtusenden där främst förlorade extremiteter har ersatts av proteser. Det är först de senare åren med den tekniska utvecklingen proteser som kan aktivt utföra komplicerade rörelser har kommit. En vanlig typ av aktiva-proteser känner av aktivitet i musklena som fortfarande finns kvar i lemmen och förmedlar signalen ut till motorer. På så sätt kontrollerar hjärnan (indirekt) rörelserna i protesen. För att förbättra känsligheten finns försök att styra direkt med signaler ifrån hjärnan men än så länge dominerar de indirekta metoderna. Det är än så länge väldigt svårt att extrahera signaler från hjärnan och här kan mycket väl intelligenta system hjälpa till med behandlingen och tolkningen av signaler som hjärnan skickar ut. I takt med att tekniken utvecklas blir proteserna mer och mer förfinade och det kräver att intelligensen förbättras i kretskorten som styr motorerna för att uppnå den precision vi har i

biologiska fingrar.Moderna proteser kan vara extremt dyra men prototyper har visat att 3D-printers kan skapa en aktiv protes för några hundra dollar[10] vilket kan öppna upp ett användningsområde i de mer fattiga delarna av världen. **TRL: 4-9**

3.3.3 Sjukvård

Operationsrobotar som Da Vinci [Da Vinci1] är robotar som styrs av en utbildad mänsklig kirurg för att göra operationer. Fördelen med ett sådant system är flera: mindre ingrepp kan göras med samma effekt, lederna på robotarmarna har större frihet att roteras än deras mänskliga motsvarighet, och kameror med högre upplösning än det mänskliga ögat kan användas. I princip behöver inte operatören ens vara på samma kontinent vilket möjliggör teleoperationer där kirurgiska experter kan verka globalt och personal alltid finns tillgängligt. Ett problem man behöver lösa är vad som ska ske när länken bryts mellan operatör och robot, eller fördröjningen blir för stor i överföringen, eftersom det idag saknas inbyggd autonomi i systemen. **TRL: 4-6**

Ett naturligt steg i utvecklingen är förstås att förse dessa robotar med någon slags intelligens så de kan utföra lättare sysslor eller assistera vid komplexa rörelser. Autonoma operationsrobotar finns för tillfället inte i bruk vad vi känner till, men det finns forskningsprojekt som har visat lovande resultat. En prototyp av en robot som kan sy ihop en patient efter utfört ingrepp har visat lovande resultat [Mayer 2006]. Inom en relativt snar framtid är det mycket möjligt att i alla fall enkla ingrepp kan göras av robotar utan att intervention eller övervakning av människor krävs. **TRL: 1-3**

3.4 Cyberhot

Konventionella attacker med soldater och stridsfordon kräver ofta stora resurser och folkligt stöd och är därför en oattraktiv metod för många stater eller organisationer. Cyberattacker däremot kan vara billiga, snabba och avsändaren kan potentiellt förbli anonym och göra stor skada, speciellt mot ett högteknologiskt land. Ett idealt mål för mindre nogräknade aktörer är till exempel ett lands kraftnät som slår mot ekonomin, moralen och försvarsförmågan samtidigt.

Exempel på hot är DDOS-attacker som görs med datorer i ett botnet, dvs. infekterade datorer som fjärrstyrs av en användare. Ett av de mer kända exemplen är kanske *Low Orbit Ion Cannon* (LOIC) attacken i 2010 som slog ut

[10] http://www.robohand.net/

MasterCard och Visa. Det iscensattes med hjälp av en svärm datorer, s.k. *botnets*, som var under kontroll av en organisation. I framtiden när svärm-teknologin (beskriven tidigare i 3.1.4) är mer utvecklad kan det öppna upp möjligheter för mer synergieffekter för större och potentare attacker.

Nedan beskriver vi kort ett av möjliga cyberhot där AI kan användas, och hur AI skulle kunna tänkas användas i försvar. För en specifik analys om cybersäkerhet ombord på en obemannad farkost se FOI rapporten [Axell 2013].

3.4.1 Intelligenta virus-attacker

I januari 2010 i Iran upptäcktes en märklig företeelse av inspektörer ifrån IAEA: centrifugerna som används för att anrika uran byttes ut i en takt som vida överskred den vanliga frekvensen vilket de borde behövas repareras. Någon eller något orsakade skada utan att upptäckas av vare sig vakter eller militär. Ett par månader senare upptäckte ett antivirusföretag ett virus som lite senare skulle visa sig ligga bakom sabotaget av urananrikningsanläggningen, viruset fick namnet *Stuxnet*. Det visade sig att Stuxnet attackerade specifikt Irans kärnkraftsprograms datorer genom att söka upp och identifiera Siemens step7-mjukvara (en mjukvara som endast används inom industriella styrsystem). Styrdatorerna i sig var inte kopplade till internet men viruset spred sig via USB-stickor från datorena med step7-mjukvaran in till kontrollsystemet. Väl på plats rekognoserade viruset systemet i två veckor innan det aktiverade sig själv. Det visade sig att viruset bar med sig information om hur målet såg ut och alla datorer som inte överensstämde med det tilltänkta målet lämnades oberörda och viruset sökte sig vidare. Viruset hade skapats för att anfalla en specifik anläggning och lämna allt annat oskadat, och väl på plats gjorde den minimala förändringar i hur styrsystemet opererade för att göra en så stor skada på infrastrukturen som möjligt. På detta sätt kunde det operera bakom linjen hos en främmande makt utan att upptäckas [Albright 2010].

Stuxnet är ett av det mest kända intelligenta virusen som har används i syfte från en nation att anfalla en annan. Det visar skadan som ett intelligent virus kan göra hos en oförberedd motståndare. Med en låg risk för upptäckt och repressalier ifrån det attackerade landet är det även attraktivt för många icke nogräknade stater och organisationer, till och med privatpersoner. Det är sannolikt att det här inte är en enskild händelse utan i och med den potential den uppvisade kommer liknande virus att användas i framtiden [Libicki 2009]. Dessa virus kan även komma att kombineras med AI-tekniker som på egen hand lär sig om systemen när det väl har kommit in, och kan skriva om sig själv för att undvika att bli identifierat och raderat. **TRL: 6-9**

