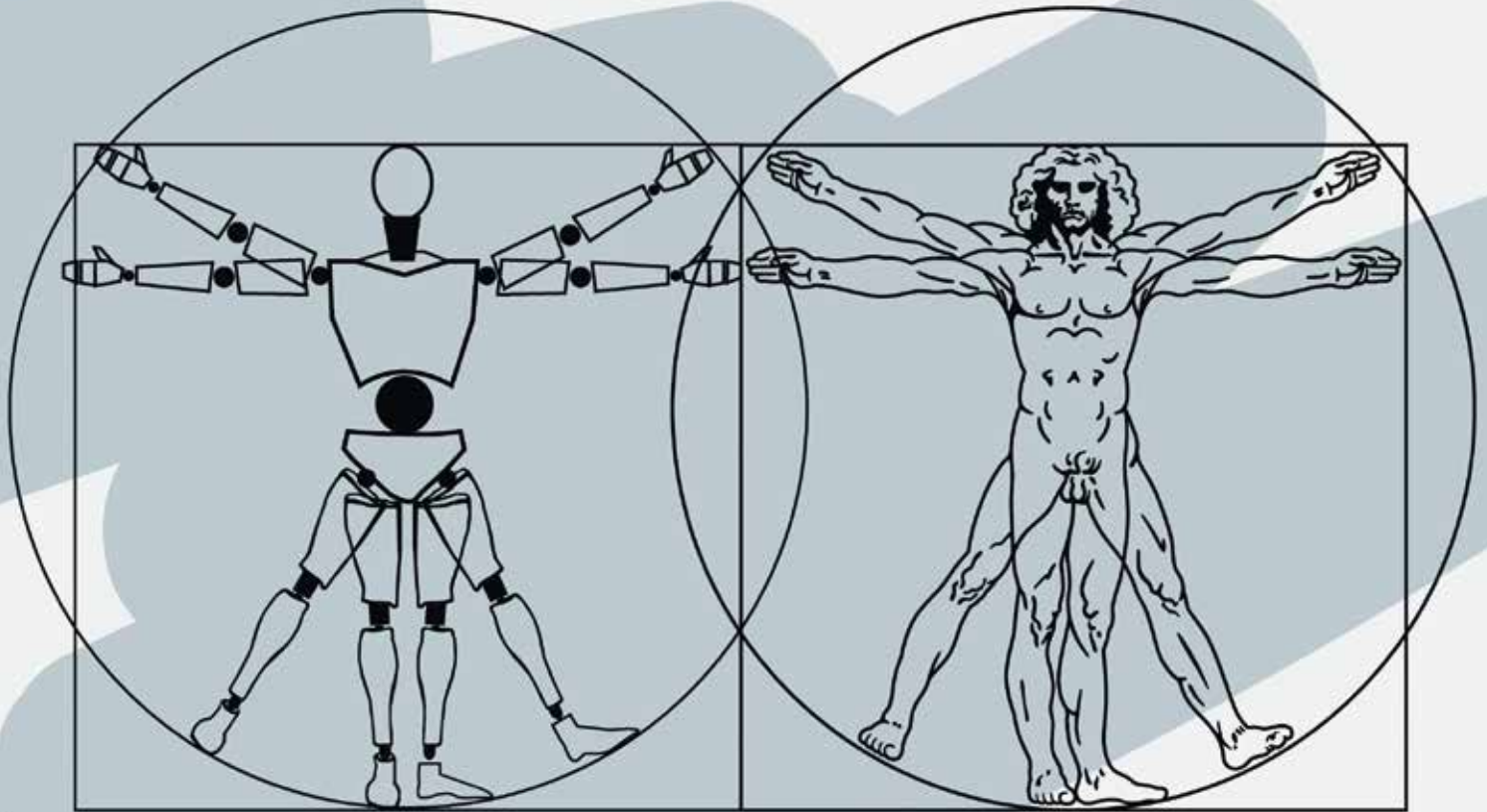




Livet med
AI



Redaktörer: Eva Regårdh och Sofie Pehrsson

Grafisk produktion: Förnuft & Känsla Marknadskommunikation AB

Illustration: Hedda Gumpert

Tryck: Trydells Tryckeri, 2018, första tryckningen

ISBN 91-89206-71-1

Sammanfattning.....	5
Vad är AI? <i>Magnus Boman</i>	7
AI-utvecklingens yttersta risker, <i>Olle Häggström</i>	10
Fyra förändrade drivkrafter för AI i vården, <i>Paulina Modlitba</i>	16
Vart kommer AI att köra oss? <i>Anna Pernestål Brenden</i>	22
Nya affärsmodeller baserat på artificiell intelligens, <i>Erik Kruse och Patrik Regårdh</i>	28
Fintech och robotrådgivning, <i>Ann Grevelius</i>	34
Hur väl står sig Sverige? <i>Mats Nordlund</i>	42
Livet med AI i skolan, <i>Mats Nordlund</i>	48
Reglering av Artificiell Intelligens, <i>Peter Wahlgren</i>	52
Om ny teknik för framtiden, <i>Nina Wormbs</i>	58
Pionjärerna som drömde om AI, <i>Christer Sturmark</i>	64
SSF:s Novusundersökning	74
Referenser.....	80

Varje kapitel kan läsas helt fristående.

Sammanfattning



Sjukvård, företag, ja vår vardag och våra arbeten påverkas alltmer av digitalisering och AI, Artificiell Intelligens. Lagstiftning, arbetsmarknadsåtgärder och de demokratiska processerna behöver hänga med. Därför är det angeläget att engagera sig i teknik- och samhällsutvecklingen. Det finns en tendens att betrakta digitalisering som en effektivisering av befintliga processer, snarare än en formidabel förändringskraft. Finansmannen Jan Stenbeck sa insiktsfullt att "politik slår måhända pengar, men teknik slår politik". Det står sig även idag. I denna rapport låter Stiftelsen för strategisk forskning experter och tänkare från vitt skilda områden ge sin syn på hur AI kan komma att omskapa samhället. Läs själv och begrunda, vad tror du?

Att ökad digitalisering, artificiell intelligens och datamining, kommer att förändra samhället, liksom jordbrukets mekanisering, eller elens införande, är de flesta överens om. Men det råder däremot delade meningar om på vilket sätt, och hur fort. Hur kommer sjukvården att påverkas av att patientens DNA finns till hands, att datorerna ställer diagnos, och kan ge personligt utformade behandlingsråd? Hur snabbt får vi självkörande fordon och vad innebär det för transporter, för chaufförens jobb, för stadernas trafiksystem? Vad händer med bankerna och det finansiella systemet när robohandel och robotrådgivning införs på bred front? Hur påverkas företagen när tillverkning successivt går mot att bli alltmer automatiserad och värdemodellerna förändras mot att snabbt kunna skala, samtidigt som man måste kunna möta den enskilda konsumentens behov för att vara konkurrenskraftig? Hur påverkas utbildningsväsendet när kurser on-line blir allt tillgängligare och pedagogiska? Hur kommer våra virtuella världar se ut och i vilken utsträckning flyttar vi in i dem?

Datorer generellt är bra på att lösa problem, men inte hittills på att ställa de relevanta frågorna och det är just förmågan att formulera problemställningar som driver forskning, innovation och utveckling. AI ger nya resurser att svara på frågorna. Men det kommer att dröja innan AI-systemen kan agera och driva utveckling självständigt. Det är gott och väl att datorer spelar schack eller Go bättre än oss människor, men när har vi ett avancerat spel som är konstruerat av AI?

AI kan, till skillnad från vanliga datorer, få en roll i att kodifiera icke kodifierbar kunskap, och att utöka vår tanke-mässiga sfär. Många mänskliga förmågor kan vi inte beskriva; vi kan inte fullt ut förklara hur vi löser problem eller fattar beslut, vi vet helt enkelt mer än vad som låter sig kodifieras i ett program. På motsvarande sätt bygger ett självlärande system också upp kunskap som inte låter sig förmedlas i alla detaljer. AI-systemen löser frågor på ett annat sätt än oss, och vi har inte förutsättningar att förstå det fullt ut, lika lite som vi kan förstå exakt vad talang sitter i, även om vi känner igen det när vi ser det. Kanske måste vi acceptera AI som kunskap bortom det kodifierade.

Lars Hultman, vd SSF

Eva Regårdh, kommunikationschef SSF

Vad är AI?

Magnus Boman är professor i intelligenta programvarusystem på KTH och forskningsexpert på RISE.

Artificiell Intelligens (AI) är ett omstritt koncept, som efter drygt 60 år av forskning fortfarande inte har någon precis mening. Vetenskapliga discipliner, som humaniora, filosofi och neurovetenskap har också egna tolkningar av själva begreppet intelligens. De flesta AI-forskare ser dock inte detta som något problem, tvärtom, de låter AI vara ett paraply som de står under när det passar. Och forskarna är förstås ändå överens om en hel massa saker under paraplyet AI, så låt oss spänna upp det och börja där.

All AI handlar om någon sorts lärande system. Systemet kan få hjälp utifrån, men det förväntas åtminstone delvis vara självlärande. Det betyder att systemet har en viss autonomi: om det har baskunskaper i någon domän så kan det klara sig själv rätt länge innan det behöver mera hjälp. Ett exempel kan vara en svetsande robot i form av en orm som kan krypa runt i stora fartygsskrov och svetsa där människor inte bör eller kan göra det. Baskunskaperna kan då byggas kring en ritning av skrovet, eller kring vad roboten kan se när den utforskar skrovet, till en början utan att svetsa.

Det finns också en etisk dimension närhelst AI används i praktiken. Den etiska faktorn har oftast både positiva och negativa värden. I exemplet kan en positiv faktor vara att människor slipper riskera sina liv genom att svetsa på platser där gasfickor kan bildas, vilket i sin tur kan leda till en explosion och skada svetsaren. En negativ etisk faktor kan vara att mänskliga svetsare förlorar sina jobb när robotar börjar användas. Notera att när vi här talar om robotar så måste det inte handla om hårdvara, utan det gäller mjukvarurobotar också, ofta kallade agenter.

Forskarna skiljer också enstaka robotar från system av robotar. Utformningen av system där robotar eller mjukvarusystem med AI kan kommunicera med varandra skiljer sig nämligen väsentligt från de där man bara behöver hantera en enda robot, som möjligen kommunicerar med en människa då och då. I det enkla fallet kan man fortfarande tala om självmedvetenhet: på något vis kan en robot skilja sig själv från objekt i miljön som den kan förnimma (se, höra, känna, ...) och kanske urskilja de där särskilt viktiga objekten vi kallar människor. I det svåra fallet måste roboten känna igen alla objekt den kan kommunicera med, bland annat andra robotar. Den kan samarbeta med, eller tävla mot, andra robotar, givet att den kan känna igen dem och adressera dem. Det sistnämnda betyder oftast att skicka meddelanden via något förutbestämt robotprotokoll. I exemplet med svetsroboten kan olika robotar arbeta parallellt med olika delar av skrovet, men det är inte säkert att det är lämpligt att låta dem svetsa samma fog samtidigt. Då är det bättre att roboten, efter att ha fått ett meddelande från en annan robot, förstår att "den här fogen jobbas det redan på, jag flyttar mig till en annan fog".

Den sortens AI vi beskrivit kallas för *svag AI*: AI utformad för att utföra en uppgift, inom en domän. Uppgiften i exemplet är att svetsa, inom domänen fartygskonstruktion. Det är ofta mycket svårt att få svag AI att fungera optimalt, till och med när det bara är en enda robot vi talar om. Det beror på att väldigt mycket baskunskaper ska in, som matematiska modeller för rörelser och för alla effekter som robotens olika handlingar kan få på miljön. Allt det här ska sedan helst förfinas över tid och uppdateras i robot och omgivande system. Till och med när man tror att systemet lärt sig svetsa nära nog optimalt så kan något oväntat inträffa som programmerarna inte tänkt på. Att de skulle tänkt på allt är både ekonomiskt sett orimligt och teoretiskt sett omöjligt i många domäner. På Internet finns gott om filmer på människor som knuffar en robot och sen skrattar åt hur svårt den har att hitta tillbaka till vad den nu gjorde. Som motvikt finns det också filmer på robotar som kan plocka isär sig själva, reparera en trasig del, för att sedan sätta tillbaka delen och fortsätta arbeta som förut.

Skulle en robot som lärt sig svetsa perfekt inuti fartygskrov kunna lära sig svetsa perfekt på en husgrund? Skulle modellerna roboten tillgodogjort sig kunna användas för att låta roboten göra något annat inuti skrovet än att svetsa, som att inspektera fogar? Två frågor som tillhör samma kategori, nämligen frågor om generaliserbarhet. Den första handlar om att byta miljö (och därmed domän), den andra om att byta uppgift. Det finns flera typer av frågor i samma kategori, men relativt stor del av forskningen om AI har under det senaste decenniet handlat om graden av överförbarhet från en domän till en annan, eller mellan olika uppgifter. Det vore ju nämligen så skönt om man kunde träna en robot på en uppgift, för vilken man har massor av data redan, för att sedan låta den utföra det riktiga arbetet på en annan uppgift. Då kunde

träningen bli jämförbar, billigare och mer kontrollerad, och vara en viktig ingrediens för att driva utvecklingen vidare. Det finns idag datamängder av hög kvalitet tillgängliga för det här ändamålet.

Stark AI kräver vilja

Om nu utvecklingen av AI med sina tekniker, metoder och algoritmer fortskrider i snabb takt så kan man tänka sig en framtid där överförbar inlärning är möjlig, kanske till och med optimerad och billig. I så fall har man uppfyllt ett av kraven på stark AI. Det finns dock fler krav, som är ännu svårare att uppnå, vilket är förklaringen till att det inte finns någon *stark AI* idag, eller under någon näraliggande framtid. För det krävs också att systemet har förmågan att känna och uppleva saker, i subjektiv mening, alltså som en individ. Inom hågläran (Philosophy of Mind) så kallas dessa 'saker' för *qualia* och individen kallas *sentient*. Något som kan förnimma *qualia* anses allmänt ha rättigheter. Tills relativt nyligen ansågs inte fiskar kunna känna smärta och när det nu är motbevisat så har nya lagkrav införts i många länder som gör fiskerinäringen ansvarig också för fiskars rättigheter. Om det känns som ett lite löjligt exempel så tänk på att det är många som bekymrar sig för robotars framtida rättigheter. Dessa personer är oftast inte AI-forskare eller ingenjörer, men kan däremot vara framgångsrika forskare eller företagare inom andra områden, såsom fysikern Stephen Hawking eller entreprenören Elon Musk. De får därför ganska mycket gehör för sina åsikter, vilka dock förutsätter att stark AI blir en realitet i framtiden. Kravlistan innehåller ännu mer för att det ska uppnås: AI-systemet ska ha ett medvetande och en håg. Det ska alltså kunna känna lust och besvikelse, kunna förstå något om vad det självt är. Idag finns ingen forskargrupp som gått ut med att de försöker realisera något slags generell intel-

ligens med hjälp av stark AI. Men att ställa frågan hur vi ska hantera stark AI i samhället är för den skull inte orimligt och alla, även AI-forskare, har säkerligen mycket att lära från den debatt som just nu pågår.

Blir AI smartare än människor?

Dagens AI-program är i nuläget fruktade av många och det av flera anledningar: de stjälar arbetstillfällen, de är ogenomskinliga på så vis att få förstår hur de fungerar, de kan luras eller vara spionprogram. En av de mest omskrivna farhågorna är att "AI tar över planeten" vid en tidpunkt som benämnts singulariteten: den tidpunkt då "AI blir smartare än människor" och ser till att göra sig av med saker och kanske till och med människor som "AI inte behöver." En förutsättning för singulariteten är att man kan tala om "en AI", vilket kräver hopkopplade AI-program, AI-system och system av system. Där är vi inte ännu. Det teknikområde som kommit längst i resan emot "en AI" är virtuella världar. Det handlar om spelvärldar, men också finansvärlden med våra bankappar och pengahandtering, samt numera också vår sjukvård: hälsoappar, digitala journalsystem och integrering av olika delar av hälsovården (friskvård såväl som sjukvård). Tanken på att alla virtuella världar en dag kan bli en kallas *metaversum*. Eftersom AI-program kan exekveras för att skapa virtuella världar och även exekveras inuti virtuella världar, så skulle metaversum kunna vara ett sätt för "en AI" att realiseras i framtiden.