3.4.2 Intelligent försvar

Artificiell intelligens kan inte bara potentiellt användas till attacker utan kan vara ett nödvändigt verktyg vid försvar mot just sådana attacker. De flesta av oss kommer dagligen i kontakt med ett framgångsrikt sådant försvar: e-postfilter som sorterar ut skräppost är självlärande mekanismer som idag har mycket god precision. När det gäller direkta attacker mot infrastruktur som kraftnät så kan en intelligent försvarsmekanism fungera som ett virituellt försvar som bara finns i systemet för att stoppa eller detektera intrång. Sådana system kallas för *intrusion detection system* (IDS) [Scarfone 2007] och man kan med hjälp av AI utöka dess förmåga [Lin 2012]. Det finns främst två metoder för detektion av attacker, där en bygger på signaturen hos attacken och främst används för att upptäcka virus. Den andra är anomalidetektion som detekterar ovanliga förändringar gentemot en normalbild. Ett problem med IDS:er är att de måste konfigureras noga för att inte ge ifrån sig ohanterliga mängder falsklarm. Enkla och intuitiva filter fungerar dåligt, men som på många andra området skulle AI-metoder kunna hjälpa mänskliga analytiker att selektera och sålla i materialet [Sommestad & Franke, 2014]. Precis som man har fysiska gräns- och säkerhetsvakter så har alla seriösa organisationer idag ”virtuella säkerhetsvakter” i form av bland annat IDS:er för att skydda sig mot attacker på nätet. Användning av AI-metoder för att sålla i stora mängder data och ge mänskliga beslutsfattare en lägesbild av tillståndet i de system man ska försvara är ett stort och växande forskningsområde [Franke & Brynielsson 2014]. **TRL: 4-6**

Ett annat intressant område är självläkande mjukvara (eng. *self-healing and self-protecting software*). Här används mer eller mindre intelligenta metoder för att låta system själva identifiera nya eller okända hot och fel, samt åtgärda dem dynamiskt. Detta är särskilt intressant i verksamhetskritiska system där tillgängligheten ofta måste prioriteras före korrigering av mindre allvarliga problem. Dessa metoder kan också tillämpas i system som kräver hög nivå av autonomi, såsom rymdsonder och obemannade farkoster [Rodhe et al. 2014]. **TRL: 2-4**

3.4.3 Summering cyberhot

Dagens virusprogram hittar oftast virus genom att jämföra signaturer hos känd skadlig kod med den exekverbara koden. Ett intelligent virus som ständigt kan byta form och ”evolvera” blir därmed väldigt svårt att ta bort och en eventuell cyberattack kan förlängas avsevärt. Stuxnet har visat hur pass intelligent ett virus kan skrivas och hur kirurgisk precisionen kan vara när kunniga personer utvecklar mjukvaran.

En nackdel med ett system byggt på AI är att den inte är lojal utan lyder bara programkoden. Byts programkoden ut av en antagonist blir maskinen som AI:n kontrollerar, t.ex. en UAV, operationsrobot eller en bil helt i händerna på

motståndaren. Lyckligtvis är sådana attacker där kontrollen tas över svåra om rätt förberedelser har gjorts – men då krävs att man är medveten om risken. Indirekta attacker som att störa GPS:en eller att slå ut elektroniken helt [Hartmann 2013] är lättare och således ett ännu större problem. Vi ser cyberattacker med hjälp av AI som ett av de större hoten mot samhället där AI är inblandat. Å andra sidan kan AI användas för att avvärja cyberattacker vilket gör det till ett tveeggat svärd. Idag använder alla delar av både applikations- och nätverkslagren olika AI-tekniker för detektera attacker – på applikationsnivån har vi exempelvis spam-filter, på datornivån har vi en sorts IDS och på nätverksnivån en annan sorts IDS [Wu & Banzhaf 2010]. För exempel på artikel på hur skador av DDoS attacker med botnets kan mildras se [Lua 2011].

3.5 Övervakning

Övervakningskameror som kan känna igen slagsmål, rån, upplopp, smuggling, inbrott för brottsbekämpning är ett prioriterat forskningsområde som beskrivet i 3.2. Fast det idag inte finns i något system som helt kan ersätta människor så pågår forskning som har nått bra resultat i igenkänning av olika mänskliga aktiviteter. Med de framsteg som finns och om trenden med förbättrad bildigenkänning håller i sig kan vi vänta oss att det kommer finnas system som mer och mer tar över människans roll i övervakning. Nedan följer några större områden som berör övervakning.

3.5.1 ”Tracking” via övervakningskameror

Att följa personer via övervakningskameror är ett forskningsområde som har kommit mycket med förbättringarna inom bild- och ansiktsigenkänning. Syftet är att kunna följa personer i realtid för att underlätta vid t.ex. brottsbekämpning och personsäkerhet. Fortfarande är det dock ett svårlöst problem och en mänsklig övervakare är fortfarande överlägsen, speciellt när personen går ur bild och sedan dyker upp i en annan kamera [Madden 2007]. Försök i London har gjorts för att följa personer längre sträckor via olika övervakningskameror automatiskt men fortfarande är det en bit kvar till felfri spårning. Med de framsteg som finns bedömer vi det som sannolikt att system som kan övervaka stora delar av städer och områden kommer inom 15 år. **TRL: 4-6**

3.5.2 Anomalidetektion

Varningssystem för t.ex. kärnkraftverk kan genom smartare detektion tidigare hitta felaktigheter eller varna om med framtida möjliga fel innan de händer. Traditionellt sätts ofta fördefinierade gränsvärden varpå systemen varnar när de överskrids, vilket inte ger ett alltför dynamiskt och intelligent varningssystem.

Ett mer intelligent system kan tidigare se när t.ex. många värden börjar bete sig ovanligt genom att se helheten, även om de är långt ifrån det hårdkodade maxvärdet de får ha. Det kan till exempel ta sig i uttryck av en plötslig hög aktivitet i ett område eller mätvärden som vanligtvis har en lägre aktivitet. Det kan tidigare stoppa system som annars tar stor skada av att operera suboptimalt under lång tid. Att försöka hitta mönster som frångår det normala kallas anomalidetektion och kan användas för att upptäcka fel som annars inte hade upptäckts men andra metoder.

Det finns främst tre olika kategorier av anomalidetektion:

1. Oövervakad anomalidetektion: algoritmen utgår främst från att majoriteten av data utgör normaltillståndet och varnar därmed när detta frångås. Den uppenbara bristen med den här blir när normaltillståndet varierar. Som exempel kan man ha obehöriga på ett område under natten där normalt det rör sig många människor på dagen. Det skulle göra att systemet, om det inte förstår konceptet ”natt”, inte ser att det är onormalt och därmed missar intrång.
2. Övervakad anomalidetektion: algoritmen tränas på vad som är normalt tillstånd och vad som inte är det. Fördelen är att man kan använda en människas erfarenhet för att styra beteendet men nackdelen blir att algoritmen inte anpassar sig efter förändringar i det övervakade systemet. Det kan även bli fel i situationer som den inte har tränats för.
3. Semi-övervakad anomalidetektion: en modell byggs runt vad som är normaltillståndet och när data kommer in beräknas en sannolikhet att den skulle kunna genereras av modellen. Är sannolikheten låg är det troligtvis något som är utanför normaltillståndet och därmed en anomali. Svårigheten blir då att konstruera en modell som representerar systemet på ett adekvat sätt.

Mer avancerade varianter av anomalidetektion tar hänsyn inte bara tillsystemets normala tillstånd, utan även till normala övergångar mellan dessa tillstånd och normala tidsintervall för övergångarna.