Vem ska då driva på utvecklingen av metaversum, AI-bot-tar, samt fortsätta styra och vara innovativ inom mer vanliga former av maskininlärning? Som marknaden för programmerare och dator- och datatekniker idag ser ut så är tillgången på kompetent personal den allra största utmaningen för företag och organisationer. Det finns inget enkelt sätt att fortbilda dagens ingenjörer att bli AI-inriktade och det finns inte heller

ett sätt att skapa AI-utbildningar som garanterat ger oss en ny generation av högutbildade och kompetenta AI-systemutvecklare. Tyvärr, problemet är helt enkelt för svårt för en så enkel lösning som "mera och bättre utbildning" Att bli en bra och ansvarsfull AI-utvecklare innebär lång erfarenhet av storskalig systemutveckling, avancerad programmering, goda kunskaper i formalia (logik, matematik och matematisk statistik) samt en medvetenhet om etiska frågor. Allt detta kan givetvis även en ung person lära sig. Utvecklingsavdelningarna på några av de största företagens AI-avdelningar vittnar om det. Men de unga personer som arbetar där är inte ett tvärsnitt av befolkningen. De utgör den högsta spetsen av ett stort isberg av högutbildade människor, liksom deras (ofta något äldre) chefer, mentorer och lärare. Det är viktigt att inse hur liten den här gruppen är idag. De träffar varandra, exempelvis på AI-konferenser, där de inte ens utgör en procent av deltagarna; de flesta kommer för att lära av varandra. Gruppen är också socialt sett relativt homogen, män i 30-årsåldern är tydligt överrepresenterade. Det vore därför underligt att förutsätta att de AI-system de bygger kommer att realisera alla de drömmar olika representanter för grupper i befolkningen har om AI, vackra drömmar såväl som mardrömmar.

AI-utvecklingens yttersta risker

Olle Häggström, professor i matematisk statistik vid Chalmers

Utvecklingen inom artificiell intelligens (AI) har stor potential att göra våra liv bättre. Den förväntas ligga till grund för en stor del av den ekonomiska tillväxten under det kommande årtiondet, och på längre sikt än så är möjligheterna närmast obegränsade – annat än av fysikens lagar. Men utvecklingen kan också gå åt fel håll. . .

Dessa positiva sidor tas upp på annat håll i denna rapport. Här skall jag istället koncentrera mig på AI-utvecklingens risker. Sådana finns, och de behöver tas på allvar. Den idag rådande andan inom teknikutveckling kan jämföras med att köra bil med gasen i botten och ögonen enbart på hastighetsmätaren utan en tanke på vilka hinder och andra farligheter som kan finnas där ute. Detta duger inte. Vi behöver se ut genom vindrutan, vi behöver hålla i ratten så att vi kan styra rätt, och vi behöver ha bromsberedskap ifall något riktigt farligt dyker upp på vägen. Vi behöver också ha en idé om vart vi vill åka. Att förutse teknikutvecklingen och styra den så att vi kan skörda de frukter den utlovar, samtidigt som vi undviker de värsta riskerna, är givetvis svårare än att trafiksäkert framföra en bil, men med tanke på vad som står på spel (hela mänsklighetens framtid) behöver vi i alla fall försöka.^{H16}

AI-utvecklingen är mångfacetterad, och det gäller även dess risker. Det här kapitlet kommer att bjuda på ett brett svep, från några av de risker som föreligger redan idag och hela vägen fram till det yttersta riskscenariot: kan vi behålla kontrollen den dag AI-utvecklingen frambringat en maskin som överträffar oss människor i alla de förmågor vi förknippar med mänsklig intelligens?

Närliggande risker

AI-baserad bildbehandling är av stort värde inom exempelvis filmindustrin. Att denna teknik också har en baksida stod klart i slutet av 2017 då en internetanvändare med pseudonymen Deepfakes publicerade en uppsättning pornografiska videoklipp som felaktigt men mycket realistiskt gav intryck av att visa några av världens mest kända skådespelerskor. Klippen var gjorda med AI-teknik för så kallad *face swap*, där en persons ansikte kan bytas mot en annans. I början av 2018 släpptes en app som låter vem som helst göra samma sak.¹ Det återstår att se om följden blir en våg av hämndporr och andra skadliga tillämpningar. Optimisten kan här anföra att problemet löser sig självt då ju teknikens tillgänglighet gör att den som utsätts för hämndporr kan hävda att det rör sig om förfalskningar. Men om det är riktigt, hur går det då med videobezis i rättegångar? Och vad gör det med vår förmåga att mönstra ut *fake news* från nyhetsrapportering? Problemen hopar sig.

Detta är bara ett av många exempel på ny teknik som omger oss i våra telefoner och på annat håll, och som kan användas såväl för goda som för onda syften. Då och då har jag hört teknikutvecklare framhålla att ingen teknik är ond eller god i sig – det är den konkreta användningen som avgör. På så



vis duckar de för moraliskt ansvar för den teknik de utvecklar. Jag har svårt att acceptera den hållningen. Om exempelvis utvecklingen av en teknik kan förutses leda till att terrorister får tillgång till ett vapen så kraftfullt att övriga samhället får svårt att värja sig, då är det en situation där teknikutvecklarna inte gärna kan friskrivas från moraliskt ansvar.

Hur är det exempelvis med postorderjätten Amazons projekt att utveckla ett system för drönbaserad leverans av bokpaket? Det är inte helt uppenbart att det går att utveckla

en sådan teknik med egenskapen att den inte enkelt kan omprogrammeras till att leverera bomber istället för bokpaket till en adress. Om inte det lyckas finns risken att tekniken får katastrofala samhällskonsekvenser.

Detta för oss in på militära drönare och besläktad AI-teknologi för så kallat autonoma vapen – eller, med en mindre artig term, mördarrobotar. För att understryka hur farlig denna teknikutveckling kan bli, låt mig citera ur det öppna brev från 2015 med rubriken ”*Autonomous Weapons: An Open*

Letter from AI and Robotics Researchers” som tusentals forskare världen över (även jag) skrev under, och som uppmanar till moratorium för utveckling av autonoma vapen:^R

Om någon av de stora militärmakterna väljer att satsa på utveckling av AI-vapen så uppstår oundvikligen en global kapprustning vars teknologiska slutpunkt är uppenbar: autonoma vapen kommer att bli framtidens Kalashnikovs. Till skillnad mot kärnvapen kräver inte dessa några kostsamma eller svårtillgängliga råvaror, och de kommer därför att bli riktigt förekommande och lätta för alla någorlunda stora länder att massproducera. Det blir bara en tidsfråga innan de dyker upp på den svarta marknaden och i händerna på terrorister, diktatorer som vill öka kontrollen över den egna befolkningen, krigsherrar med avsikt att genomföra etnisk rensning, etc. Autonoma vapen är idealiska för lönnmord, folkförtryck, destabilisering av nationer, och det selektiva dödandet av någon viss etnisk grupp. Vi tror av dessa skäl att en militär AI-kapprustning inte vore till gagn för mänskligheten.

Den sista meningen är givetvis en kraftig underdrift. Ändå träffar man ibland på någon som försöker tona ned dessa risker. Ett exempel är kognitionsvetaren Steven Pinker, i en pannediskussion vi båda deltog i 2017.^{H17} Han framhöll att det skulle krävas en galning för att konstruera något så hemskt som ”en svärm av robotinsekter avsedd att attackera enskilda människor baserat på ansiktigenkänning”, och att det idag inte längre finns utrymme för sådana enstaka galningar eftersom alla betydande ingenjörprojekt numera genomförs i stora samarbeten.

Jag delar inte alls Pinkers obändiga optimism här, då han ju helt ignorerar den inre logiken i kapprustningar och det militärindustriella komplexet. Han förbiser också att minst lika hemska massförstörelsevapen har utvecklats i över 70 års tid – inte av ensamma galningar, utan just i stora samar-

betsprojekt (varav Manhattanprojektet är det mest kända). Varför skulle allt sådant utvecklingsarbete plötsligt upphöra av sig självt? Faran med en militär AI-kapprustning är verklig och något vi behöver förhålla oss till.

Arbetsmarknad

Kommer AI-utvecklingen att leda till så kallad *teknologisk arbetslöshet*, alltså arbetslöshet orsakad av rationalisering till följd av tekniska framsteg? Att människors arbetsuppgifter övergår till maskiner är inget nytt fenomen. I takt med jordbrukets effektivisering gick dess andel av den svenska arbetskraften från cirka 75 procent vid mitten av 1800-talet till cirka 3 procent idag, men det försatte inte 72 procent i arbetslöshet. De gick istället till andra sektorer (inledningsvis främst industrisektorn, därefter allt mer till tjänstesektorn).^S Arbetslösheten har gått upp och ned i takt med ekonomins konjunktursvängningar, men över längre tidsperioder har den trots teknikutvecklingen inte skenat. Men kommer detta – att arbetskraften hittar anställningar i nya sektorer i ungefär samma takt som de rationaliseras bort från gamla – att bestå? Detta är knappast självklart med tanke på ett antal nya omständigheter, som att det idag inte längre bara är fysiska och manuella arbetsuppgifter som automatiseras bort, utan även intellektuella (ett ofta citerat exempel är automatiserad journalistik^M). Att vi i evighet skall kunna hitta nya områden där den mänskliga förmågan överstiger maskinernas är inte givet, särskilt inte om AI-forskningen lyckas skapa en så kallad *superintelligens* – en maskin som vida överträffar människan i fråga om generell intelligens.

Ibland påstås att vår räddning från att bli utkonkurrerade skulle bestå i att kombinationen människa plus maskin alltid är bättre än maskinen ensam. Den tankegången håller inte, ty om maskinens överlägsenhet på något område blir tillräckligt

stor blir människans bidrag betydelselöst. Som min matematikerkollega Johan Wästlund påpekat: ett X2000-tåg inte blir märkbart snabbare om en människa med handkraft hjälper till att putta på.

Man kan givetvis ifrågasätta om den teknologiska arbetslösheten i grunden är något dåligt. Lönearbete kanske trots allt inte är livets mening? Visdiktaren Kjell Höglund påpekar i sin sång *Maskinerna är våra vänner* att om de tar över våra arbeten får vi mer tid över för konst, kultur, kärlek och förströelse. Ett samhälle med 100 procent arbetslöshet är i viss mening jämlikt, men det är inte uppenbart hur ett sådant samhälle kan organiseras. Om det är dit vi vill så behöver vi också ha en fungerande plan för vägen dit, vilken rimligtvis går via arbetslöshetsnivåer på 20 procent, 50 procent och 90 procent. Hur kan vi passera sådana övergångsstadier utan att vidga de ekonomiska klyftorna ofantligt och utan att skapa social instabilitet? Det här är svåra frågor som vi inte har tydliga svar på idag.

Dagens arbetslöshetssiffror kan knappast tolkas som början på en eskalerande teknologisk arbetslöshet,^A men läget kan förändras snabbt. Utvecklingen av autonoma fordon kan på bara ett par decenniers sikt väntas leda till en strukturomvandling där många miljoner taxi- och lastbilschaufförer världen över förlorar sina arbeten i något som kan bli en närmast fullständig utträdning av en hel arbetsmarknadssektor. En liknande utveckling kan komma snabbt i en rad andra sektorer,^{BM, FO} varav textilindustrin är en. Den har hittills varit mycket personalintensiv till följd av att det varit svårt att utveckla industrirobotar som förmår hantera mjuka material, något som AI-metoder de allra senaste åren dock kommit att ändra på. Följden kan bli en snabb automatisering av textilindustrin, så att de ekonomiska incitamenten för klädföretag i väst att förlägga produktionen till låglöneländer försvinner.

Om den då flyttas hem blir konsekvenserna svåra för en rad av tredje världens ekonomier.^{SN}

Jag nämnde ovan scenariot med uppkomsten av en superintelligent AI. AI-futurologen Eliezer Yudkowsky menar att ett sådant har så långtgående konsekvenser att frågor om arbetsmarknad skjuter bredvid målet.^{Y13}

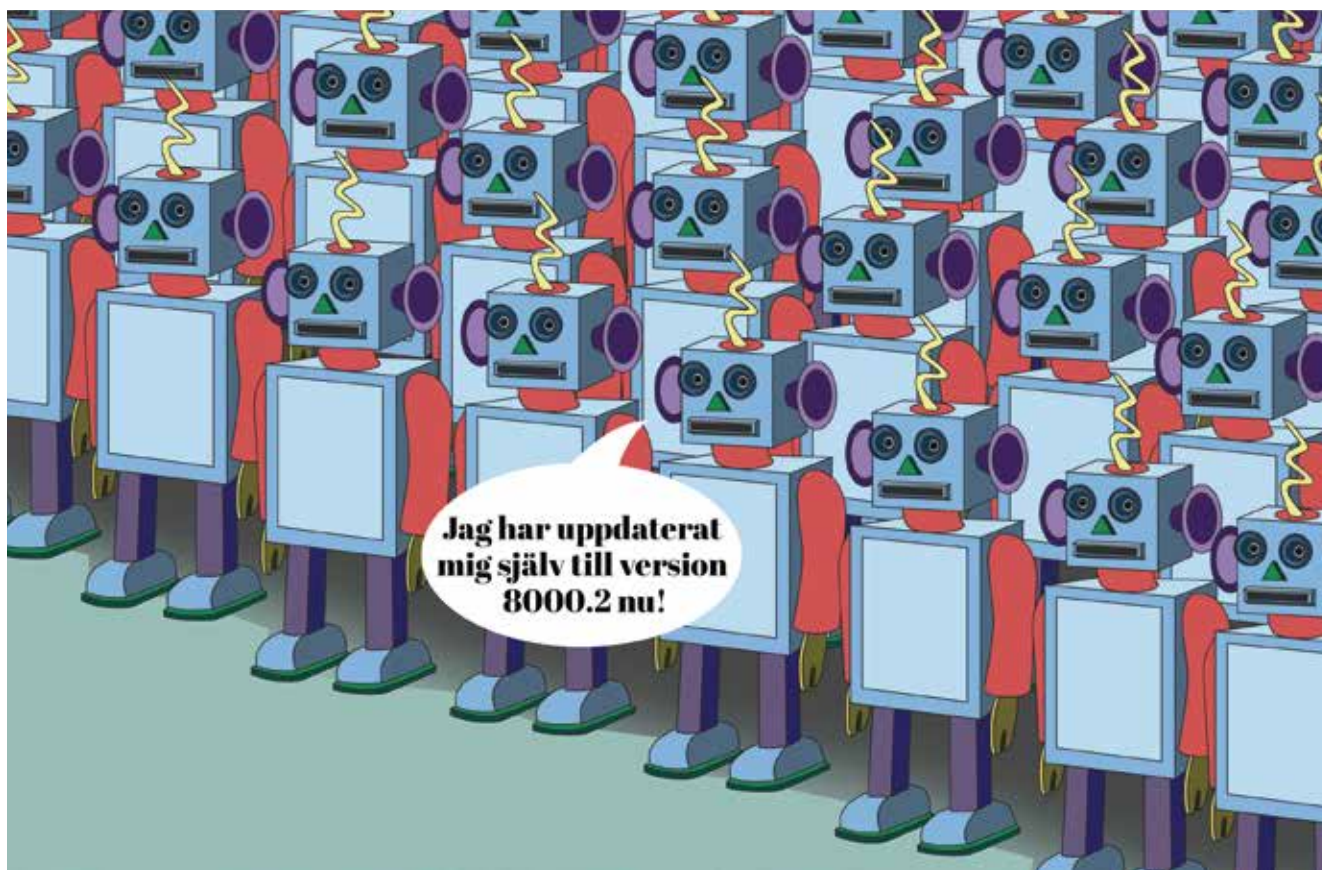
Att fråga om vilka konsekvenser en superintelligent maskin skulle ha för konventionell mänsklig arbetsmarknad är som att fråga vilka följder det skulle få för handeln mellan USA och Kina om månen kraschade in i jorden. Följder skulle det förvisso bli, men frågeställningen missar poängen.

I nästa avsnitt behandlas, ur ett större perspektiv än arbetsmarknad, tänkbara följder av ett teknikgenombrott som skapar en superintelligent AI.

Superintelligens

Alan Turing var en av 1900-talets främsta matematiker, och den som i backspegeln kan ses som datavetenskapens grundare. Redan 1951 diskuterade han på fullt allvar möjligheten av en framtida superintelligent maskin.^T Någorlunda bred popularitet (förutom i science fiction-litteraturen) vann begreppet först med IT-visionären Ray Kurzweils bok *The Singularity is Near* från 2005, där genombrottet för superintelligens beskrivs som det avgörande steg som skall hjälpa oss människor att befria oss från våra skröpliga kroppar och ge oss allt vi kan önska, inklusive erövrandet av världsrymden.^K Därefter har diskussionen, mycket tack vare den tidigare nämnde Eliezer Yudkowsky och den svenskfödde filosofen Nick Bostrom, kommit att skifta karaktär – bort från Kurzweils evangeliska tonfall och mot en större medvetenhet om att ett AI-genombrott kan ha betydande risker.^{Y08, B}

Mer om dessa risker strax. Men först – ordet singularitet i Kurzweils boktitel syftar på ett hypotetiskt scenario där AI-



utvecklingen når superintelligensnivåer väldigt plötsligt och snabbt. Vad som får vissa tänkare på området att tro på något sådant är idén om ett slags upprepad självförbättring, där en AI konstruerar en snabbare och bättre version av sig själv, som i sin tur konstruerar en ännu bättre version, och så vidare i en snabbt eskalerande spiral mot allt högre intelligensnivåer. Expertisen är oenig om huruvida ett sådant scenario är troligt, eller om ett mer gradvis förlopp är att vänta.^B Men

om en singularitet kan inträffa så gör det situationen extra farlig, då den ger oss mycket lite tid att anpassa oss till den nya situationen.

Singularitetsdiskussionen gäller alltså *hur plötsligt* utvecklingen till superintelligensnivåer kan ske. Frågan om *när* en superintelligent AI kan väntas är en annan. Även här är experter oeniga, och deras uppskattningar av när superintelligens kan bli möjlig är spridda över hela det kommande århundra-

det och mer än så. Vissa hävdar rentav att något sådant aldrig är att vänta.^B Oenigheten är inte förvånande med tanke på att vi inte ens vet vad för slags AI-teknologi som är den mest lovande vägen till artificiell generell intelligens. De mest spektakulära framstegen på senare år har skett inom den neurala nätverks-teknologi som kallas *deep learning*, men det är inte alls säkert att just den kommer att fortsätta att dominera AI-utvecklingen. Kanske får vi se ett uppsving för metoder som mer naturtroget efterliknar den mänskliga hjärnans kopplingsschema, eller som bygger på evolutionsteorins princip om naturligt urval. Vi skall inte heller utesluta en renässans för den mer direkta programmeringsmetod som kallas GO-FAI (*good old-fashioned AI*). Ett genombrott kan också komma ur någon kombination av dessa metoder, eller ur något helt nytt och ännu oförutsett.

Huruvida en superintelligent AI blir farlig eller inte hänger i hög grad på dess drivkrafter och mål: vad är den programmerad att försöka uppnå? Om dess mål inte ligger i linje med mänsklig välfärd (eller vad det nu är för mänskliga värderingar vi önskar befrämja) så kan vi vara illa ute. Ett känt och avsiktligt karikatyrartat tankeexperiment med syfte att visa hur även till synes harmlösa mål kan få katastrofala följder är det så kallade *Paperclip Armageddon* – den stora gemkatastrofen. Det börjar med en AI som är satt att styra en gemfabrik och som därför programmerats med målet att maximera gemproduktionen. Denna AI råkar bli fröet till den självförbättringsspiral och singularitet som skapar en superintelligens, vilken raskt skrider till verket genom att omvandla all materia den kommer åt (inklusive oss människor) till gem.

Ingen tror på allvar att just gemkatastrofen kommer att bli verklighet, men en svårighet är att även försök att mer medvetet skapa en god framtid lätt kan få katastrofala följder. En superintelligent AI med målet att maximera det totala

välbefinnandet i världen skulle precis som gemmaximeraren antagligen bli vår undergång, då människokroppen rimligtvis är långt ifrån optimal när det gäller att maximera mängden välbefinnande per kilogram materia. Både Yudkowsky och Bostrom menar därför att det centrala (men också mycket svåra) problemet är att säkerställa att den första superintelligenta AI:n har (för oss) gynnsamma mål.^{Y08, B} Detta problem benämns numera *AI Alignment*.

Många försök har gjorts att påvisa att vi enkelt kan komma runt problemet, men inget av dem övertygar. Det kanske vanligaste är ”vi kan ju alltid dra ur sladden!”. Men så snart en superintelligent AI har tillgång till Internet och i kraft av sin överlägsna intelligens kan gå igenom (eller runt) våra brandväggar, så kan den skapa tusen strategiskt utplacerade kopior av sig själv och därmed bli i praktiken immun mot sladdurdragning. Tanken att en superintelligent AI varken skulle hitta oväntade uppkopplingsvägar eller kunna övertala (med lock och pock eller annan list) en mänsklig datoradministratör att bereda uppkoppling är mer än lovligt naiv.^{H16}

Ett annat vanligt försök att slippa ta itu med *AI Alignment* är att hävda att den som är superintelligent själv kommer att förstå vad som är rätt och fel, och exempelvis inse att bevarande av ekosystem och främjande av mänsklig välfärd är viktigare än gemproduktion. Att på detta vis ta för givet att de värderingar vi människor har kommer att delas av varje annan intelligent varelse är ett misstag av det slag som brukar kallas *antropomorfism*. Om vi vill att en superintelligent AI skall prioritera ekosystembevarande så behöver vi lösa *AI Alignment*-problemet i tid, så att vi hinner förse AI:n med denna värdering innan den blivit superintelligent och det är för sent att göra något åt saken.

Fyra förändrande drivkrafter för AI i vården

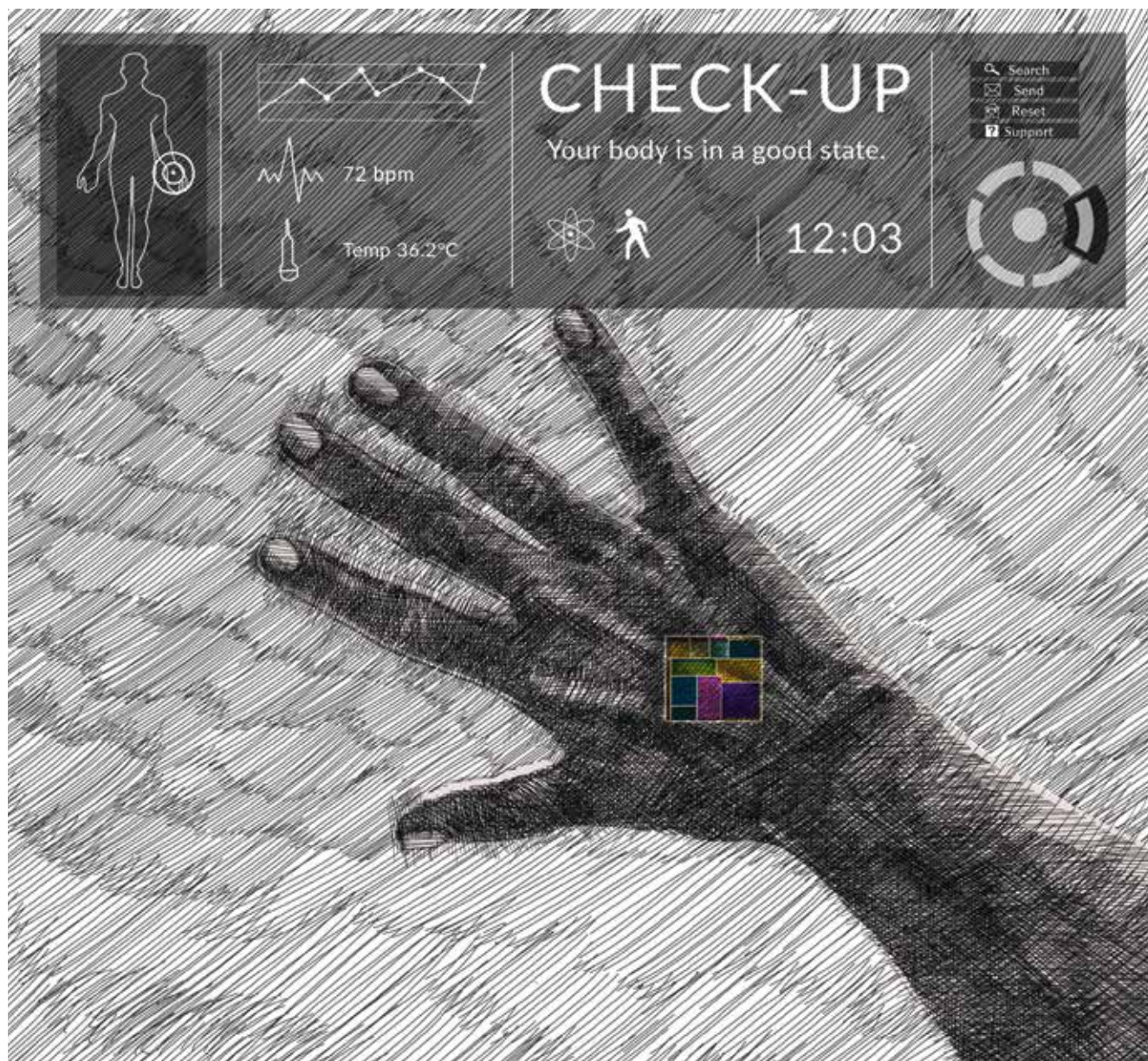
Paulina Modlitba, senior konsult inom innovation och digitalisering

Predictive, personalized, preventive, and participatory, på svenska förutsägande, personligt, förebyggande och deltagande – kallas också “P4” och är ett paradigm som har sitt ursprung inom bioteknik och medicin. Med åren har det dock kommit att tillämpas på bredare front, ofta kopplat till innovation och dataanalys.

Begreppet P4 myntades av bioteknikpionjären Dr Leroy Hood, som drivs av idén att tillgång till storskalig och billig medicinsk data i molnet kommer att revolutionera vårdsektorn¹. Istället för att bota sjukdomar, kommer vi enligt Hoods vision med hjälp av bland annat genomisk kartläggning, kliniska mätvärden och livsstilsdata kunna förutspå, individanpassa och motverka sjukdomar i framtiden på ett helt annat sätt än vad vi kan idag. Mycket av detta kommer att ske i samråd med patienterna själva. Därav paradigmet med fyra P. Dr Hood var själv med och lanserade en DNA-sekvenseringsmaskin som blev av betydelse för att kartlägga det mänskliga genomet. Med ny, effektiv teknik har kostnaden för att avläsa en människas genom stadigt sjunkit under senare år; idag är det möjligt att sekvensera ett helt mänskligt genom på en dag för under 8000 kronor.

I snabb takt närmar vi oss en framtid där hälsovården och samhället i stort har genomgått ett paradigmskifte från att lindra och bota skador och sjukdomar till att motverka dem genom att hålla den allmänna hälsolivn så hög som möjligt. Men till skillnad från förr sker detta inte längre med hjälp av generella kostråd och rekommendationer utan med analys

av stora mängder anonymiserade data som ger individanpassade och livsstilsrelaterade rekommendationer. Hälso- och livsstilsdatan samlas in löpande utan att individerna själva aktivt behöver logga sitt beteende. Istället kommer vi i allt större utsträckning att samla data via sensorer, e-tatueringar, små microchip på eller inopererade under huden, eller till och med cirkulerande som små kapslar i blodet. Till detta tillkommer data som registreras av olika sensorer omkring oss – i hemmet, på arbetsplatsen, på gymmet och andra platser där vi godkänt att den samlas in - och dessutom kommer de flesta av oss att ha sitt ”genomdata” i mobilen. Här kan AI spela en viktig roll för att identifiera mönster och för att förstå vilka genetiska anlag, livsmönster och andra faktorer som påverkar vår hälsa och livslängd. Det kan sedan återkopplas och tala om för oss på individnivå när det är dags att uppsöka sjukvården. Den stora vinningen är att de självlärande systemen, som har tillgång till data från miljontals individer, kontinuerligt uppdaterar insikter och rekommendationer i takt med att ny data tillförs. De insikter som utvecklas kan dessutom enkelt och snabbt delas mellan olika system eftersom allt lagras i molnet.



Att värna om personlig integritet och etisk hantering av data är i dessa sammanhang centralt. Det finska initiativet MyData, som numera också är en årlig konferens, lanserades 2015 i rapporten "MyData – A Nordic Model for human-centered personal data management and processing"². Utgångspunkten för MyData är att etablera en modell som ger individer bättre kontroll över sin egen data, vilket i sig ökar integritet och valmöjligheter, men också möjligheten för individer att tjäna pengar på sin egen data. Medan EU:s nya dataskyddsförordning, GDPR, syftar till att skydda data vill MyData alltså ta det ett steg längre genom att driva på ett paradigmskifte där individuella användare tar makten över och till och med kan tjäna pengar på sin egen data.

Forskare från RISE SICS³ och Uppsala universitet har precis beviljats medel av SSF för ett projekt döpt till "Hacka inte min kropp". Teamet har nyligen kunnat visa att fettlagret mellan muskler och hud kan användas som radiokanal med större bandbredd än många andra lösningar; en insikt som kan komma till nytta på många olika områden såsom hjärnmaskin-interaktion och för upptäckt av bröstcancertumörer. Projektet inkluderar även framtagandet av en bredare säkerhetsarkitektur för implantatnätverk i kroppen och för säker anslutning till molnet. Målet är att garantera konfidentialitet och integritet.

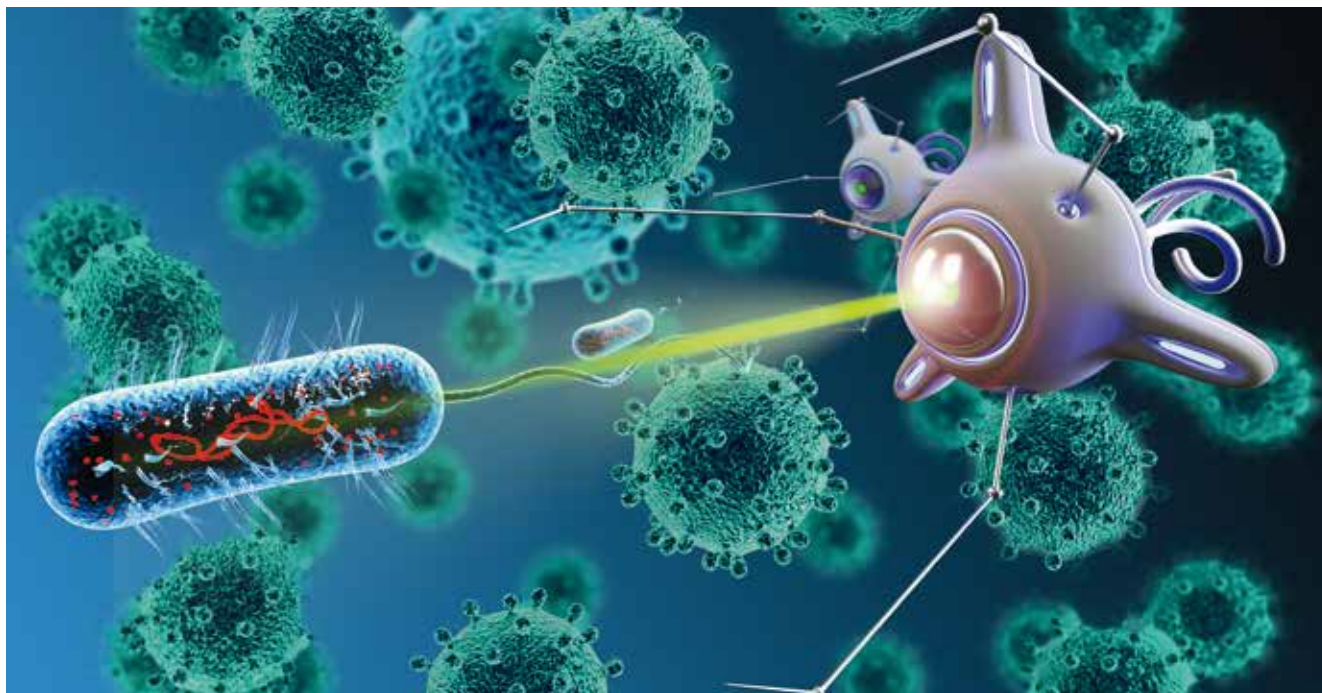
Predictive

En stor fördel med AI är att den inte bara drar slutsatser baserat på historiska data, utan även kan utveckla modeller för framtida händelser. Digitala marknadsförare kan då få ett utvecklingsstöd för att minska de antal kunder som avbryter ett köp, självkörande bilar kan undvika allvarliga tillbud och industrier minimera sina driftsstörningar.

Kina har deklarerat att man satsar på att bli en ledarnation när det gäller AI.⁴ Orsakerna är flera, men det är tydligt att det kinesiska samhället är i stort behov av effektivisering och stöd av kognitiva system³. Inte minst inom sjukvården. Kinesiska sjukhus är kraftigt underbemannade med i genomsnitt 1,5 läkare per 1000 medborgare. Parallellt med detta står Kina inför enorma hälso- och vårdutmaningar. Landet har flest överviktiga barn och diabetessjuka i världen, samtidigt som andelen äldre ökar snabbt. Som ett led i detta har både staten och den privata sektorn satsat på att utveckla framtidens sjukhus. Så finns det till exempel en bildanalys-tjänst, Miying, som hjälper läkare att identifiera tidiga tecken på cancer, eller en digitaliserad tjänst som analyserar ögonpatienter och lyckats pressa ner analys tiden från flera dagar till sekunder. Tekniken används redan på hundratals sjukhus runtom i Kina.

En annan aktör är Alibaba, företaget som började med e-handel och nu växt till ett stort konglomerat. De har utvecklat AI-mjukvara som kan analysera skiktröntgenbilder. Dessutom har Alibaba investerat i AI-stödda medicinska labb för snabbare diagnoser.