Det är tydligt att alla metoder har sina för och nackdelar men används rätt metod för situationen kan den vara mycket effektiv. Anomalidetektion kan inom övervakning användas inom till exempel:

- Intrångsdetektion: intrång kan potentiellt upptäckas tidigare i om med att förberedelser för intrång ger ofta upphov till ett ”onormalt” beteende innan intrånget sker eller till och med när planering sker.
- Patientövervakning: som beskrivet tidigare (i 3.3) kan anomalidetektion leda till bättre patientvård där onormala nivåer av värden kan ge indikationer att något är på väg att inträffa som t.ex. innan en hjärtattack eller bara en allmän försämring av hälsotillståndet. Det öppnar upp

möjligheter till tidigare detektion och därmed preventiv vård. Samma princip kan användas för att detektera exempelvis fallolyckor med äldre i hemmet [Savage et al. 2012].

- Bankintrång och pengatvätt: banker och spelsiter lägger redan stora pengar på att upptäcka och förbygga felaktig användning av deras system. Här används anomalidetektion till en stor utsträckning för att fånga upp beteenden analytikerna inte har tid eller kapacitet för att göra.
- Varningssystem: kärnkraftverk, lastbilar och flygplan är några av de många komplexa system som består av många samverkande delar. Det är svårt att fånga upp alla sätt de systemen kan gå fel på så där finns ett stort användningsområde [Bitéus 2007].

Anomalidetektion är ett enormt område så det är svårt att ge någon överskådlig bild. Tekniken används redan idag och utan tvekan kommer den att användas i framtiden men till vilken utsträckning är svår att svara på. **TRL: 6-9**

3.5.3 Sjukvård

Övervakning av patienter med stöd av datorer har funnits länge men även här finns en uppenbar tillämpning av AI där uppgifterna passar maskiner väldigt bra: det är ett väldefinierat problem med gränsvärden och nivåer som är kvantifierbara. Maskinerna kan till exempel fungera som assistenter och påminna om personalen har missat att göra någon medicinering eller procedur under ronden. Hjärtstartare som bara behöver fästas på bröstet för att optimera defibrillering och sluta när hjärtat börjat slå finns sen länge [Kerber 1997] och kan göra skillnad på liv och död när hjärnskador tar ett fåtal minuter att få och ambulanspersonal är långt bort. I förlängningen finns det ingen som talar mot att autonoma system kan bli ännu mer effektiva och kan sköta doseringar, behandlingar och diagnostiseringar på egen hand. Utan domänkunskap inom medicin och sjukvård är det dock svårt att säga hur långt automatiseringen kan gå. **TRL: 6-9**

3.5.4 Övervakning av samhällsfunktioner

I maj 2010 föll Dow Jones plötsligt med 1000 punkter (motsvarande c:a 9%) för att inom några minuter återhämta sig till normala nivåer igen. Det här var det största intra-dag fallet i Dow Jones historia någonsin. Även om stora fall på börsmarknaden har hänt förr var det mest anmärkningsvärda hastigheten på fallet [Easley 2011]. Det är fortfarande omtvistat om vad som egentligen hände men troligtvis var anledningen maskiner som utförde så kallad högfrekvens-handel (HF-handel). I HF-handel varar positioner ofta inte mer är några milli- eller mikrosekunder och saknar egentlig övervakning av människor på grund av hastigheten. Om de var bakomliggande orsak eller bara bidrog till eskaleringen är

oklart men utan automatiserad handel hade förloppet varit långsammare och mer överskådligt. Förloppet visar att utan säkerhetsspärrar kan oreglerade algoritmer kan göra stor skada utan att någon hinner reagera. I värsta fall kan dessa börskrascher leda till instabilitet i finansiella systemet och orsaka stor ekonomisk, samhälles och personlig skada. AI-system används sen länge i HF-handel och kommer att bli mer använt med tiden om trenden håller i sig. Ska övervakning ske så måste ett system som inte är beroende av mänsklig input för att ta beslut användas, eftersom ingen människa är snabb nog. Det talar mycket för att det kommer vara just AI-algoritmer som kommer att övervaka andra AI-algoritmer i framtiden. **TRL: 4-6**

3.5.5 Sammanfattning övervakning

Med AI har det inte bara öppnats upp nya metoder för effektivare övervakning utan även nya områden där AI krävs för att övervakning överhuvudtaget ska kunna ske. Ett exempel på det är högfrekvenshandel där algoritmerna är för snabba för att en människa kan övervaka effektivt. Vi kan identifiera ett behov av bevakning av samhällskritiska funktioner som har blivit helt eller delvis automatiserade som börser, elsystem och kommunikationssystem, dels mot avsiktliga attacker (3.4), men även mot oavsiktliga effekter av automatisering där oförutsedda situationer uppstår.

Andra högst troliga områden som AI-tekniker för övervakning kommer att bli mer använt är fängelser, större tillställningar, trafikflöden, drönare för att övervaka skördar, smarta klimatsystem i byggnader etc. Det är snarare svårt att hitta områden där det inte finns en möjlig tillämpning.

3.6 Prognoser

Att förutsäga händelser innan de händer är en utmaning som aldrig kommer att lösas inom många områden. Men inom specifika domäner finns det möjlighet att göra bra och användbara prognoser om förutsättningarna kan kvantifieras. Väderprognoser, valresultat, och finansprognoser är exempel på där detta i alla fall är teoretiskt utförbart. Tekniker baserad på AI för att analysera stora mängder data har kommit i och med att datamängden har blivit för stor för människor att bearbeta. Speciellt gäller detta i realtid men även i många fall är mängden data för stor för att en stab med analytiker kan gå igenom den inom sin livstid. Här listan vi några av de områden som vi har identifierat kommit starkt de senaste åren.

3.6.1 Väderprognoser

Speciellt viktigt för många samhällsfunktioner är kvalitén på väderprognoser och här sker ständiga förbättringar med hjälp av datorer och simuleringar. National Weather Service (NOAA) ger i år (2014) ut en ny modell som ska ge en högupplöst bild hur vädret ser ut i USAs 48 fastlandsdelstater. Det utnyttjar ett nätverk av mark- och satelitbaserade sensorer och radar för att mata systemet med väderdata. Upplösningen ska då gå ifrån 13 km ned till bara 3 km. Systemet ska även kunna se skillnad på roterade och icke-roterande stormar för att ge en bättre indikation ifall en tornado håller på att bildas[11]. Neurala nätverk är exempel på en AI metod som länge har används inom väderprognoser. De är populära eftersom de kan göra generaliseringar (vilket innebär att även om de inte har sett ett specifikt tillstånd kan de dra slutsatser grundat på tidigare erfarenheter) och eftersom de kan lösa icke-linjära prognoser [Taylor 2002]. Om det är AI-metoder eller traditionella metoder som ”drar det längsta strået” i slutändan är svårt att säga men AI-metoder har en potential att bli dominerande med tanke på genombrotten som kommit inom andra områden. **TRL: 6-9**