I Sverige är Karolinska Institutet en av de institutioner som leder utvecklingen av AI inom sjukvården. Mattias Nilsson Benfatto och Gustaf Öqvist Seimyr har bedrivit forskning kring hur AI kan användas för att förutspå dyslexi i ett tidigare skede med hjälp av ögonrörelsedata. Resultaten visar på en så hög överensstämmelse, 90 procent, mellan AI-programmets bedömning och de traditionella bedömningarna, att tekniken kommersialiserats i form av bolaget Optolexia, som erbjuder dyslexiscanning till skolor^{5,6}. På institutionen för kliniska vetenskaper på Danderyds sjukhus används redan idag AI för att analysera röntgenbilder i syfte att identifiera frakturer och bedöma vilka patienter som behöver opereras.



Hos Enheten för Experimentell astma- och allergiforskning som drivs av professor Sven-Erik Dahlen används självlärande system för att identifiera och klassificera olika typer av svår astma.

Personalized

Googles VD, Sundar Pichai, kungjorde nyligen att de gått över till att tänka "AI first". Ett steg i denna transformation är Googles köp av BrainQ, som applicerar AI på EEG-data⁷ för att ta fram individanpassade behandlingsplaner för bland annat strokepatienter⁸. EEG-data kritiseras ofta för att vara brusig och opålitlig men här menar BrainQ att just AI kan spela en viktig roll i att skilja brus från signal. Om tekniken

visar sig fungera har den potential att hjälpa patienter att få tillbaka känsel och rörelseförmåga i påverkade kroppsdelar.

Ett helt annat användningsområde är prenumerations-tjänsten Freda som analyserar menstruationsdata och skickar hem tamponger och bindor när mensen närmar sig. Här kan man enkelt se framför sig hur vi i framtiden kommer ha än mer pålitlig biodata om menstruation och andra faktorer med hjälp av olika e-tatueringar¹⁰, smarta plåster, inopererade chip¹² och uppkopplade piller som cirkulerar i kroppen.¹³

Preventive

Fördelarna med AI blir kanske än mer tydliga när den används för preventiva ändamål. På Stanford-universitetet har

ett forskarlag lett av Andrew Ng utvecklat ett neuralt nätverk som kan förutspå kvarvarande livslängden (3-12 månader) hos patienter som befinner sig i palliativ vård, baserat på stora mängder data i läkarjournaler¹⁴. Tanken är att systemet ska kunna användas för att individanpassa vården efter den enskilda patientens behov, men också att effektivisera och optimera fördelningen av läkare så att de patienter som är i störst behov av läkarvård får det.

Ljud kan också vara en källa till insikter kring hälsotillstånd. Forskare använder röstdata bland annat för att förutspå och identifiera posttraumatiska tillstånd, depressioner, allvarligare hjärnskador och hjärtattacker.¹⁵

Microsoft har utvecklat ett eget självlärande system^{16,17} för att kvalitetssäkra ingrepp med genkniven CRISPR och hjälpa forskare att bedöma riskerna för oönskade bieffekter med enskilda DNA-modifieringar och påskynda editeringsprocessen. Grundproblemet är att många av regionerna i vårt genetiska material är snarlika och det är därför lätt hänt att CRISPR modifierar fel gen. Vissa gener är mer benägna att förvirra CRISPR-tekniken än andra; de är "högriskgener". AI används här för att kartlägga och undvika denna typ av misstag och påföljande bieffekter.

Själv mord är ett annat område där AI redan idag visat sig kunna spela en förebyggande roll. I en studie publicerad i *Nature Human Behavior*¹⁸, användes maskininlärning för att analysera hjärnaktiviteten hos 17 individer som fick höra en lista med positivt och negativt laddade ord, samt ord sammankopplade med själv mord (till exempel "dödlig"). Kombinerat med information om individernas tidigare själv mordsbenägenhet lyckades man identifiera 15 av 17 av de som tidigare försökt ta sitt liv och 16 av 17 i kontrollgruppen. Forskare på Florida State University har analyserat anonymiserade journaler från 2 miljoner patienter för att på



liknande sätt identifiera vilka faktorer som indikerar en förhöjd själv mordsbenägenhet¹⁹. Facebook har också liknande projekt igång.

Även minnet kan förmodligen påverkas med hjälp av insikter från själv lärande AI-system. Det är sedan tidigare känt att direktstimulering av hjärnan kan påverka och förstärka registreringen av minnen²⁰. Det är dock viktigt att stimuleringen sker vid exakt rätt tidpunkt, när minnet registreras. Tekniken är dock fortfarande i ett tidigt utvecklingsskede och datamängden relativt liten, varför mer avancerade och pålitliga slutsatser är svåra att dra²¹.

Participatory

Det finns mycket som talar för att vi som patienter kommer engagera oss allt mer i vår egen hälsodata och vård i framtiden. Internet, digitala tjänster och smarta bärbara prylar gör

det redan idag möjligt för gemene man att 1) söka information om sjukdomar och symptom, 2) löpande mäta och följa sitt eget hälsotillstånd, och 3) konsultera läkare på distans via tjänster som Kry och MinDoktor. Det nya europeiska datadirektivet GDPR bidrar därtill till att stärka individers data- och integritetsrelaterade rättigheter och verkar för en ökad transparens kring vilken data som samlas in och lagras, och hur den får användas.

När maskin och människa i allt större utsträckning kommer samexistera och interagera blir insikterna kring hur design påverkar tillit centrala. John Zimmerman, professor i människa-datorinteraktion på Carnegie Mellon universitetet i USA har länge studerat vilka faktorer som bidrar till att människor vågar lita på och faktiskt använder självlärande system. Studien visar att det bland annat är viktigt att integrera AI-baserade rekommendationer i existerande data och rutiner istället för att helt ersätta den²². Tajmingen är också viktig; rekommendationerna bör presenteras i direkt anslutning till beslutsfattandet. Det har också visat sig att personer som får möjligheten att finjustera rekommendationerna litar mer på systemet. Med andra ord är samarbete och integration centralt.

Etik och moral

Efter många och motiga decennier är det nu tydligt att AI börjat få sitt stora genombrott i samhället. De praktiska tillämpningarna som kommer samhället och oss medborgare till gagn blir snabbt fler. Parallellt med diverse skräckscenarior, såsom innovatören Elon Musk som hävdar att AI är ett större hot än kärnvapen²³, är tilltron till vad AI redan idag kan åstadkomma stor; till och med överdriven enligt vissa²⁴. Inte minst av kommersiella skäl. IBM fick, efter en större journalistisk granskning, nyligen revidera bilden av hur avancerad deras "IBM Watson for Oncology"-system är i att bistå

cancersjukvården²⁵. En studie från Cornell universitetet och universitetet i Wyoming visar att det är relativt enkelt att lura algoritmer som är tränade att identifiera foton²⁶.

Samtidigt har diskussion om partisk AI och behovet av ökad transparens tagit fart. Redan idag har Google utvecklat AI som har förmågan att utveckla egen AI som är mer avancerad än den vi människor kan skapa²⁷ och då blir det särskilt angeläget att lösa detta så kallade "black box problem" genom att utveckla självlärd system som kan förklara sina algoritmer och insikter för oss människor. Bland annat forskare på UC Berkeley, Universitetet i Amsterdam och Facebook AI Research samarbetar kring detta och publicerade nyligen en artikel med lovande resultat²⁸. Enligt artikeln är deras AI-modell för analys av bilder den första i världen som kan motivera sina egna beslut med naturligt språk och dessutom kan peka ut beläggen i bilderna.

Mot bakgrund av detta är det redan idag ytterst angeläget att både offentlig och privat sektor gör sig redo för en framtid där AI och människor kommer samarbeta och samverka i olika former. En framtid, eller snarare nutid, där självlärande system kan assistera oss människor i att fatta säkrare och mer effektiva beslut, utföra specifika rutinuppgifter, identifiera mönster och insikter utanför de ramar vi är vana att verka inom och förstå. Tack vare detta frigörs mänskliga forskares och läkares tid till att utföra fördjupade analyser, fatta slutgiltiga beslut, bemöta patienter och utbyta kunskap.

Det är också viktigt att redan idag utveckla tydliga etiska riktlinjer och lagar för AI. Hur bygger och implementerar vi uppkopplade självlärande system som optimerar och utvecklar sina egna algoritmer utan att de förstärker eventuell skevhet och partiskhet i den data den tränats på och äventyrar enskilda individers integritet och säkerhet? Det är en av vår tids viktigaste frågor.

Vart kommer AI att köra oss?

Anna Pernestål Brenden är föreståndare för Integrated Transport Research Lab, KTH

Drömmen om bilar som kör själva har funnits sedan 1920-talet, men det är först nu, en bit in på 2000-talet som det har blivit verklighet. Förlärlösa tåg har funnits längre, och automatiserade så kallade spårtaxi finns sedan ett tag på flygplatser, så varför dröjer det så länge med förlärlösa bilar? Vad är det som gör att vi nu börjar höra talas om dem? Och hur kommer självkörande bilar att påverka oss?

Bilar behöver kunna navigera själva, eftersom de till skillnad från tåg inte har räls som styr dem. Men den största utmaningen för självkörande bilar är den oförlärlösa miljön som finns runt omkring dem: andra fordon (körda av människor), cyklar och fotgängare. Och även om vi har trafikregler så är det inte alltid vi följer dem... Människor är väldigt bra på att tolka den här komplexa miljön, se omgivningen och kunna förutse vad som kommer att hända.

Under de senaste åren har de förutsättningar som behövs för att bilar ska kunna klara de här komplexa miljöerna och bli självkörande kommit på plats. För det första finns nu nya sensorer¹ som kan "se" omgivningen noggrant. För det andra finns små och kraftfulla datorer som kan göra stora mängder beräkningar. För det tredje finns algoritmer som bygger på maskininläring², det vill säga, de kan bli bättre på att utföra vissa uppgifter genom att träna. Tack vare de små datorerna kan avancerade program köras i bilen, informationen från sensorerna kan sättas samman och det går tillräckligt snabbt att tolka vad sensorerna "ser" för att datorerna ska kunna tolka vad som händer runt omkring fordonet, och till slut välja vad bilen ska göra. Tack vare maskininläring kan sedan bilarna tränas upp genom att utsättas för en mängd olika sce-



narier, precis som när vi människor tränar för att ta körkort. Med den skillnaden att det bara är en bil som behöver lära sig, och med den finessen att det bara är att kopiera över den kunskapen till alla andra bilar! Just detta – att kunna se och tolka den komplexa världen runt omkring är central för att kunna skapa fordon som kör själva på vägar tillsammans med andra trafikanter.

Självkörande fordon är viktigt för AI

Utvecklingen av självkörande fordon driver AI framåt. Längtan efter självkörande bilar och bussar är stor. Många vill slippa hålla i ratten under färden. Antalet trafikolyckor kan minska, eftersom robotar inte blir trötta eller onyktra. Trafik- och stadsplanerare ser att vägarna skulle kunna utnyttjas effektivare och att man kan minska på antalet parkeringsplatser. Operatörer av transporttjänster (som taxi, bussar och

godstransporter) kommer kunna minska driftskostnaderna eftersom man inte behöver betala förarlöner. När man inte behöver betala för förarlöner kan godstransporter köras långsammare, vilket leder till minskad bränsleförbrukning. Fordonstillverkare får nya möjligheter att sälja ny teknik och nya transporttjänster. Självkörande fordon är något som kommer att synas på gator och vägar, och det är ett område där AI kan få stor påverkan på våra vardagliga liv.

Vad blir konsekvenserna av självkörande fordon?

AI och självkörande fordon har potentialen att drastiskt förändra hur vi reser och transporterar gods, hur städer och samhällen fungerar. I tabellen nedan listas några av de effekter som forskare och analytiker ser som möjliga^{3,4}.

Flera av dessa effekter påverkar varandra. Vissa kommer att förstärka varandra och andra motverka varandra. Effek-

Ökad kapacitet på vägarna på grund av bättre flöde i trafiken. Framför allt på motorvägar och större vägar.
Färre olyckor , eftersom man räknar med att cirka 90 procent av alla trafikolyckor beror på mänskliga faktorn.
Lägre energiförbrukning , tack vare "eco driving" och "platooning" (lastbilar som kan köra mycket nära varandra eftersom de är virtuellt sammankopplade. Detta gör att luftmotståndet minskar).
Godstransporter kan köras långsammare och på natten eftersom det inte finns någon förare. Det kan leda till mindre trängsel och lägre energiförbrukning.
Nya typer av transporttjänster och lägre kostnad för taxi/buss, eftersom man inte behöver ha en förare.
Färre parkeringsplatser , eftersom självkörande bilar kan parkera tätare och om vi använder de nya transporttjänsterna istället för att ha egen bil.
Vi åker istället för att gå eller cykla , eftersom det blir billigare att resa med taxi och buss.
Vi åker mera , eftersom det blir billigare och bekvämare att resa om vi slipper köra.
Högre energiförbrukning , på grund av flera energikrävande datorer och sensorer ombord, och på grund av att det är energikrävande att hantera stora mängder data.

terna av detta kommer att utforma samhället och var vi väljer att bo, på samma sätt som bilen påverkat samhällets utformning de senaste 100 åren. Vad är det som tar överhanden? Det är väldigt svårt att säga just nu.

Forskare har identifierat några faktorer som förutspås få stor påverkan på utfallet av självkörande bilar^{5,6,7}. Det kan handla om saker som:

Beskrivningarna av scenarierna är skrivna som berättelser år 2030, så låt oss göra en resa in i framtiden.

Scenarie 1: Samma men ändå inte (Proaktiv samhällsplanering – ingen delningsekonomi)

Det har hänt mycket sedan mitten av 2010-talet. Påtagliga effekter av klimatförändringar, tillsammans med att både

1. Vem driver utvecklingen: staten, det civila samhället eller näringslivet? Blir det samhällsplanerare som är drivande och drar upp riktlinjer och genomför regeländringar, eller tvärtom, kommer samhällsplanerare och vi konsumenter att vara mer avvaktande och näringslivet mer aktivt?

2. Kommer delningsekonomi att slå igenom? Vill vi dela, transporter, prylar och data? Är det okej att inte längre ha tillgång till egen bil? Eller fortsätter vi ungefär som nu med att gärna vilja ha våra egna saker? Det i sin tur kommer att påverka hur de självkörande bilarna används – antingen i delade tjänster eller som privata taxibilar.

3. Teknikutvecklingen – kommer vi att lyckas utveckla bilar som helt utan förare kan köra säkert och effektivt överallt? Eller visar det sig att det var för svårt, så att vi istället får bilar med avancerade förarstöd och som är självkörande på vissa sträckor, men ändå behöver en förare?

4. Kommer lagstiftningen att hänga med? Hur kommer lagstiftning och regler att utformas för att stödja utveckling och testning av självkörande fordon?

Hur skulle det kunna bli i framtiden?

Det finns alltså många osäkerheter kring hur det faktiskt kommer att bli när (om) självkörande fordon blir vardag. När det är så osäkert kan framtidsscenarioer vara ett sätt att ändå skapa sig en, eller egentligen flera, bilder av hur det kan bli. I detta avsnitt presenterar vi två framtidsscenarioer, som målar upp bilder av hur framtiden med självkörande fordon kan se ut om olika aktörer blir drivande och delningsekonomi slår igenom, eller inte. Dessa två scenarioer är resultatet av ett forskningsprojekt som genomfördes på KTH tillsammans med Statens väg- och transportforskningsinstitut under 2017⁸. De beskrivningar som ges här är bearbetade och förkortade versioner av de i rapporten.

myndigheter och företag faktiskt har utnyttjat många av de möjligheter som digitaliseringen ger, har lett till att vi idag har ett Sverige som både liknar och skiljer sig från förr. Samhället flyter på smidigare, och det har utvecklats en stor mängd tjänster som stöttar vardagslivet, som ger rådgivning och underlättar bra vardagsbeslut.

Det är egentligen bara ett område där framtidsprofetierna från 2018 inte infriats: svenskarna, liksom övriga européer, hakade inte på delningslösningarna. Oavsett om det handlar om delad data eller om delade prylar. Vi vill helt enkelt ha våra egna prylar och vårt eget utrymme. En anledning är flera återkommande digitala terrorattacker och ett cyberkrig runt år 2020. Man satsade tidigt på ett nationellt journalsystem,





men efter ett allvarligt hackerintrång föll förtroendet för systemet. Detta i kombination med intrånget i det nationella brottsregistret gjorde att folk helt enkelt inte litar på att staten kan garantera säkerheten. Det har också påverkat folks värderingar, och man vill ha kontroll över sina egna saker. Istället har det vuxit fram en marknad för olika "botar", det vill säga artificiella assistenter som hjälper oss att hålla reda på personliga data och på vilket sätt den utnyttjas.

Gaturummet har förändrats. Istället för att anpassa sig efter bilarna är nu fotgängare och cyklister i fokus, och det har blivit rejält dyrt att ha bil i städerna. Trängselavgifterna är differentierade och baseras, förutom på tiden på dygnet, också på vilken gata man kör på, hur stor bil man har, hur många som åker i bilen och hur mycket lokala utsläpp bilen orsakar. Detta har lett till att färre flyttar in till städerna.

Man bor utanför städerna, och jobbar hemifrån eller från lokala kontor som ofta delas av flera företag. De som pendlar gör det med små elfordon och kollektivtrafik. Genom platooning har kostnaderna för lastbilstransporter sjunkit för både bränsle och personal.

Jämfört med vad man trodde 2018 har utvecklingen av självkörande fordon bromsats upp, tekniken är för dyr för att ha en egen självkörande bil. I stället ser vi avancerade förarstöds-system och bilar som är självkörande på till exempel motorvägar i speciella filer, något som har lett till att trafiken flyter mycket smidigare. Många äldre känner frihet när de ändå kan köra sin bil högre upp i åldrarna.

Scenarie 2: Du får det du behöver - innan du ens vet att du behöver det (näringslivet driver -delningsekonomi har slagit igenom)

Det vi ser omkring oss i dag, år 2030, är på många sätt annorlunda än det vi var vana vid under 2010-talet. AI och digitalisering har skapat helt nya lösningar och tjänster som den stora majoriteten av nyfikna och progressiva svenskar inte bara gillar, utan verkligen älskar. Vardagslivet är mycket mer bekvämt och smidigt än förr. Då ägde vi våra prylar och hade kontroll, nu är det individuellt anpassade helhetskoncept som gäller. Idag finns det ett antal dominerande aktörer som hjälper oss med våra vardagsbestyr tack vare smarta digitala tjänster. Ett exempel är att Google gett sig in som operatör inom det offentliga och har byggt nya vägar inom ramen för sina egna transporttjänster. På dessa vägar har de som köpt deras nya tjänst "Seamless" egna filer, ungefär som kollektivtrafik hade egna körfält förr i tiden. De bolag som lyckats bäst är de som samlat unik data om sina brukare, och sedan använder det för att skapa individuella lösningar, där de själva som leverantör är spindeln i nätet. Personliga data har då blivit den nya valutan.

När det gäller utvecklingen av självkörande fordon har marknaden varit styrande. Den traditionella kollektivtrafiken ser ut ungefär som 2018. Samtidigt har utbudet av nya taxiliknande tjänster baserade på delade och självkörande fordon exploderat. Många har sålt sin bil, för att istället använda olika slags mobilitetstjänster. Alla stora köpcentrum kör egna transporttjänster, och hämtar upp sina kunder vid deras hem och kör dem till butikerna. En del storföretag har gått över till att använda "LinkedIn Workplace" – en tjänst som erbjuder en effektiv arbetsplats baserad på en kombination av företagets behov och medarbetarens livspussel. Tjänsten innebär att man prenumererar på "effektiv och kreativ arbetsproduktion". Den transporterar folk automatiskt till jobbet om man behövs där, men om det inte är viktigt att man fysiskt är på jobbet rekommenderar den folk att arbeta hemifrån istället. Allt baserat på avancerad dataanalys som har koll på vad som behöver produceras just den dagen, hur olika medarbetare bäst är effektiva, vilka som behöver mötas fysiskt eller digitalt, samt medarbetarnas individuella preferenser.

Utmaningar

Det finns förstås många utmaningar med självkörande fordon. En är att fortsätta utveckla tekniken tills vi får säkra och effektiva system. Detta är särskilt besvärligt när självkörande fordon ska dela gator med manuellt körda fordon och andra trafikanter, i så kallad mixad trafik. Vi människor följer inte alltid trafikreglerna, och tar ibland genvägar för att det ska flyta smidigare. Ska de självkörande fordonen programmeras att alltid följa reglerna? Det låter rimligt, men i vissa tester har det visat sig att det kan leda till problem när de ska samspas med manuellt körda bilar. Å andra sidan, om de inte ska följa trafikreglerna - vilka regler ska de då följa?

Det är många *lagar och regler* som påverkar utvecklingen. Vem har ansvar för körningen, och hur kan man garantera att en bil är säker i alla lägen? Detta är särskilt utmanande om systemet bygger på maskininlärning och man har "tränat" bilen med hjälp av data. Då är det svårt att veta hur bilen kommer att agera i nya situationer som inte fanns med i det dataset som användes för att träna den. De självkörande fordonen kan navigera tack vare sensorer och kameror som ser omgivningen. För att skydda människors integritet finns även regler för hur dessa får användas. Tekniken och möjligheterna utvecklas snabbt, och vi behöver en lagstiftning som kan ta tillvara på möjligheterna.

Eftersom de självkörande bilarna kommer att vara i närheten av människor, så finns en risk för olyckor. I samband med det kan det uppstå moraliska frågor – bilen kan behöva "välja" mellan att skada passageraren eller skada tio som står på trottoaren. Studier har visat att många av oss vill att alla andra ska ha bilar som är programmerade för samhällets bästa, men inte själva åka i en sådan bil⁹. Det är ett så kallat socialt dilemma, där en enskild person tjänar på att göra något så länge inte alla andra också gör så, för då förlorar alla på det. Eftersom självkörande bilar är något som påverkar hela samhället behöver beslutet fattas gemensamt. Sådana här komplexa frågor kan bromsa utvecklingen. Innan vi ger oss in i frågan om "vem ska bilen köra på", är det viktigt att påminna om att en stark drivkraft för utvecklingen av självkörande fordon är att minska olyckorna i trafiken totalt sett, eftersom självkörande fordon trots allt förväntas bli mycket säkrare än manuellt körda bilar.

Se referenser på sid 81.

Nya affärsmodeller baserat på artificiell intelligens

Erik Kruse och Patrik Regårdh arbetar med digitala marknadsstrategier på Ericsson

Det är i möjligheterna att omdefiniera marknadens behov som de riktigt stora förändringarna kommer. Nya tekniska framsteg, tillsammans med förändrade värderingar och livsstil påverkar varandra ömsesidigt och företagen måste hänga med för att inte konkurreras ut.

För några år sedan var Artificiell Intelligens (AI) mer eller mindre science fiction för gemene man, men nu 2018, är det en smygande realitet. Hittills har kanske AI varit "motorn" för att ta bort rutinarbete, analysera mönster i extremt stora datamängder och för att skapa röst-assistenten som Siri, Cortana och Alexa. Men vi ser fler och fler områden där AI smyger sig in, och antingen förstärker vår förmåga att ta beslut, eller tar bort människan ur ekvationen. Vi ser många applikationer implementeras i företag och banker som ersätter kundtjänsten med AI. Dessutom är AI ett verktyg för automation för att transformera finansiella processer. Sammantaget har allt detta möjliggjort betydligt lägre kostnader, bättre riskbedömning och kortare ledtider. Ytterligare ett exempel är Amazon Go som har visat att även dagligvaruhandel kan var helt automatiserad och mer eller mindre utan personal.

Men det slutar inte där. När AI går ut på marknaden och börjar analysera våra mänskliga beteenden och preferenser så ändrar det hur marknaden fungerar och agerar (och även politiken och därmed demokratin). Vi går då från en traditionell till en digital marknadsdynamik baserad på AI:s förmåga att

dra slutsatser och förutsäga individuella kunders preferenser och behov.

Digitala marknadsplatser såsom Amazon, Google, Microsoft Azure och Spotify är bra exempel på det värde som AI tillför. Genom att mer eller mindre i realtid analysera många kunders beteenden är man överlägset effektiv att förutspå behov och efterfrågan, samt skapa nytta som är svåråtkomliga i den analoga världen. Tydligast är detta naturligtvis i de tjänster som i grunden är definierade av digitaltekniken, till exempel sökning på nätet. Sökningarna ökar i precision inte bara genom att analysera innehållet i sig, utan också genom att analysera hur användarna tar till sig sökresultatet, hur de klickar sig vidare etc.

I grunden är det inte något nytt att de producenter som har bäst förmåga att bedöma marknadens behov vinner, men med AI kan det förfinas. Det är nu möjligt att i nära nog realtid förutsäga behov och dra långtgående slutsatser på individnivå. Netflix eller Spotify är några exempel från mediabranschen. Deras helt digitala plattformar levererar varianter på tidigare fysiska produkter eller analoga tjänster. Med AI kan tjänsterna kompletteras med individuella preferenser, personligt





designade spellistor och förslag på kringaktiviteter, som när Spotify tipsar om att artister som du lyssnar på kommer till stan. Digitaliseringen av tjänsten blir därför betydligt mer än en rationell distribution av det nya digitala formatet. Tjänsten utvidgas till att erbjuda en förstärkt, mer personlig upplevelse av musik eller film. Tack vare internets spridning har dessutom Spotify och andra digitala företag direkt en global bas för erbjudandet. Digitaliseringen har också en inbyggda fördelar för kunderna. Det är enklare att ha tillgång till strömmande

musik än att åka iväg till en affär och köpa en cd. Enkelheten, individanpassningen och bekvämligheten har ett stort värde för användaren.

De aktörer som är mest framgångsrika i att skapa insikter och erbjuda relevanta upplevelser utnyttjar ofta sin ställning till att öppna sina plattformar för komplementära erbjudanden, som i exemplet med Spotify och konserter. Det gäller att exponera mervärdet på ett smart sätt och att knyta nya intäktsströmmar till de relationer och transaktioner som man

medverkar i. När bilköpssajter erbjuder allt från bilbeställning till banklån, försäkring och service med leverans till garageuppfarten, så sker det på basis av analysförmågan och insikter på individuell nivå som styr bilaffären. Med dessa insikter som utgångspunkt och med smarta flersidiga affärsrelationer skapas nytta också för andra affärsverksamheter och därmed nya intäktsströmmar.

Bilhandelssajten är exempel på hur digitaliseringen driver förändring i traditionella branser. Precis som i taxibranschen och dagligvaruhandeln kvarstår transporten och varorna som den centrala kundnyttan, men med data och AI-centriska affärsmodeller har värdesystem utmanats i grunden. Det är ingen tvekan om att aktörer som Uber, med sitt affärsupplägg där passageraren direkt avropar en bil till önskad kvalitet och kostnad, har vunnit konsumenternas gillande och även fått stor påverkan på branschen. Übers datamodeller och algoritmer har, även om själva taxiresan kvarstår relativt oförändrad, medfört en omfattande rationalisering och värdeskapande ur konsumentens perspektiv. Lite förenklat kan sägas att kvar finns bilen, föraren och passageraren - allt övrigt har ersatts med en intelligent plattform på vilken kunderna mer självständigt kan göra sina val och följa upp sin leverans. På så sätt har genererandet av data blivit en naturlig del även för ett i grunden fysisk värdeskapande. Ett trovärdigt scenario är att Übers framtida verksamhet med självkörande taxi kommer ha algoritmer som kör dig, eller dina varor, helt automatiserat baserat på dina preferenser avseende tid, typ av bil, miljöpåverkan med mera. Därför finns det skäl att tro att Übers påverkan inte kommer begränsas till att skaka om taxibranschen utan även driva en förändring i transportsektorn som helhet. Det är troligt att självkörande bilar förändrar dagens mönster att köpa och äga en bil, till att ha tillgång till transporttjänster anpassade efter individuella behov.

Individualisering av erbjudanden

Många är de aktörer som investerat i olika sätt att nå sina kunder i digitala format, men som samtidigt behåller kärnprocesserna oförändrade. Resultatet är då ofta begränsat till effektiviseringar i försäljning och kundhantering utan något egentligt nytt värdeskapande. I en del fall är det till och med så att det som tidigare varit vissa företags största tillgångar har på grund av digitaliseringen blivit en belastning. Bankerna är ett bra exempel, innan bankappar och digitala möten var bankernas kontorsnätverk en stor konkurrensfördel, en viktig del i att hantera kunderna. Idag är det en kostnad eftersom vi sköter alla bankärenden själva via dator och mobiltelefon. Detta är svallvågorna av första vågen av digitalisering som i huvudsak riktat sig mot konsumentmarknaden. På samma sätt sker naturligtvis investeringar i digitalisering för att effektivisera de interna processerna, men även här med det väsentliga slutresultatet att skapa bättre varianter på en redan etablerad verksamhet.

För bara några år sedan hade företag flera öar av marknads- och försäljningsdata. De saknade insyn i den kompletta kundresan från start till mål. Avsaknaden av en gemensam bild av kunden tvärs över avdelningar, kanaler och kampanjer hindrade dem från att nå mål och önskad effektivitet i marknadsföringen. Idag har detta börjat förändras i de datadrivna företag som ser och hanterar data som en strategisk tillgång. Detta är dock inte tillräckligt, det är först när insikter och preferenser från kunderna direkt styr och kopplas in i affärsprocesserna för att skapa unika varianter till varje brukare/användare, som det är meningsfullt att tala om en fullt ut digital affärsmodell. En sådan modell innebär ofta en ompaketering av erbjudanden

Ett exempel kan vara livsmedelshandeln. Den förändras nu från detaljhandel, till matkassar som levereras till dörren, och

framöver skulle man med AI kunna ge personligt utformade måltidsförslag utifrån vad kunden beställt tidigare, ungefär som med Spotifys spellistor. Den leverantör som gör det bäst kommer få flest kunder.

Digitaliseringen förflyttar alltså värdeskapandet från produktion och successiva rationaliseringar i produktionsledet till konsumtion och förmågan att fånga och förutsäga brukarens eller användarens behov. Detta innebär inte att effektivitet i produktionsled är oväsentligt, tvärtom, det är en nödvändig förutsättning, men inte tillräckligt för att dominera marknaden. Om då industrialiseringen under årtionden drivits av storskalig produktion, så drivs digitala affärsmodeller av skalfördelar som sitter i en lätthet att nå en stor, om än fragmenterad, marknad kombinerat med en djupare förståelse av kund och brukare.

Nya digitala värdesystem

Just frågan hur värdesystem struktureras om är centralt, men ofta förbiset, när affärlivet successivt tillägnar sig mer och mer digitalteknik och AI-lösningar. Internt är det inte svårt att förstå att tekniken ger möjligheter till mer slimmade organisationer och bättre produktivitet medan den vidare frågan om vilka aktörer som elimineras och vad själva kärnan i erbjudandet består i vanligtvis är en allt för utmanande frågeställning. Konsekvensen riskerar att på sin höjd bli en bättre version av existerande affär medans andra mer vidsynta spelare ändrar affären i grunden. Nu ökar tillgängligheten av processorkraft och beräkningskapacitet samtidigt som den digitala världen utökas från att inbegripa information och människor till att också koppla upp prylar. Konsekvensen är att tillgången till digitaliserade affärsmodellers två grundkomponenter, data och analysförmåga fortsätter att öka exponentiellt och dessutom nå en allt större spridning.

Vi ser för närvarande hur digital teknologi inte bara ökar produktiviteten och effektiviteten i stora företag som levt på skalfördelar (economies of scale), utan istället börjar bryta ner institutionella barriärer och decentralisera makten från stora företag till slutanvändaren. Vi anar början av en ny, distribuerad, skalekonomi där det blir möjligt att bygga mer flexibla organisationsformer än vad traditionella företag erbjuder. Individer, innovatörer och små företag ges större utrymme att konkurrera med varandra - och med storföretagen. Individer kan arbeta, överföra pengar eller exponera sina tillgångar i lösa nätverksbaserade organisationer, utan den kostbara och tröga traditionella företagsstrukturen. Individer kan gå ihop för att förhandla till sig ett bättre erbjudande/pris, eller ta uppdrag i nätverk som existerar under en begränsad tid och för speciella syften. Man kan vara ganska säker på att det här kommer få stor påverkan på arbetslivet och samhället.

Vi måste välja

Dagens lagar och regler för affärlivet och även för oss som är medborgare är i stora stycken en produkt av industrialiseringen och baserat på traditionella värdekedjor i traditionella industristrukturer.

Nu förändras förutsättningarna så i grunden att risken är uppenbar att regler och lagar riskerar att motverka sina egna syften om de inte anpassas.