3.6.2 Samhällsprognoser

Under den arabiska våren upplevde många att sociala medier för första gången spelade en avgörande roll i samhällsomvälvande händelser. Idag är bilden av ”Twitter-revolutionen” mer nyanserad: De digitala nätverken möjliggör både för människor att utbyta information och avslöja missförhållanden nya möjligheter till censur och desinformation [Eriksson et al. 2013]. Icke desto mindre kan den digitala information som finns på sociala medier vara högst relevant i en situation som exempelvis insatser som att få ut medborgare eller skicka fredsbevarande trupper till ett kollapsande land. Genom analys av medier, sociala medier och data-flöden på internet kan man potentiellt förutse vad som kommer att hända i ett land innan det faktiskt händer. Det pågår för närvarande intensiv forskning och utveckling inom detta område. **TRL: 1-3** (På avgränsade områden med större kommersiella möjligheter, exempelvis prediktion av intäkter från filmpremiärer eller topplisteplaceringar för nya låtar, är mognaden högre [Goel et al. 2010].)

3.6.3 Pandemi- och epidemiprognoser

Ett av de större hoten mot civilisationen är utbrott av pandemier eller epidemier. Även om sannolikheten inte är stor så kan det ha katastrofala följder. Därför

11 http://research.noaa.gov/News/NewsArchive/LatestNews/TabId/684/ArtMID/1768/ArticleID/10458/NOAA%E2%80%99s-Upcoming-Weather-Forecast-Model-Zeros-In-Earlier-on-Severe-Weather.aspx

satsas det mycket pengar på att kunna förutse och ge en tidig varning och på så sätt reducera skadan av, eller förhindra dessa händelser. Ett välkänt och effektivt sätt är att analysera sökresultat på sökmotorer eller diskussioner på sociala medier. Man har sett att det finns en koppling i ökning av sökord som ”Influensa symptom” och ”Influensa behandling” med ökning av patienter som söker vård veckor senare. Som exempel så använde 40 procent av alla brittiska hälsovårdsorganisationer Twitter i smittskyddssyften år 2010 [Morris 2011]. **TRL: 4-6**

3.6.4 Sammanfattning prognoser

Vi har bara visat ett fåtal av de möjliga områden där det faktiskt är möjligt att ge indikation om, och i bästa fall förutse, vad som kommer att hända med hjälp av intelligent analys av data. Prognoser ger en möjlighet att förbereda och handla för att minimera skadan av annars oförutsägliga händelser. Eftersom de flesta bygger på algoritmer som kräver datorkraft kommer prognoserna bli bättre i takt med att beräkningskraften ökar. Det innebär att även om, mot förmodan, algoritmerna inte blir bättre med tiden så kommer vi i vilket fall ändå se en utveckling mot bättre prognoser i allmänhet.

3.7 Expertsystem

Beslutstöd med hjälp av data har vuxit explosionsartat de senaste åren. Det man vill uppnå är expertsystem som kan till exempel ge råd, larma eller helt enkel bara svara på frågor snabbt och korrekt. Ett välkänt exempel på detta är IBMs dator *Watson* som slog två stormästare i Jeopardy genom att ha tillgång till en enorm databas och avancerade algoritmer. Man visade med detta att man kan få en maskin att tolka och svara på relativ abstrakta frågeställningar som ”Proverbially, it's ’where the heart is’”[12] och ”It's a poor workman who blames these”[13]. Tanken är när tekniken är utvecklad så istället för att behöva söka väderprognosen för sitt närområde så kan frågan bli ”Behöver jag ett paraply idag?” så kan maskinen förstå vad personen menar och svara därefter. Ett expertsystem försöker alltså efterlikna människans förmåga att tolka indata och använda tidigare erfarenheter och tillgänglig information för att ge ett svar. Problemet ligger ofta inte i tillgången till data, ofta finns det en rik tillgång, utan svårigheten ligger i att konstruera algoritmerna så att de kan göra om data till meningsfull information. **TRL: 1-6**

[12] Rätt svar: Home
[13] Rätt svar: Tools

3.7.1 Kliniska expertsystem

Expertsystem inom klinisk tillämpning är system som hjälper till att ta beslut om behandling, diagnos, medicinering, prioritering, osv. [Garg 2005]. Ett system som kan användas på patienter med Alzheimers beskrivs här [Corchado 2008]. Algoritmer som kan diagnosticera patienten som ett komplement till läkarens utlåtande kan förbättra patientsäkerheten väsentligt. Kliniska expertsystem har funnits länge men fördelen med AI-system är att istället för att bygga på enbart statistik och probabilistiska metoder så kan de bygga på modeller av sjukdomstillstånden och patienterna, och därmed ge en bättre diagnos. Precis som i många andra områden handlar det än så länge mycket om att assistera och ge råd till läkaren men det är inte teoretiskt omöjligt att expertsystemen kan ta över många uppgifter ifrån personalen. Den potentiella mängd information ett system kan behandla kan vida överstiga en människas och därmed ge en mer precis behandling. De kan även ha ett bredare spektra av domänexpertis, dygnet-runt bevakning och vara mer kostnadseffektiva än människor. Kostnadseffektiviteten kommer främst av att de inte kräver lön och av att mjukvaran kan kopieras så utbildningstiden blir obefintlig när väl första versionen av programmet blir tillgängligt. Nackdelarna är idag avsaknande av kreativitet och människans ”sunda förnuft” men morgondagens kliniska expertsystem kan mycket väl komma att utrustas med ett mer mänskligt ”intellekt”. Eftersom kliniska expertsystem redan används idag så är det inte frågan om när de kommer utan mer hur bra de kan bli. Om man inte kan visa att människor besitter någon diagnosticerande förmåga som inte en maskin kan ha är det bara en tidsfråga innan diagnoserna har passerat mänsklig precision. **TRL: 1-6**

3.7.2 Personlig assistent/rådgivare

De flesta har numera kommit i kontakt med, eller hört om, datorstyrda personliga assistenter som Apples SIRI eller någon av de motsvarigheter som finns för Android-telefonerna. De är assistenter som kan svara på enkla frågor om vädret eller avståndet till närmaste korvkiosk, men även utföra enkla uppgifter som att sätta larm, skicka email eller spela musik. Även om det inte är komplexa uppgifter så ju mer utvecklingen går framåt desto mer användbara funktioner kommer att finnas. Många minns ju deras förgångare *Microsofts assistent Clippy,* ett pratande gem som var mer irriterande än hjälpsamt. För att verkligen fylla en funktion som personlig assistent i dess rätta bemärkelse behöver dock många element tillkomma som abstrakt förståelse, resonerande och logiskt tänkande. Det finns inga indikationer att dessa rådgivare ska kunna göra komplexa uppgifter inom de närmsta åren eftersom ingen forskning pekar på att vi står inför ett sådant genombrott för artificiella personliga assistenter. **TRL: 1-3**

3.7.3 Sammanfattning expertsystem

När det gäller mer abstrakta beslut/råd/användning så ser vi att dagens teknik fortfarande är begränsad. Även om det inte egentligen finns något som talar emot, så ser vi inget som tyder på att det kommer bli ett stort genombrott för abstrakt förståelse inom de närmsta åren. Dock ser vi att även utan högnivåförståelse kommer många expertsystem att få en viktig roll inom många domäner. Förutom nämnda områden så till exempel jordbruk, utbildning, militär och finans för att nämna några få.