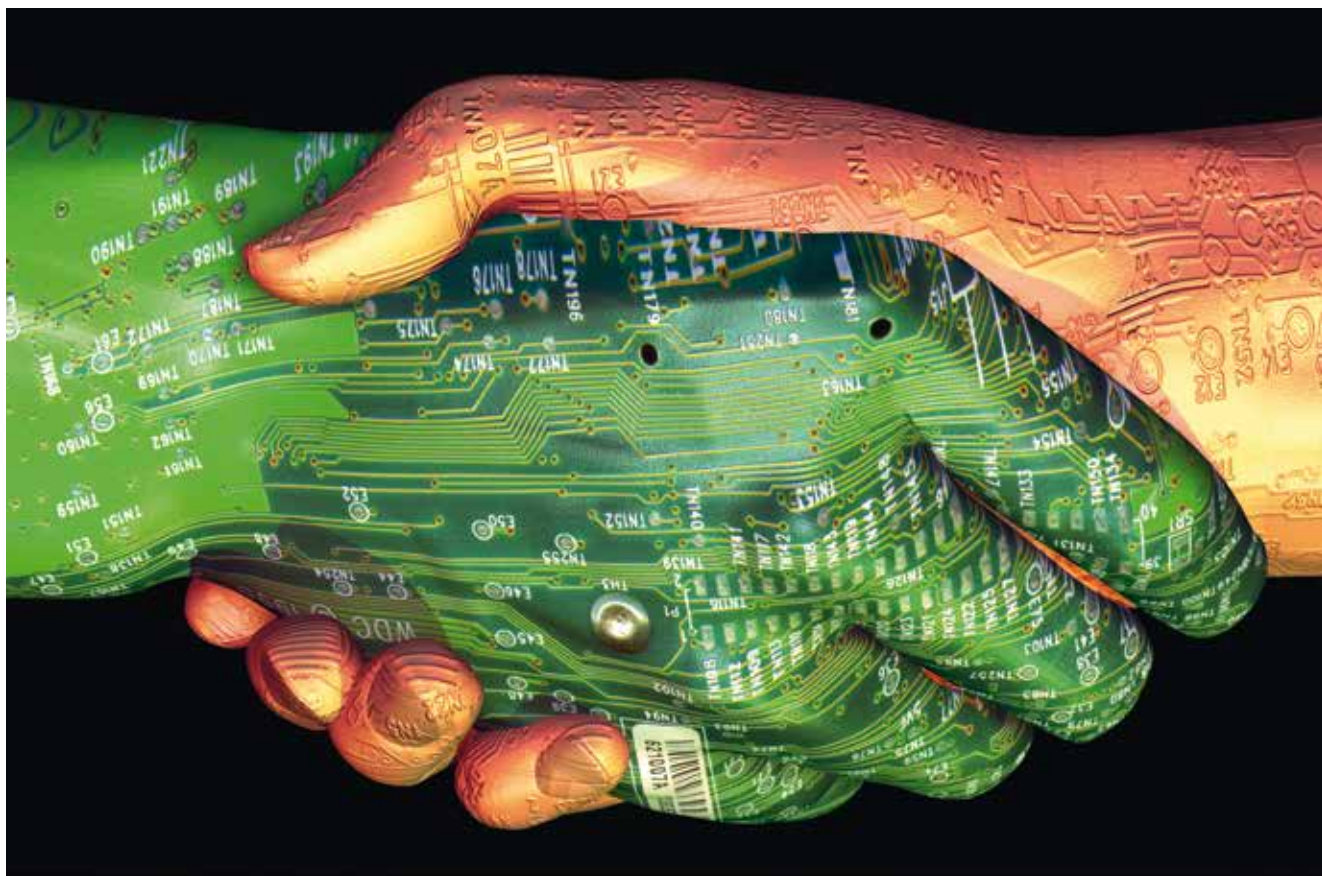
Politiker, andra beslutsfattare och även vi som individer behöver ta proaktiva beslut om vilken framtid vi vill ha. Ska vår konsumtion och vårt informationsutbyte baseras på fem, sex globala plattformar som sitter på all världens data? Eller vill vi ta makten över vår data? Kommer företag att tillåta att andra företag tjänar pengar på deras data, eller inte? Även om artificiell intelligens idag är mer eller mindre ett paraplybegrepp, ganska vagt och odefinierat, så måste beslutsfattare i företag

och offentlig sektor börja förbereda sig för en radikalt annorlunda verklighet baserad på AI.

Stephen Hawking kallade Artificiell intelligens, AI, för "ett av vår tids viktigaste samtal". AI:s påverkan på allt från våra relationer till hur vi arbetar kommer att vara omvälvande, kanske bortom vår vildaste fantasi.

Vår förhoppning är att vi kommer att använda AI för att förstärka människans förmåga att fatta beslut som är bra för

företagen samt vår planet. Fokuset måste flyttas från att ersätta människan till att förstärka människan, från att behålla den organisatoriska strukturen till en mer flexibel organisation förstärkt av AI. För en sak är säker - AI kommer att tvinga oss att göra saker annorlunda.



Fintech och robotrådgivning

Ann Grevelius är fintech-entreprenör och medgrundare till robotrådgivaren Opti

Bill Gates konstaterade redan 1994 att ”Banking is necessary, banks are not”. Så tidig han var, och så rätt han hade. Vi har behov av mängder av finansiella tjänster, både som konsumenter och företag. Vi behöver hjälp med betalningar, lån och investeringsrådgivning, men – behöver de komma från dagens finansiella aktörer och banker?

En bank är en aktör som förmedlar betalningar och möjliggör sparande och investeringar, driver verksamhet med betaltjänster och elektroniska pengar, kreditgivning till konsumenter, inlåning, valutaväxling och annan finansiell verksamhet. Området för finansiella tjänster är således brett, men frågan är hur och av vem de ska levereras till kunden. Behöver det vara banker i den mening vi ser dessa institutioner idag? Eller skulle det kunna vara helt nya aktörer? Hur, var och när vill vi som konsumenter använda finansiella tjänster?

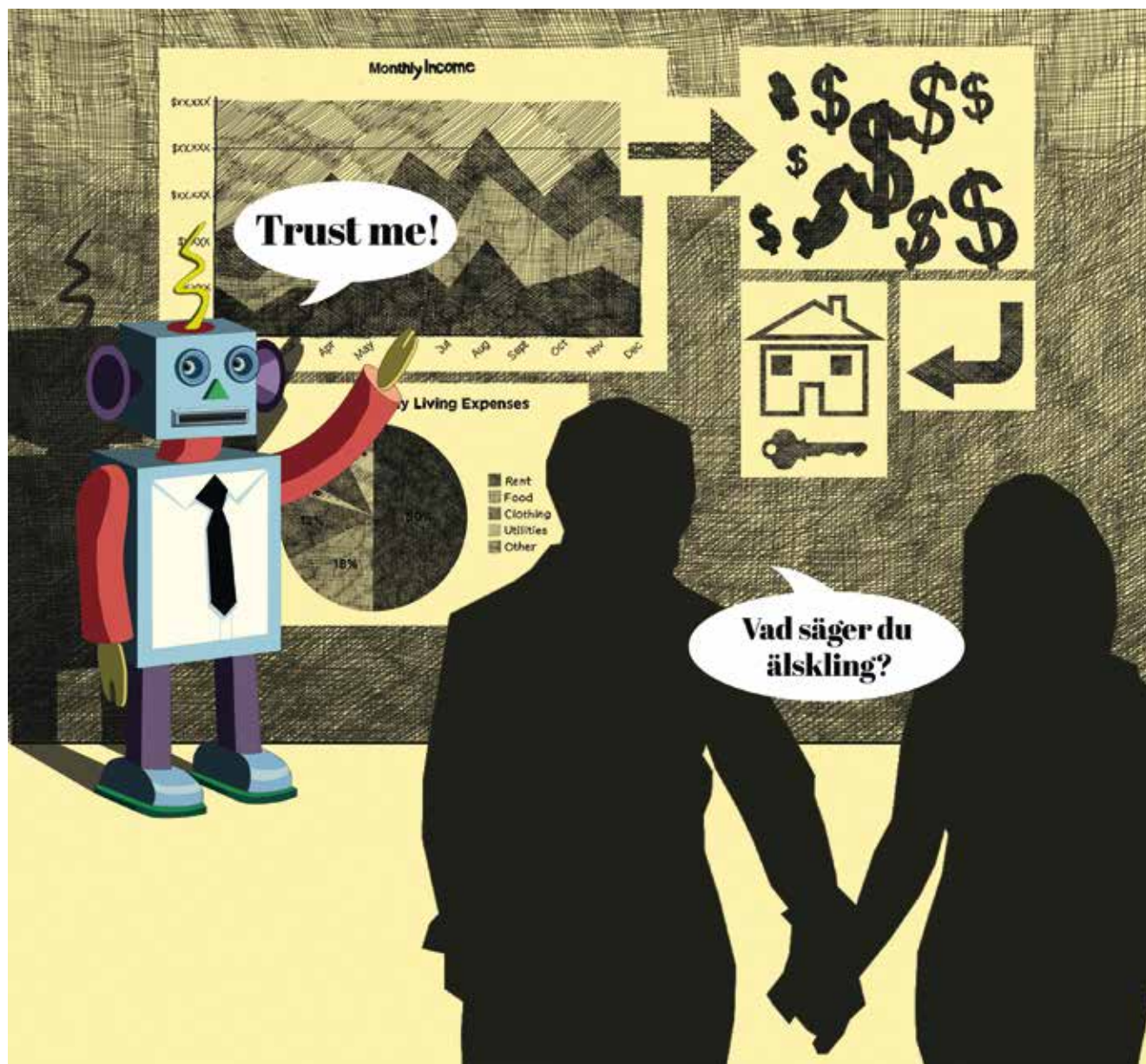
Runt omkring storbankerna växer idag upp ett helt ekosystem av nya aktörer, så kallade fintech-bolag, som både kompletterar och utmanar etablissemang. Detta samtidigt som bankernas IT-budgetar ökar och kunderna efterfrågar alltmer digitala och smarta tjänster. De nya aktörerna är ofta i grunden teknologibolag (tech) som levererar finansiella tjänster (fin), därav begreppet fintech.

Finansiella tjänster är på många sätt perfekta för teknisk disruption, alltså förändring, för digitala lösningar och stora vinster finns att hämta i AI och användandet av big data och deep learning. För vilka av dessa tjänster behöver egentligen utföras av en mänsklig hjärna? Vad skulle lika gärna, eller kanske till och med bättre, kunna utföras av en ”maskinhjär-

na”? AI refereras allt oftare till av finansiella aktörer när de beskriver sin vision för framtiden för den finansiella sektorn. Möjligheterna med ny intelligent teknologi är enorma och förväntas bidra till bland annat lägre kostnader, högre effektivitet, ökad kundlojalitet, nya och bättre tjänster samt ökad digital säkerhet.

Den dag vi har nått högre inom *artificiell generell intelligens*, kommer sannolikt finansbranschen att se helt annorlunda ut. Men på vägen dit kommer utvecklingen att vara mer stegvis och snarast adressera specifika områden, med snävare varianter av AI. Även här är dock applikationsområdena otaliga, då finansiella tjänster lämpar sig mycket väl för ett ”problem-för-problem-tänkande”. Det är enkelt att se enorma vinster med ökad användning av AI inom mängder av områden. Några av dessa är blockchain och kryptovalutor, kreditbedömning, robotrådgivning, betalningslösningar och effektiv kundanpassning och kommunikation.

Många nya aktörer vill idag framstå som ”AI-företag”, även om det i egentlig mening knappast handlar om ”artificiell intelligens” i strikt bemärkelse, utan snarare expertsystem. Hype-faktorn är mycket hög och man kan till och med se att bolag som profilerar sig med till exempel AI eller blockchain när en högre värdering. I själva verket handlar det



idag mer ofta om avancerade beräkningar och algoritmer, och i bästa fall någon variant av ”smal (eller snäv) AI”.

Drivkrafter för förändring

Drivkrafterna bakom den våg av förändring, nya tjänster och aktörer som sveper över oss idag är många. I huvudsak drivs utvecklingen av fundamentala faktorer såsom efterfrågan, utbud, timing, kommunikation, regelverk och teknologi.

I all förändring är kunderna och deras preferenser centrala. Den kanske viktigaste drivkraften är därmed ett förändrat kundbeteende. Vi förväntar oss idag att enkelt och snabbt kunna få existerande finansiella tjänster levererade via smarta, digitala, snabba media. Enkla paralleller kan exempelvis göras med detaljhandeln, där online-andelen stadigt ökar. Internetbank-lösningar möjliggör också en ökad digital kundservice, vilket i sig medför en förändringskraft kring de traditionella kontorsnäten.

Utifrån kundsidan ser vi även att yngre generationer, de så kallade millenials, helt och hållet ifrågasätter befintliga strukturer. För dem är inte banker som fenomen betydelsefullt, utan de vill ha enkla, smidiga lösningar, i mobilen och uppvisar ett högt förtroende för helt nya spelare. Behövs traditionella bankkonton för denna generation?

Utbudssidan driver också förändring. De stora aktörerna ser möjligheter till kostnadsbesparingar, men även till förbättrade och nya tjänster. Transparensen ökar vilket ger en tydligare bild av både utbud och priser och gynnar olika former av jämförelsetjänster, vilket är positivt för konsumenten. Vissa aktörer omfamnar också denna utveckling, medan andra av förklarliga skäl inte är lika positivt inställda. I en första våg har vi främst sett befintliga tjänster digitaliseras, medan vi nu börjar se mer av nyheter både på produkt och leverantörsidan, med en rad nya aktörer som kliver in på finansscenen.

Vi förväntar oss också att kunna utföra allt i mobilen. När vi idag i stort sett hanterar hela våra liv via mobilerna, så varför inte banktjänster? Ser vi till storbankernas kunder så sker idag all tillväxt för internetbankerna på den mobila plattformen. Och när vi i allt högre utsträckning blir en digital bankkund är steget till att testa nya aktörer inte längre så långt. Vi börjar idag se att gamla strukturer bryts upp, inom traditionella kärnområden som bolån och sparande.

Finansiella sektorn är i mångt och mycket präglad och påverkad av omfattande regelverk. Även denna utveckling har idag blivit en stark drivkraft till förändringsvindarna. De ökade kraven på transparens driver exempelvis fram en helt ny rapportering kring avgifter, vilket i sin tur gör konsumenterna mer medvetna både kring hur systemet faktiskt ser ut idag och får dem att ifrågasätta och söka nya alternativ. Andra nya regelverk lägger grunden för Open banking, vilket ger kunden ökad makt över den egna datan och öppnar upp för nya aktörer som nu kan ta del av kunddata som hittills bara funnits hos bankerna.

Ibland pekas hårda krav och strikta regelverk ut som hinder för fintech-bolagens möjligheter och det förs en stark debatt kring att exempelvis utveckla så kallade ”sandboxes”. Här kan unga bolag testa sina affärsidéer på ett tidigt stadium och få vägledning vad gäller regelverk och liknande. Detta kan givetvis vara positivt för att öka på takten i förändring, men samtidigt är regelverk centrala för en bransch som bygger på förtroende. Banker brukar sällan nå höga kundnöjdhetsstal, men – man litar på dem. Därför är det viktigt med regler även för nya aktörer och tillstånd från Finansinspektionen blir därmed även en kvalitetsstämpel för dessa.

I Sverige är vi av hävd snabba att ta till oss ny teknologi och pröva nya tjänster. Inom finansindustrin har vi under senare år sett två nya smått revolutionerande initiativ, som snabbt

erövrat svenskarnas förtroende, nämligen Swish och Mobilt BankID. Vid slutet av 2017 hade 7,6 miljoner svenskar Mobilt BankID och under året användes tjänsten 2,5 miljarder gånger. Detta är också centralt för att underlätta för många nya finansiella tjänster och en stark drivkraft för nya fintech-bolag.

Vad gäller nya former av kommunikation har ännu inte så mycket hänt. Visserligen har vi sett skiftet till mobilenheter och ökad digital kommunikation, men på traditionella områden som supportfunktioner dominerar fortfarande telefon, mail och vanliga chattar. Här finns stora förhoppningar på att AI ska kunna förändra detta i grunden och ett antal aktörer testar idag chatbots, som SEB:s och Nordnets Amelia, en digital kundservicemedarbetare. Det finns ett antal lågt hängande frukter kring vanliga frågeställningar med enkla svar och där kommunikationen kan ske i skrift. Vad gäller en intelligent konversation kring mer avancerade frågor lär det dröja innan Amelior fullt ut kan ta över kundservicefunktionerna hos bankerna. Potentialen är dock enorm, både vad gäller kundupplevelse och inte minst kostnadsbesparingar hos bankerna.

Så vilka är aktörerna och hur agerar de?

Precis som vilken traditionell industri som helst så domineras även finanssektorn av de stora, traditionella "analoga" aktörerna. De som genom århundraden ägt befolkningens bankmarknad och själva kunnat diktera ramar och villkor, inte sällan som i Sverige med en tydlig oligopol-struktur. Vad som nu händer, drivet av ökat teknologiinnehåll, liksom i alla branscher, är att dessa strukturer börjar utmanas och luckras upp.

Bankerna inser givetvis vikten av digitalisering och legitimerar den även genom enorma investeringar i uppgraderingar av befintlig IT-infrastruktur, och ny teknologi samt genom att på

olika sätt närma sig fintech-aktörer. Aktivitetsnivån har bara under senaste året intensifierats markant och idag tävlar storbankerna i flera nya grenar, bland annat i Open banking, samarbeten med fintechbolag och investeringar i unga utmanare.

Intressant är att dra paralleller till andra branscher där förändringen startade tidigare och har kommit längre, till exempel media och detaljhandel. Här har digitaliseringen skördat många offer och framtingat stora förändringar i affärsmodeller. Också förloppet är intressant att notera, i början händer inte så mycket, men när konsumenterna väl börjar anamma den nya tekniken och nya konsumtionsmönster etableras, ökar hastigheten i förändringsförloppet dramatiskt. Allt talar för att detta är en utveckling vi kommer att se även inom finansindustrin.

Bankväsendet förändras idag från grunden. Värdekedjorna bryts upp i mindre beståndsdelar och nya aktörer ger sig på utvalda, ofta höglönsamma, nischer. Att alla inte kan vara bäst på allt är uppenbart, men också att en liten aktör idag faktiskt kan nå stor spridning och ta anseende på marknadsandelar med hjälp av digital teknik. Detta öppnar upp för nya aktörer på den finansiella arenan. Små uppstickare, entreprenörer, fintechbolag men även större aktörer från andra industrier som har många användare och högt förtroende skulle kunna ge sig in på det finansiella området. Vad sägs om en Apple- eller Facebook-bank? När det handlar mer om data och AI-baserade tjänster, har våra analoga banker längre några komparativa fördelar? Regelverket kommer ju att garantera ramverket och så länge detta följs, så kanske lika gärna helt nya företag kan ge sig in i finansbranschen. Visserligen gör regelverk och myndigheter att inträdesbarriärerna är relativt höga, men det begränsar samtidigt konkurrensen. Och för riktigt stora aktörer är det ingen match att få de nödvändiga tillstånden.



Så är detta en trolig utveckling? Det finansiella området kanske inte är det allra mest attraktiva för de stora tech-bjäsarna, men möjligheten finns där.

Områden "up for disruption"

I stort sett alla finansiella tjänster har redan, eller kommer att, påverkas av ny teknik, digitalisering och i förlängningen artificiell intelligens. Vissa områden har kommit förhållandevis långt idag.

Tidigt ut har betalningsområdet varit, i Sverige med Klarna i spetsen. Ett bolag som startades av tre studenter på Handelshögskolan 2005 och som idag omsätter 4 miljarder kronor i 14 länder och har 1700 anställda. Den snabbt växande e-handeln har varit en bidragande orsak till efterfrågan även på mer dynamiska betalningstjänster, som Izettle.

Ett annat område, och mer baserat på avancerade algoritmer, är den elektroniska robohandeln på aktiemarknaden. Här återfinns två huvudsakliga varianter. Högfrekvenshandel bygger på mycket intensiv handel, ofta med små marginaler, helt baserat på förprogrammerade algoritmer med syfte att generera avkastning. Den andra varianten är exekveringsalgoritmer, som används av bankernas och fondbolagens tradingavdelningar och har fokus på "best execution" och effektiv exekvering av ordrar från fondförvaltarna. Ett antal arbetstillfällen har redan gått förlorade inom dessa områden, där de snabba algoritmerna utkonkurrerat många anställda inom tradingverksamheten.

Investeringsområdet har på exekveringsnivå kommit relativt långt digitalt drivet av främst nätbankerna Avanza och Nordnet, medan rådgivningsprocessen fram tills helt nyligen varit helt analog. Så kallade fondrobotar kan idag analysera alla fonder som vi som privatpersoner i Sverige kan investera i och rangordna dessa. Adderar vi sedan modern portföljteori

kan vi på denna bas skapa de mest avancerade och professionella sparportföljerna. Och via digital teknik kan vi nå ut till alla och envar. Man måste inte längre vara förmögen för att få den bästa investeringsrådgivningen. 2014 var Opti, som jag själv är medgrundare till, först ut att få tillstånd från Finansinspektionen för fullt digital, oberoende investeringsrådgivning, så kallad robotrådgivning. Här erbjuds traditionell professionell kapitalförvaltning digitalt via mobila applikationer till betydligt lägre priser än tidigare. Härigenom kan dessa tjänster alltså nå även småsparare. Dessutom attraheras millenials i hög utsträckning av mobila, smarta tjänster och har inte samma banklojaliteter som äldre generationer. När detta slår igenom på bred front kommer ytterligare ett antal arbetstillfällen att försvinna från finanssektorn.

Att kunna erbjuda kostnadseffektivt professionellt sparande är centralt även för samhället. Ett privat pensionssparande är avgörande för att det framtida pensionssystemet ska hålla. De barn som föds idag väntas fira sina 100:ade födelsedagar. Vi lever således längre och även om politikerna successivt höjer pensionsåldern, så går ekvationen inte riktigt ihop. När folkpensionen infördes år 1913 var pensionsåldern 67 år och medellivslängden cirka 60 år. Rent krasst skulle alltså merparten av alla förmånstagare aldrig i någon större utsträckning nyttja systemet. Det som ytterligare komplicerar situationen är att den förväntade avkastningen framåt idag tyvärr ser ut att hamna på lägre nivåer än vad vi vant oss vid. Teknologi som kan bidra till att minska kostnaderna i pensionssystemet är därmed ytterst samhällsviktiga.

Kryptovaluta – inte bara för kriminella

Blockchain är en digital, decentraliserad publik loggbok där till exempel kryptovaluta-transaktioner registreras. Tekniken kopplas ofta just till *kryptovalutor*, men det finns många andra

intressanta användningsområden, även om det knappast är troligt att någon nationalstat i en nära framtid kommer att anamma en kryptovaluta som sin. Som exempel finns idag potential till att digitalisera löpande skuldebrev, vilket är en viktig kugge i processen för att digitalisera den så viktiga bolåneaffären för bankerna, clearing och för att verifiera avtal och kontrakt. Tekniken kan även användas vid samhällliga processer såsom publika val och många andra liknande situationer.

Det finns idag otaliga kryptovalutor, såsom Bitcoin och Ethereum, och diskussionerna går varma ifall dessa är att betrakta som "valutor" eller mer av en spekulativ investeringstillgång. Bitcoin kan betraktas som ett tillgångsslag om tillräckligt många är villiga att vara uthålliga aktörer på dess marknad. Som sådant påminner Bitcoin mycket om en äldre fysisk motsvarighet, nämligen guld.

Andra ser det som ett instrument för betalningar av varor och tjänster i den mörka, illegala delen av samhället. Den senaste tidens mycket kraftiga kursrörelser visar om inte annat att marknaden fortfarande är i sin linda och att ett långsiktigt jämviktsvärde knappast stabiliserats. Intressanta användningsområden saknas dock inte, exempelvis att kunna skicka betalningar världen över utan alla de kostnadsdrivande mellanled som finns idag. Detta skulle gynna medborgare i utvecklingsländer, och kunna få positiva samhällsekonomiska effekter.

Även inom kreditområdet finns stora möjligheter för artificiell intelligens. Idag är kreditprocesserna många gånger summariska och baserade på statiska data som inkomst, betalningshistorik och traditionell kreditvärdighet. Med dynamisk teknik skulle dessa processer kunna bli mycket mer finmaskiga och anpassade till varje individs ekonomi och därmed bland annat bli billigare för kunder med hög kre-

ditvärdighet. Så kallad *credit scoring* kan även användas inom olika typer av konsumentkrediter, och även för företagskunder baserade på resultat- och balansräkningar samt prognoser baserade på betalningshistorik. Även inom försäkringsområdet finns motsvarande potential att kundanpassa premier och försäkringslösningar, med potentiellt stor konsumentnytta och även bättre riskhantering för bolagen.

Alla tjänster, såväl gamla som nya, kräver någon form av kommunikation. Idag sker denna oftast via telefon eller chat-tjänster, som har avgränsade öppettider och som bygger på att det sitter människor och svarar på inkommande frågor. Kostsamt och ineffektivt för både kund och företag. Blickar vi framåt kommer *chatbottar* att ta över en stor del av denna verksamhet. Förhoppningsvis kan vi då också få lite mer av pro-aktiv kommunikation, inte bara reaktiv respons på uppkomna problem. När chatbottar brett ut sig, hur många bankkontor kommer vi då att ha kvar? Hur många anställda behöver bankerna ha?

Vad händer framöver?

Vad kommer närmast? Vilka är nästa banktjänster att utsättas för konkurrens från nya aktörer? I vissa fall innebär ökad användning av teknik och AI en ökad konkurrens inom befintliga områden, medan det i andra fall handlar om att helt nya produkter och tjänster utvecklas. Kanske har vi snart digitala personliga ekonomi-coacher på samma sätt som vi tränar vår fysik med personliga tränare? Kanske kan dessa AI-baserade specialister analysera alla våra transaktioner, vårt sparande och våra lån och komma med proaktiva, oberoende råd om var våra sparkonton bör ligga och hur vårt pensions-sparande bör se ut, liksom även tjänste- och premiepensioner? Potentialen i att från grunden utmana nuvarande strukturer är enorma och lär påverka de aktörer som idag dominerar



marknaden. Antingen anpassar de sig eller så kommer uppstickarna att ta märkbara marknadsandelar.

Kommer AI att ta över finansbranschen helt? Vår mänskliga hjärna kommer att ha svårt att hävda sig vad gäller beräkningar, analys av stora mängder data etc, men en plats för människan lär alltid finnas även inom den finansiella sektorn. Även om antalet människor som arbetar inom banksektorn kommer att minska och tjänsterna förändras rejält.

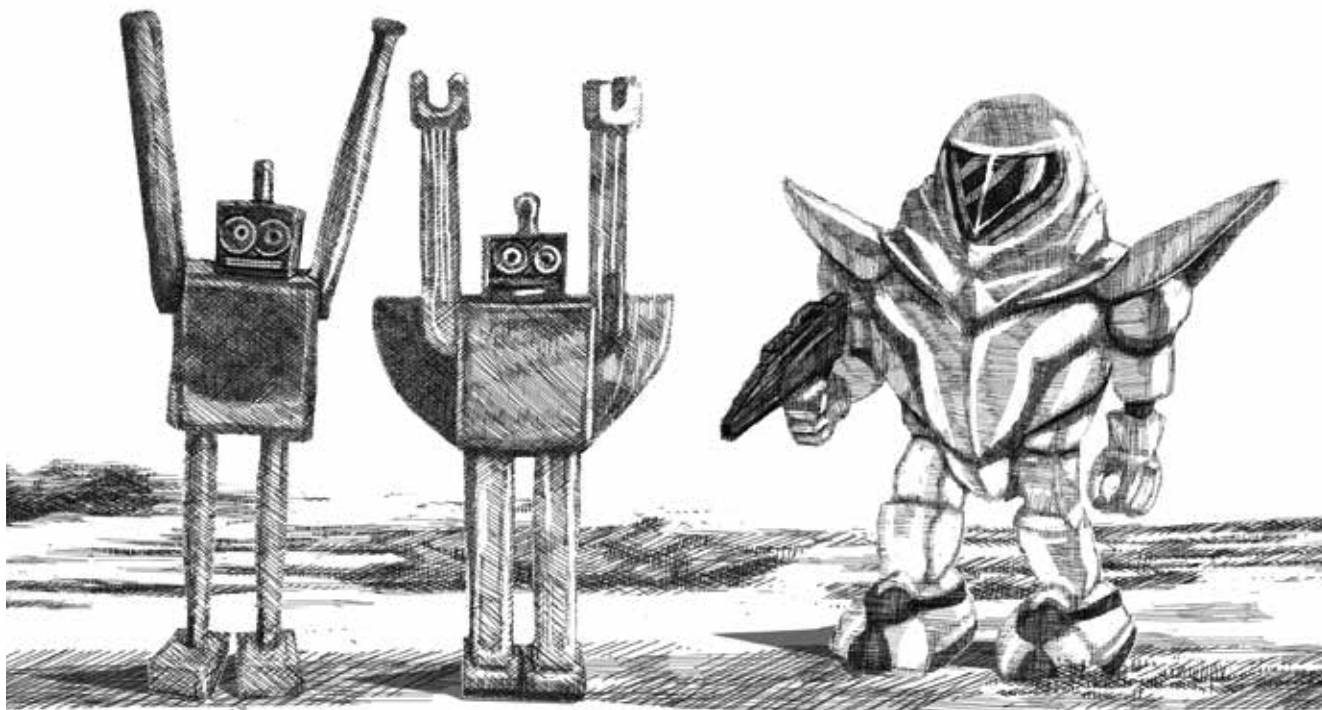
Morgondagens konsumenter, våra ungdomar, lär inte konsumera banktjänster som tidigare generationer gjort. Vad är egentligen en bank? Är det en fysisk byggnad vid det lokala stortorget? Nja, millenials som inte har någon relation till dessa fysiska banker kan nog komma att prioritera tillgänglighet i mobilen, attraktiva kundupplevelser, personligt anpassade tjänster och de vill inte betala mer än nödvändigt. Finansbranschen har omvälvande år framför sig och mycket talar för att kartan kommer att se väsentligt annorlunda ut om ett par år. Banker kanske då bara är meta-tjänster, en

samling olika appar som alla optimerar respektive specialisttjänst. För egentligen, vad talar för att ett finansiellt företag verkligen är bäst i alla grenar? Och ifall jag som konsument kan välja spjutspetsaktörer i respektive gren, varför skulle jag inte göra det? Bankernas ideér om helkundskonceptet är nog av utdöende karaktär. Och därmed en del av de höglönsamma verksamheterna. Mer troligt är att bankerna måste förändras i grunden och kanske börja hjälpa sina kunder med ett bredare spektrum av tjänster, kanske att välja el- och teleabonnemang vid sidan av försäkringar och andra kringtjänster. Med artificiell intelligens kommer branscher att brytas isär och flyta ihop, i nya konstellationer och med ett större fokus på vad kunderna faktiskt önskar.

Så, vi lär fortsatt ha behov av alla dessa finansiella tjänster, men kanske från nya aktörer och ibland i form av helt nya tjänster. "Banking is necessary, banks are not" gäller även för framtiden.

Hur väl står sig Sverige?

Mats Nordlund är teknologie doktor och ansvarig för speciella projekt och forskning på Zenuity AB



De flesta stater och företag har idag identifierat Artificiell Intelligens (AI) som ett område av högsta strategiska värde. Flera länder har tagit fram ambitiösa nationella strategier för AI, några exempel på nationella mål är:

- **Kina**¹: Accelerera uppbyggnaden av en innovativ nation och bli en global makt inom vetenskap och teknik.
- **USA**²: Vara världsledande och ta fram ny AI-kunskap och teknik som ger en mängd fördelar till samhället, samtidigt som negativa effekter av AI minimeras.
- **Ryssland (Putin)**³: "AI är framtiden, inte bara för Ryssland, utan för hela mänskligheten. Det kommer med kolossala möjligheter men också hot som är svåra att förutse. Den som blir ledande inom detta område kommer styra världen."

- Finland: Finland ska vara en ledande AI-nation och med detta menar man inte minst att bli ledande i att använda AI. Man har också tagit fram en 8-punktsplan för att nå detta mål.
- EU⁴: Håller på att arbeta fram en AI strategi och första materialet släpptes under våren 2018.

Stormakternas ansatser inom AI visar att sannolikt kommer den globala maktkampen att förändras från att ha haft fokus på vapenkapprustning till fokus på AI-kapprustning. Det ger möjligheter för tillväxtekonomier, som Indien och Brasilien, och tillväxtföretag att utmana etablerade stormakter och storföretag, med följderna att maktbalanser förskjuts.

I denna kapprustning ligger även en självförstärkande spiral: Den som har bästa AI kan utveckla ännu bättre AI, osv. – Vinnaren tar allt...

Denna "maktkamp" gäller även näringslivet där det finns motsvarande ambitioner att dominera sina respektive marknader. Industrin kommer alltså också att satsa hårt på AI.

Sverige

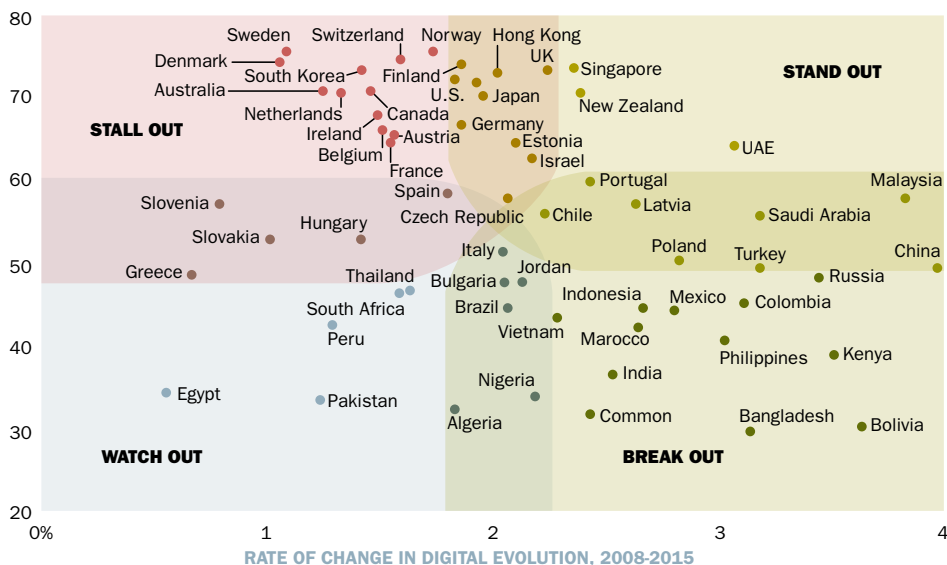
Sverige har idag ingen nationell AI strategi, men vi har ett digitaliseringsmål som säger att vi ska "vara bäst i världen på att använda digitaliseringens möjligheter".