3.8 Big data och informationshantering

Det finns ingen enhetlig definition på ”Big data” men oftast syftar man på analys av enorma mängder data. Svårigheten att analysera beror då på data-mängden, data-hastigheten, data-heterogeniteten osv. och är långt mycket mer komplext än vad en människa kan analysera utan hjälpmedel. AI har alltså ett uppenbart användningsområde här gällande filtrering, analys, sortering och annat för att presentera information för en analytiker, eller helt enkel agera autonomt på egen hand. Exempel där det redan används är låt-rekommendationerna i Spotify där alla användares musiksmak analyseras för att ge en enskild rekommendation grundad på vad liknande användare lyssnar på. Ett företag som har specialiserat sig på att ge analytiker ett verktyg att söka terrorister, människosmugglare mm. är till exempel Palantir. Deras koncept bygger på interaktion mellan dator och människa för att få en synergieffekt i samarbetet. Grundidén är att använda AI för att göra en förstärkning av den mänskliga analysförmågan (s.k. *Intelligence augementation*) genom att tillgodo se analytikern med relevant data ifrån en stor mängd data och datakällor. Deras största kunder är USAs olika statliga institut och som exempel användes Palantirs mjukvara för att avslöja det så kallade ”GhostNet” (ett kinesiskt spionprojekt som spionerade bland annat på ambassader och utrikesministar).

Så kallade ensamvargs-terrorister använder sig numera ofta av internet för att sprida sina åsikter och eventuella manifest. Till exempel var Stockholmsbombaren Taimour Abdulwahab al-Abdaly aktiv på internet där han kanaliserade sina åsikter långt innan dådet. Dessa personer ger ofta från sig flera så kallade *svaga signaler*, som var för sig inte betyder så mycket, men om man lägger ihop alla indikatorer kan man se ett mönster. Till exempel så postade Anders Behring Breivik en gravt islamofob film på Youtube timmar innan sitt terrordåd. Tillsammans med publiceringen av sitt våldsglorifierande manifest och registrerade uppgifter att han hade köpt kemikalier och stora mängder gödslingsmedel, finns det skäl att tro att om någon hade alla dessa uppgifter framför sig hade reagerat. En erfaren analytiker kan mycket väl identifiera dessa personer som potentiella våldsverkare men resurserna räcker inte till på långa

vägar i ett land som har miljontals invånare, speciellt när tiden är knapp. Även många skolskjutningar och terroristaktioner inleds med att gärningsmannen/gruppen uppger sina intentioner på internet timmar innan själva dådet. Ett intelligent system som skannar av publika uppgifter och forum kan mycket väl fånga upp dessa samband. En mer anlyserande studie om hur man kan hantera svaga signaler finns här [Brynielsson 2012]. **TRL: 1-6**

Analys av stora mängder data är inte bara användbart i säkerhetssyfte. Företag som inriktar sig på text-mining av till exempel Twitter, Facebook och sociala medier har exploderat i mängd. Bara i USA förväntas över en miljon personer med kompetens att hantera data att behöva anställas i framtiden[14]. Forskning, reklam, sjukvård, samhällsanalyser är några av de andra områden som Big-data har, och kommer att ha stor påverkan på i framtiden. För den här delen av informationshantering bedömer vi **TRL: 5-9**

3.9 Språk och kommunikation

3.9.1 Sökfunktioner

Den kanske mest utbredda och använda AI-liknande systemen är sökfunktioner som Google och Bing. Googles mest kända algoritm PageRank [Page 1999] baseras på länkar som människor har gjort mellan sidor för att ranka relevans. Ju fler länkar desto viktigare är sidan och mer sannolik att hamna långt upp på resultatlistan. Tillsammans med andra kriterier, till exempel nyckelord, får användaren en filtrerad och destillerad version av möjliga sidor i form av en lista. Med andra ord är skillnaden mellan sökfunktioner och personliga assistenter att de senare bara ger tillbaka ett svar medan här gör användaren den sista avgörande beslutet på vad som verkligen är relevant baserad på en lista (söklistan du ser efter en internetsökning). Inom det här området sker en ständig förbättring av filtrering- och sorteringsalgoritmerna men det finns en låg kontextuell förståelse än vilket gör att de bara kan nå en viss gräns i användbarhet. Som nämnts tidigare är det svårt att göra en uppskattning av när algoritmerna blir bättre på att förstå sådan kontext som är viktig för att göra bra sökningar. Även när det gäller bild och videosökningar görs dessa idag fortfarande främst på vilka taggar mediet har, inte själva innehållet i sig, så algoritmerna förlitar sig på att människor har satt rätt beskrivning på filen. **TRL: 7-9**

[14] http://rkbusinessschool.com/uploads/2_The_New_York_Times_on_The_Age_of_Big_Data.pdf

3.9.2 Taligenkänning

Taligenkänning är ett populärt område av många anledningar, dels för de ofta används av röststyrda personliga assistenterna nämnda tidigare, dels för att tal i råformat tar stort utrymme som ljudfiler. Kan man omvandla tal till text kan mycket större mänger information sparas. Lagrad text är dessutom lättare att analysera, söka och kategorisera än motsvarande ljudfiler. Dagens teknik är väl utvecklad där mjukvaran gör relativt få fel, som exempel Nuance's program "Dragon". I princip är alltså det här ett löst problem i många språk, förutom när ljudkvaliten är dålig eller när man har begränsad datorkapacitet som i mobiltelefoner. Om bara ett fåtal år kan man anta att dessa problem är lösta även de. **TRL: 8-9**

3.9.3 Språköversättning

De flesta som har använt verktyg för språköversättning som Google translate förstår att det är en bit kvar till fulländning. Översättning mellan liknande språk fungerar relativt bra jämnfört med översättning mellan till exempel tyska och mandarin. Även om det kan vara bra nog i många sammanhang kommer problemen med att översätta ironi, liknelser, kontext m.m. att inte lösas förrän text- och talförståelse har blivit löst (se nästa sektion) vilket gör det svårt att säga hur tekniken ser ut i framtiden. **TRL: 4-6**

3.9.4 Text- och talförståelse

Natural language processing (NLP) eller talförståelse är ett populärt område inom AI och även om stora framsteg har gjorts är det långt kvar till mänsklig nivå på tolkning. I de applikationer som finns måste texten vara förenklad och helst grammatiskt felfri för att ens förstå de mest enkla kommandon. När väl tekniken är utvecklad kan den bland annat användas till automatiserad sammanfattning, översättning, analys, svarextraktion, informationshämtning och sökning. Eftersom NLP är starkt kopplat till att göra maskiner lika intelligenta som människor behöver många komponenter för stark AI komma till innan problemet helt kan lösas. Innan dess får vi nöja oss med att kommunicera med en dator via strikta regelverk. **TRL: 1-3**

3.9.5 Sammanfattning språk och kommunikation

Ett av de stora hindren för språk och kommunikation som vi ser det är förståelsen för *kontext.* Alltså vad den högre innebörden är av en mening är förutom nyckelorden. Att använda nyckelord räcker bara en bit men för att tolka språk på samma sätt som människor krävs mycket mer forskning. Vi tror att det blir en gradvis förbättring med tiden men det är omöjligt att säga när vissa genombrott uppnås. Vi ser dock redan idag att även utan den kontextuella förståelsen

kommer ändå användbarheten vara mycket utbredd. Det är svårt att tänka sig att använda internet utan sökfunktioner och tekniken kommer bara att bli bättre framöver.

4 Bevakningsbehov

I detta kapitel försöker vi lista intressanta trender som har ett bevakningsbehov. Dels gäller det tekniker i sig men även hur stater och företag satsar på olika AI-områden. Det här är självklart bara ett urval och kommer med all säkerhet ändras när nya rön inom AI kommer.

4.1.1 Tekniktrender att följa

Det är få av de nämnda områdena inom AI som inte har potential att bli viktiga i framtiden. Vi försöker här göra en sammanställning av de mest intressanta teknikerna ur ett användningsperspektiv. De beskrivs alla mer detaljerat i kapitel 3.

- Autonoma farkoster(3.1): Det börjar nu bli tekniskt möjligt att ha självstyrande bilar, stridsfordon, tåg och flygplan. Det kommer sannolikt att påverka samhäller i stort, bara i USA finns det 5 miljoner som är chaufförer till yrket. Lägger man till antalet timmar som går åt för att köra till och från jobbet för alla privatpersoner blir det en enorm samhällskostnad. Autonoma flygmaskiner kommer även att blir mer avancerade och använda, både civilt och militärt. Bara i USAs armé finns tusentals i drift, där många modeller även är vapenbestyckade för strid.
- Ansiktsigenkänning (3.2): Tekniken börjar nu närma sig samma nivå som vi människor kan känna igen ansikten vilket skulle innebär ett stort genombrott. Denna teknik är bra nog att användas både civilt och militärt och det finns många potentiella användningsområden så det är en viktig teknik att följa.
- Händelseigenkänning (3.2): När det blir möjligt för en dator att förstå vad som händer via en kamera eller videosekvens kommer vi att få se intressanta användningsområden som t.ex. brottsbekämpning via CCTV kameror där miljontals kameror kan övervaka stora områden dygnet runt.
- Operationsrobotar (3.3): Möjligheten att låta robotar operera skulle kunna rädda människor i områden där operationsläkare är en bristvara. Första användningsområdet skulle dock vara vid assistans och endast mindre komplexa operationer men utvecklingen kan gå snabbt efter det.
- Väderprognoser (3.6): Genom bra väderprognoser kan man bland annat öka produktionen i jordbruket och minimera kostnader för naturkatastrofer som stormar och översvämningar. Genom att öka

exaktheten i väderprognoser kan AI bidra till stora samhällsekonomiska vinster. En speciellt intressant teknik bygger på (djupa) neurala nätverk.

- Samhällsprognoser (3.6): AI kan användas för att se sociala strömmar som kan till exempel leda till effektivare fattigdomsbekämpning, förutspå epidemier, förutspå revolutioner, mm. Används redan fast i väldigt begränsad utsträckning.
- Expertsystem, kliniska (3.7): AI kan användas för diagnosticering och medicinering vilket kan ge stora samhällsvinster dels i form av minskade personalkostnader och dels i form av färre feldiagnosticeringar och mindre felmedicinering. Det krävs självklart att tekniken blir bättre/billigare/snabbare än människor på diagnosticering vilket inte är givet. Tekniken är inte där än men det finns goda möjligheter att det kommer i framtiden.
- Expertsystem, jordbruk (3.7): För att användas till övervakning och optimering av bevattning, gödningsmedel, skördetid bland annat. Liknande system kan även användas till att övervaka och styra om trafik.
- Analys av data, människa-datorinteraktion (3.8): Tillsammans med en AI kan en mänsklig analytiker få en ”hävarm” som hjälper denne att analysera mängder av data som aldrig skulle vara möjlig annars. Vi ser inte att en människas analytiska förmåga kan bli ersatt i en nära framtid utan AI blir ett hjälpmedel för att filtrera, samla och presentera enorma mängder data på ett lättolkat sätt. Speciellt inom underrättelsetjänst och analys kommer det här att bli viktigt.
- Analys av data, automatiserat (3.8): Kan användas till bättre användning av resurser i liten skala som till exempel att hålla ett rum svalare där ingen är för att spara energi, till omallokering av resurser på nationell nivå för att optimera energianvändning för miljontals människor. Området är enormt stort så mognadsgraden varierar.

4.1.2 Intressanta trender inom länder och stater

- Kina har seglat upp som en stor robotnation och var 2013 största uppköparen av industrirobotar med Japan som tvåa.
- USA och DARPA satsar väldigt brett och stort inom AI. Vi har beskrivit de olika projekten i kapitel 3 för vidare läsning.
- Den årliga tillväxten av forskare inom teknik i OECD länder mellan 1995 och 2005 var ungefär 3.5% [NSF]. Motsvarande lista för Kina var 10.8% vilket motsvarar en fördubbling var 7:e år.

- The Economist uppskattar att data-analys-industrin i Europa runt 500 miljarder SEK[15] med en årlig tillväxt på 10%. Global Industri Analysts uppskattar att vid nästa år (2015) så kommer AI-marknaden vara värd nära 300 miljarder SEK[16].

4.1.3 Intressanta trender inom industrin

- Många bolag satsar stort inom AI men Google står i en klass för sig. De har en egen AI-grupp och har i år (2014) även köpt upp en av världens största fristående AI-grupp DeepMind för $500 miljoner. Tidigare gjorde de ett uppköp av ett bolag som tillverkar termometrar för inomhusbruk. Man spekulerar i att de bland annat vill integrera AI system för att optimera elkostnader för uppvärmning i hem. Som en av världens största bolag och stor aktör inom AI är Google en trendsättare att följa.
- BI (Business Intelligens) används i analys av kunders beteenden för att till exempel anpassa reklam och rekommendationer. King, ett svenskt bolag som gör spel för Facebook, gör analys av alla spelare och anpassar spelet och betalningssystemen individuellt genom AI-metoder. Även Spotify som är ett annat svenskt bolag gör musik-rekommendationer baserat på liknande algoritmer.
- Amazon och DHL vill använda drönare för att dela ut varor för att göra utkörningen billigare och snabbare. Amazon har redan semi-automatiserad pakethantering i lagret med hjälp av robotar. Deras troliga mål är att göra hela processen automatiserat, från köp till leverans.
- Bilföretag är bara en av många industrier som gör sina löpande band mer och mer robotinriktade för att effektivisera produktionen.

[15] http://www.economist.com/node/15557443
[16] http://www.strategyr.com/Artificial_Intelligence_AI_Market_Report.asp

5 Säkerhetsaspekter

Vi överlåter mycket till läsaren att avgöra säkerhetsaspekter när det kommer till nämnda tekniker. Vi vill dock peka på några särskilt viktiga säkerhetsaspekter som kan tänkas komma att bli aktuella:

- Drönare och andra autonoma farkoster (3.1): många länder satsar stort och en autonom pilot har potentialen att bli en mycket bättre flygförare än sin mänskliga motsvarighet. FBI har även pekat ut att självkörande bilar kan användas för att begå brott.
- Ansiktsigenkänning(3.2): det är risk att tekniken till exempel används för att känna igen soldater i strid och hota deras anhöriga i hemländerna.
- Cyberhot (3.4): vi tror att det blir vanligare med attacker som grundar sig i AI-metoder. Med enkla medel kan de ställa till med stor skada och därför är det viktigt att följa den här utvecklingen.
- Kartläggning av människor via sociala medier kan göras automatiskt med AI-tekniker och kommer att bli mer och mer avancerade. Speciellt för personer med känslig information (politiker, analytiker, militär personal, mm.) kan spårbarhet innebära ett problem.

En teknik vi har inte nämnt mycket här eftersom vi anser att den inte kommer att komma inom bara ett fåtal år är stark AI (nämnt i kapitel 2). Skulle den komma kan den potentiellt vara extremt farlig så därför vill vi nämna den. Det vill säga om vi lyckas bygga en AI som har en nära-mänsklig intelligens så skulle vi kunna låta den bygga en förbättrad version av sig själv. Den i sin tur skulle kunna bygga ännu en som är än mer intelligentare och snabbare. Det här blir en iterativ process som skulle accelerera väldigt snabbt tack vare lättheten att kopiera data-kod och med datorhastigheten. Det skulle potentiellt leta till en så kallad super-intelligent AI som med lätthet kunna utmanövrera människor i kapacitet speciellt om den har tillgång till ett världsomspännande nät som till exempel internet [Bostrom 2014]. En okontrollerad super-intelligens har i undersökningar bland forskare hamnat tillsammans med kärnvapenkrig som en av de möjliga anledningarna till mänsklighetens undergång. Även om det här låter som science fiction så är det inget osannolikt scenario i en alltför avlägsen framtid. Det finns inget som talar emot att AI kan bli minst lika intelligent som en människa och då finns det egentligen inget som talar mot att något i de här banorna kan hända.

Sammanfattningsvis är det författarens åsikt att för att ha ett säkert samhälle i framtiden är det viktigt att ha stor kompetens och forskning inom AI. Den fördelen man kommer att få med AI-system i framtiden är oöverskådlig och samhälleliga förluster potentiellt stora om ett samhälle skulle halka långt efter.

6 Referenser

Ahlberg, J.; Johansson, F. Johansson, R.; Jändel, M.; Linderhed, A.; Svenson, P. & Tolt, G. (2011) **”Content-based image retrieval. An introduction to literature and applications”**, FOI-R—3395-SE, Technical report, FOI.

Albright, D.; Brannan, P.; Walrond, C. (2010) **“Did Stuxnet Take Out 1,000 Centrifuges at the Natanz Enrichment Plant?”** ISIS Reports

Axell, E., Jändel, M., Lindahl, D.,Linder, S., Rodhe, I., & Westerdahl, L. (2013) **”Robust kommunikation, cybersäkerhet och artificiell intelligens för säkerhet i RPAS”** FOI-R—3825—SE.

Biteus, J. (2007). **“Fault isolation in distributed embedded systems”** doktorsavhandling, Linköpings universitet.

Bostrom, N. (2014). **“Superintelligence: Paths, dangers, strategies.”** Oxford University Press.

Brynielsson, J., Horndahl, A., Johansson, F., Kaati, L., Martenson, C., & Svenson, P. (2012, August). **”Analysis of weak signals for detecting lone wolf terrorists.”** In Intelligence and Security Informatics Conference (EISIC), 2012 European (pp. 197-204). IEEE.

Casal, A., Hogg, T., & Cavalcanti, A. (2003). **“Nanorobots as cellular assistants in inflammatory responses.”** In IEEE BCATS Biomedical Computation at Stanford 2003 Symposium.

Ciresan, D., Meier, U., Masci, J., & Schmidhuber, J. (2011, July). **“A committee of neural networks for traffic sign classification.”** In Neural Networks (IJCNN), The 2011 International Joint Conference on (pp. 1918-1921). IEEE.

Corchado, J. M., Bajo, J., De Paz, Y., & Tapia, D. I. (2008). “**Intelligent environment for monitoring Alzheimer patients, agent technology for health care.**”, Decision Support Systems, 44(2), 382-396.

DeBenedictis, Erik P. (2005). **"Reversible logic for supercomputing"**, Proceedings of the 2nd conference on Computing frontiers. pp. 391–402

Ducatelle, F.; Di Caro, G. A.; & Gambardella, L. M. (2010). **”Principles and applications of swarm intelligence for adaptive routing in telecommunications networks.”** Swarm Intelligence, 4(3), 173-198.

Easley, D., M. López de Prado, M. O'Hara: (2011) **“The Microstructure of the ‘Flash Crash’: Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading”**, The Journal of Portfolio Management, Vol. 37, No. 2, pp. 118–128, Winter.

Eriksson, M., Franke, U., Granåsen, M. & Lindahl, D. **“Social media and ICT during the Arab Spring”** (2013) FOI-R--3702--SE.

Estlin, T., Gaines, D., Chouinard, C., Castano, R., Bornstein, B., Judd, M., Nesnas, I. & Anderson, R. (2007, April). **“Increased Mars rover autonomy using AI planning, scheduling and execution.”** In Robotics and Automation, 2007 IEEE International Conference on (pp. 4911-4918). IEEE.

Fagnant, D. J.; Kockelman, K.; (2013) **”Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations for capitalizing on self-driven vehicles”**, Transportation Research.

Flushing, E. F., Kudelski, M., Gambardella, L. M., & Di Caro, G. A. (2013, October). **”Connectivity-aware planning of search and rescue missions.”** InSafety, Security, and Rescue Robotics (SSRR), 2013 IEEE International Symposium on (pp. 1-8). IEEE.

Fleischer, M. (2005). **”Foundations of swarm intelligence: From principles to practice.”** arXiv preprint nlin/0502003.

Forsell, L. (2013) **”Obemannade luftfarkoster för territorialövervakning, tekniktrend 2030”** FOI-R—3662—SE.

Franke, U., & Brynielsson, J. (2014) **“Cyber situational awareness–A systematic review of the literature”** Computers & Security, 46, 18-31.

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO. (2013, September 10). **”Scientists aim to develop intelligent and wearable body exoskeleton.”**, ScienceDaily.

Garg, A. X., Adhikari, N. K., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., ... & Haynes, R. B. (2005). **”Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review.”** Jama, 293(10), 1223-1238.

Goel, S, Hofman, J M, Lahaie, S, Pennock, D M, Watts, D J, **“Predicting consumer behavior with Web search”**, PNAS, vol. 107 no. 41 17486-17490, October 12, 2010.

Hartmann, K., & Steup, C. (2013) **“The vulnerability of UAVs to cyber-attacks - An approach to the risk assessment.”** In Cyber Conflict (CyCon), 2013 5th International Conference on (pp. 1-23). IEEE.

Huo, Y., & Li, Z. (2009). **”Mechanism design and simulation about the exoskeleton intelligence system.**” In Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on (Vol. 2, pp. 551-556). IEEE.

Johansson, K. & Mårtenson, C. (2006**) “Mixed-initiative-interaktion för militära beslutsstödssystem”** FOI MEMO 1864.

Kerber, R. E., Becker, L. B., Bourland, J. D., Cummins, R. O., Hallstrom, A. P., Michos, M. B., ... & Zuckerman, B. D. (1997). **"Automatic External Defibrillators for Public Access Defibrillation: Recommendations for Specifying and Reporting Arrhythmia Analysis Algorithm Performance, Incorporating New Waveforms, and Enhancing Safety"** Subcommittee on AED Safety and Efficacy. Circulation, 95(6), 1677-1682.

Libicki, M. C **"Cyberdeterrence and Cyberwar"** 2009 Rand.

Lin, S. W., Ying, K. C., Lee, C. Y., & Lee, Z. J. (2012). **"An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection."** Applied Soft Computing, 12(10), 3285-3290.

Lua, R., & Yow, K. C. (2011). **"Mitigating DDOS attacks with transparent and intelligent fast-flux swarm network."** *Network, IEEE*, *25*(4), 28-33

Lyon, D. (2006) **"Theorizing surveillance, The panopticon and beyond"**

Madden, C., Cheng, E. D., & Piccardi, M. (2007). **"Tracking people across disjoint camera views by an illumination-tolerant appearance representation"**.Machine Vision and Applications, 18(3-4), 233-247.

Mayer, H., Gomez F., Wierstra D., Nagy I., Knoll A., and Schmidhuber J. (2006). **"A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks."** Proceedings of the International Conference on Intelligent Robotics and Systems (IROS-06, Beijing).

Minsky, M. & Papert, S. (1969) **"Perceptrons: an introduction to computational geometry"** MIT press.

Morris, K. (2011) **"Tweet, post, share - a new school of health communication"** The Lancet Infectious Diseases, vol. 11 (7), ss. 500-501.

Ni, B., Wang, G., & Moulin, P. (2013). **"RGDB-HuDaAct: A color-depth video database for human daily activity recognition."** In Consumer Depth Cameras for Computer Vision (pp. 193-208). Springer London.

NSF (US National Science Foundation). **"Science and Engineering Indicators"** 2010, Chapter 3. Science and Engineering Labor Force.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). **"The PageRank citation ranking: Bringing order to the web."**

Rodhe, I., Westring, E. & Karlzén H. (2014) **"Self-healing and self-protecting software. Scanning the research frontier."** FOI-R--3836--SE

Ryoo, M. S., & Aggarwal, J. K. (2011). **"Stochastic representation and recognition of high-level group activities."** International journal of computer Vision, 93(2), 183-200.

Savage, S. J., Allberg, H., Larsson, B., Nilsson, T., Peolsson, M., Rydell, J., Stenborg, K-G., Emilsson, E. (2012) **”Ett proaktivt trygghetssystem”**, FOI-R--3515--SE

Silberglitt, R., Anton, P. S., Howell, D. R., Wong, A., & Gassman, N. (2002). **“The global technology revolution 2020, in-depth analyses: Bio/nano/materials/information trends, drivers, barriers, and social implications.”** Rand Corporation.

Sommestad, T. & Franke, U. (2014) **“A test of intrusion alert filtering based on network information”** Security and Communication Networks, under tryckning.

Taylor, J. W., & Buizza, R. (2002). **“Neural network load forecasting with weather ensemble predictions.”** Power Systems, IEEE Transactions on, 17(3), 626-632.

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). **”Deep-Face: Closing the Gap to Human-Level Performance in Face Verification.”** In IEEE CVPR.

Wu, S. X. & Banzhaf, W. (2010) **“The use of computational intelligence in intrusion detection systems: A review.”** Applied Soft Computing 10.1, pp. 1-35.

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). **“Face recognition: A literature survey.”** Acm Computing Surveys (CSUR), 35(4), 399-458.

Att filtrera, sortera, förstå, manipulera, extrahera i det dataflöde vi omges av har blivit informationsålderns stora utmaningar. Från att fysiska tillgångar historiskt sett har lett till en nations välmående och rikedom, har ett skifte skett till att det är de mer abstrakta tillgångarna som påverkar framgång mest. Det kommer mer och mer bli förmågan att använda och förstå information som blir avgörande för utveckling. För att göra detta måste vi bygga system som kan hjälpa oss människor med de delar vi inte förmår, kan, är för långsamma för, eller helt enkelt tycker är för tråkiga att hålla på med.

Det man bland annat försöker uppnå inom forskningsområdet artificiell intelligens (AI) är att förmå en maskin att lära sig av situationer som den upplever och med dessa erfarenheter lösa uppgifter. Redan idag används AI inom vitt skilda domäner såsom bilkörning, spel, ansiktsigenkänning, övervakning, diagnosticering, epidemiprognoser, expertsystem och sökfunktioner. Användningsområden blir fler och fler allt eftersom tiden och utvecklingen går framåt. Vi kan vänta oss stora framsteg framöver eftersom området till stor del är outforskat. Området artificiell intelligens har bara funnits i ungefär 60 år vilket är väldigt kort jämfört med många andra forskningsområden. Även om man fortfarande är långt ifrån en ”människolik” artificiell intelligens har man kommit långt inom många specifika områden. Det är dessa områden den här rapporten handlar om för att ge läsaren en inblick i den aktuella forskningsfronten och möjliga framtida tillämpningar. Denna tekniska prognos sammanfattar kortfattat kunskapsläget, lyfter fram aktuella tekniker och analyserar trender som kan leda till nya tillämpningar.

FOI

ISSN1650-1942

www.foi.se