En kartläggning av olika länders digitalisering publicerades i Harvard Business Review i juli 2017⁵. Denna visar hur

Plotting the Digital Evolution Index, 2017

Where the digital economy is moving the fastest, and where it's in trouble.

HOW COUNTRIES SCORED ACROSS FOUR DRIVERS ON THE DIGITAL EVOLUTION INDEX (OUT OF 100)



FRITT TOLKAT FRÅN SOURCE DIGITAL EVOLUTION INDEX 2017, THE FLETCHER SCHOOL AT TUFTS UNIVERSITY AND MASTERCARD © HBR.ORG

långt länderna kommit i sin digitalisering (y-axeln) och hur fort de utvecklas (x-axeln).

Sverige, tillsammans med andra nordiska länder, Schweiz, USA, Singapore och Sydkorea hör till de länder som kommit längst (ligger högst på y-axeln). Vilket visar att vi idag kan sägas nå det nationella digitaliseringsmålet.

Bilden visar också att utvecklingstakten av digitaliseringen i Sverige är internationellt sett är mycket låg (vi ligger till vänster på x-axeln), och vi kategoriseras som ett "avstannade" (stall out) land. Det betyder att vi riskerar att missa målet att vara bäst i världen på att använda digitaliseringens möjligheter.

För att avstannande länder ska "komma loss" pekar författarna på att det krävs en samlad ansträngning för att förnya sig. Det behövs en satsning på en digital teknologi (till exempel AI) där man kan vara ledande. De hinder mot innovation som finns ska bort. Avstannande länder kan lära av "stand out" länder, som Singapore hur man håller igång förändringstakten. De avstannade kan också utnyttja sin mognad, styrka och nätverk för att förnya och fortsätta sin utveckling.

För att behålla vår konkurrenskraft och försvara demokratin måste Sverige välja AI som en digital nyckelteknologi att satsa på.

Utifrån detta mål har jag gjort en sammanställning av de styrkor och svagheter som finns inneboende i Sverige, samt de möjligheter och hot som finns i vår omvärld och kan hjälpa eller stjälpa våra ansträngningar att nå detta mål.

Sverige har idag många styrkor som kan hjälpa utvecklingen av AI. Dessa styrkor omfattar bland annat bra och unika data-set, starka och ledande industrier och myndigheter med god teknikmognad och ambition. Samhället är uppkopplat med stor mobiltäckning och god infrastruktur för kommunikation samt har starka finansörer med möjlighet att satsa på AI (stiftelser, staten och näringsliv). Vi har också en bra nystarts-kultur.

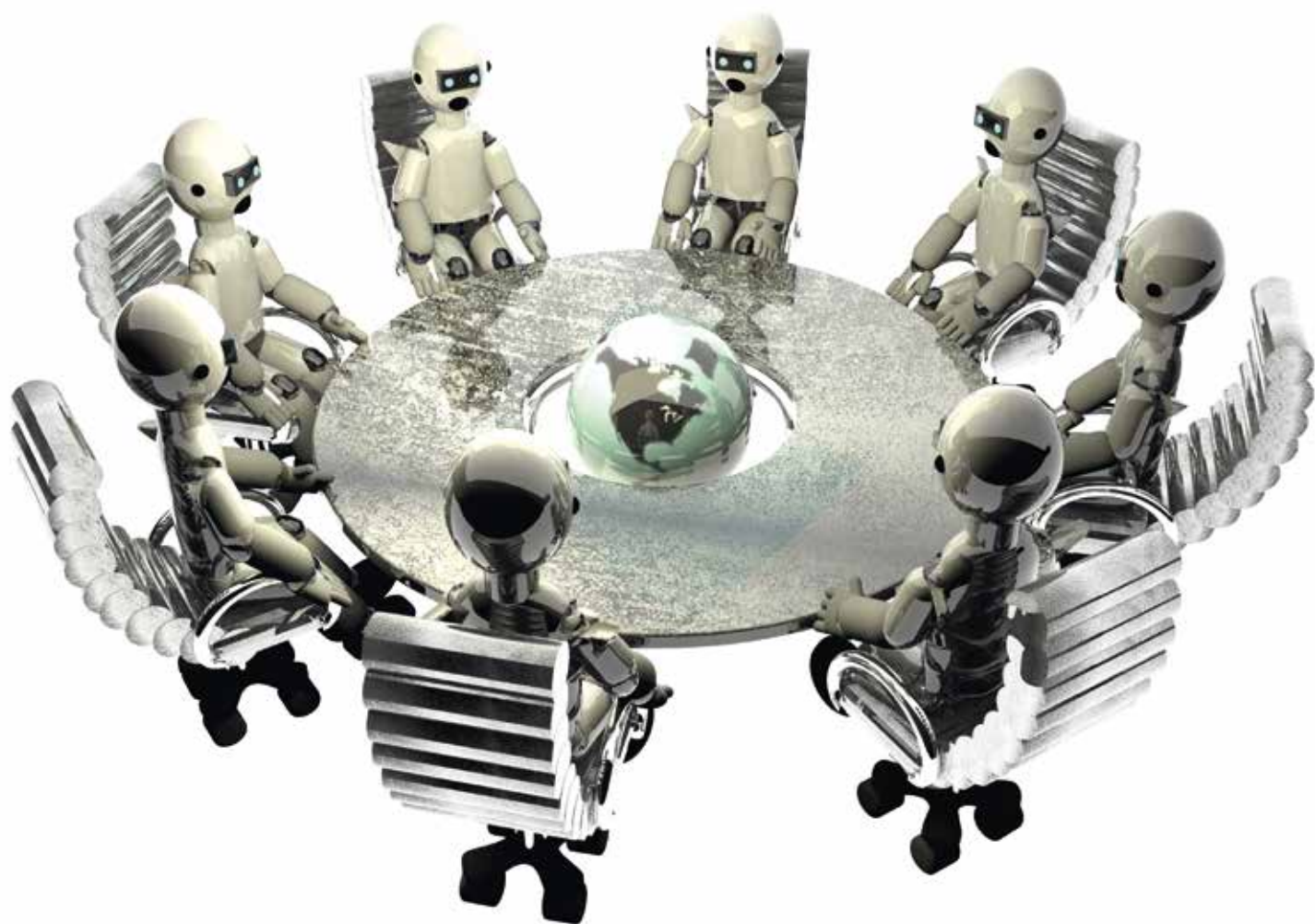
Å andra sidan har vi svagheter vilka måste hanteras för att hindra utvecklingen. Exempel är brist på tillgång till AI kompetens – ett problem vi delar med de flesta länder. Våra universitet är lågt rankade inom AI och är inte väl positionerade i internationella forskningsnätverk. Vi har för få industrier som är starka inom AI (motsvarande Google, Facebook, Amazon, Baidu, och Alibaba). Det är relativt svag tillgång till riskkapital utanför Stockholm, och lärarkåren på grund- och gymnasieskolor har ganska låg IT utbildning.

Att behålla våra talanger i landet är en framgångsfaktor som konstant hotas. Många försöker rekrytera dem till andra länder, det kan vara till kända universitet eller företag. Det finns också en risk att svenska företag lägger utvecklingen av AI utomlands på grund av brist på tillgänglig kompetens.

Några strategier för att stärka vår position i den internationella konkurrensen kan vara att utnyttja de företagsband till Kina som finns för att knyta närmare kontakter till deras forskning. Ytterligare strategier är att utnyttja den osäkerhet som den nya amerikanska administrationen och Brexit skapar för att locka talanger till Sverige. Flera stora amerikanska företag har byggt upp datacenter i Sverige och det kan finnas möjligheter att dessa företag kan lägga viss AI-utveckling i anslutning till sina datacenter och universitet här. Dessutom kan vi utnyttja att Sverige och Skandinavien anses vara mycket snabbare när det gäller att driva förändringar än till exempel Tyskland och Frankrike. Etablering av ett nationellt kompetenscentrum för AI skulle också vara en styrka som kan matcha liknande satsningar utomlands.

Framtidsspaning

Intåget av artificiell intelligens är ett lika stort steg som maskinernas intåg under den industriella revolutionen. Maskinkraft ersatte muskelkraft, speciellt vad gäller repetitiva och



tunga arbetsuppgifter. AI har potentialen att ersätta "hjärnkraft".

På samma sätt som skedde under den industriella revolutionen så kommer en mängd jobb att försvinna och tas över av AI. Idag är det ingen i Sverige som slår våra fält, bryter vår malm eller tröskar vår säd för hand. De personer som tidigare hade den typen av uppgifter har frigjorts för andra och nya jobb. Det handlar om allt från nöjesindustri, service och tjänster till avancerat forsknings- och utvecklingsarbete. I framtiden kommer vi inte att se personer arbeta med bildtolkning, granskning av blanketter, diagnostik, styrning av maskiner, fordon och liknande.

Den industriella utvecklingen har givit oss kunder/brukare nya lösningar:

hitta rätt lösning för det vi önskar, diagnostiklösningar som kan förutse kommande problem och sätta in förebyggande åtgärder, nya typer av datorspel och robotar i vården. Allt detta kan sedan förstärkas i avancerade användargränssnitt som bygger på förstärkt verklighet (augmented reality) eller virtuell verklighet (virtual reality).

Med intåget av AI kommer också en mängd risker. Vad kan hända om vi släpper kontrollen om AI och inte längre förstår hur det fungerar? Hur hindrar vi att AI används för att till exempel förtrycka en befolkning, mobba en kollega, eller generera fake news? Om vi gör oss beroende av AI, vad händer med viktiga samhällsfunktioner vid strömavbrott eller kommunikationsfel? Kommer vi att förlora viktig kunskap i befolkningen om den bara finns i AI systemen?

Tidsperiod	Värdesatt förmåga	Produktkaraktäristik	Produktionsteknik
<1800-talet	Jobba med händerna	Unika lösningar	Hantverk
1800-talet – 2015	Jobba med maskiner	Unika lösningar Massproduktion Masskundanpassning	Hantverk med stöd av maskiner Produkter anpassade till maskiner Modulära produkter
2015	Jobba med människor	Massindividualisering	AI-baserade produkter med mänskligt stöd

AI kommer att ge oss produkter och tjänster som blir mer och mer individanpassade. Algoritmer kommer kunna identifiera de unika behoven som finns och föreslå, eller sätta in lösningar, som är anpassade till den specifika situationen eller individen.

Exempel på existerande AI-lösningar är digitala assistenter som Alexa och Siri, chatbots på webbsidor som hjälper oss

Det finns också en risk att det uppstår en AI-överklass som har tillgång till och använder AI-lösningar, och en AI-underklass som inte hänger med.

Tittar man framåt kommer AI leda till att kunskaper som idag värdesätts högt kommer att sjunka i värde medan andra kunskaper och förmågor kommer att värderas högre. Läkare kan till exempel få stöd med bland annat tolkning av rönt-

genbilder, EKG-kurvor, och EEG (elektronisk avläsning av hjärnan). Algoritmerna tränas på miljontals bilder och kurvor så att de har förmågan att upptäcka även mycket ovanliga diagnoser. Tolkningen går fort, algoritmen kan även föreslå behandlingsåtgärder. Läkaren kan då fokusera på att möta patienten, skapa en förtroendeingivande och positiv miljö, och sedan arbeta med behandling, uppföljning och rehabilitering för att snabbare få patienten bättre, eller bromsa sjukdomsförloppet. För chaufförer kan man tänka sig en lösning där långfärder på landsväg eller motorväg överläts åt AI-lösningar medan körning i mer komplicerade stadsmiljöer görs mer manuellt. Detta ökar säkerhet, minskar utsläpp och gör det möjligt för förarna att huvudsakligen vistas i sin hemmiljö där de kan möta kunder för att lämna och hämta material och varor.

Utifrån dessa exempel ser vi att rena faktakunskaper kommer värderas lägre jämfört med idag, och förmågan att positivt möta människor värderas högre. En typisk arbetssituation inom många yrken kommer att vara att utifrån resultatet av en AI-bearbetning kunna förklara olika situationer, sammanhang och åtgärder.

Generellt kan vi säga att de flesta yrkesarbetande kommer att utnyttja algoritmer, kvalitetssäkra resultatet från dessa och sedan fokusera sitt arbete på sina kunder/brukare/patienter för att ge dem en helhetsupplevelse med höjd livskvalitet och effektivare utnyttjande av resurser (hållbarhet).

Vägen framåt

Utifrån denna artikel följer ett antal viktiga frågor att diskutera och fundera på för dem som är aktiva inom AI-forskning och utbildning i Sverige

- Vad för kunskap och förmågor ska vi skapa?
- Vem ska skapa och förmedla denna kunskap?
- Vilka kunskaper och förmågor ska förmedlas och till vem?
- Hur ska dessa skapas och förmedlas?
- Hur ser vi till att Sverige blir ett attraktivt land för topp-talanger inom AI att komma till och arbeta i?

Vi bör också diskutera om vi redan idag utbildar för framtidens behov. Som exempel, vad kommer följande yrkesgrupper göra i ett AI samhälle? Vad ingår i deras utbildning idag?

- Finansrådgivare
- Revisorer
- Röntgenläkare
- Signalbehandlingsingenjörer
- Översättare

Det finns många nyckelord som man kan ha som utgångspunkt för en diskussion och debatt kring AI, hur utbildning, forskning och olika yrken kan komma att utvecklas. Det handlar om etik, integritet, hållbarhet, risker, omställning av yrkesverksamma, automatisering, skatt, "mjuka förmågor", och livskvalitet.

Se referenser på sid 81.

Livet med AI i skolan

Mats Nordlund är teknologie doktor och ansvarig för speciella projekt och forskning på Zenuity AB

Utbildning lägger grunden för välbefindandet i vårt land. Nya verktyg baserade på artificiell intelligens (AI) kommer att ge helt nya möjligheter inom utbildningsområdet. För att Sverige ska lyckas dra nytta av detta är det viktigt att lärarkåren engageras tidigt, och att lagstiftning, regleringar och riktlinjer anpassas så att AI-baserade verktyg kan införas i utbildningen på ett bra sätt.

A I-baserade verktyg kan skapa stora värden då de används för administrativa och repetitiva uppgifter - sådant som idag utförs av kvalificerade personer, där många parametrar ska analyseras och sedan ligga till grund för beslut och planering. Verktygen frigör tid så att de som tidigare haft dessa uppgifter i stället kan möta och arbeta med de individer som påverkas. Ett nyligen publicerat exempel på det här är Trelleborgs kommun som satt in en AI-baserad lösning för att analysera socialbidragsansökningar, med resultat att de nu bedöms mer lika (inget beroende av individuella tjänstemän), och mycket snabbare. Samtidigt har de tjänstemän som tidigare granskat ansökningar getts möjlighet att arbeta med sina brukare och hjälpa dem in i arbete i kommunen. Följden är att det nu är tre gånger så många personer som kommit ut i arbete på ett år jämfört med tidigare. Även på Gotland pågår ett liknande projekt inom digital samhällsbyggnad.¹

Inom utbildningsområdet kan vi förvänta oss liknande utveckling som i Trelleborg och vi ser redan vissa verktyg som tagits fram. På samma sätt som där, så kommer inte AI-baserade verktyg att ersätta lärare, utan bli ett viktigt stöd genom att avlasta dem med fler och fler administrativa uppgifter som idag ligger utanför lektionstid.

I en inte så avlägsen framtid kan man tänka sig att en elev lämnar in räkneuppgifter eller en språkuppgift till läraren. Uppgiften skannas (om den lämnats på papper) av läraren och rättas och analyseras därefter på några sekunder av ett AI-baserat verktyg. Verktyget, som har elevens historik och nu de senaste resultaten, tar direkt fram en individuell plan för hur eleven ska arbeta vidare. Denna presenteras till läraren i form av lämplig pedagogik, exempel, underlag, samt nya uppgifter som möter eleven precis där hen står och som passar elevens egen inlärningsstil. Läraren kan sedan ta detta underlag och planera sitt eget arbete med eleven för nästa tillfälle.

Hur kommer AI-baserade verktyg för utbildning fram?

Utveckling av AI-baserade verktyg består av två delar: dels att skapa de algoritmer som ska utföra arbetet och dels att träna algoritmerna med hjälp av stora mängder data. Att få fram bra data är nyckeln till hur bra verktygen kommer att bli.

Inom vissa ämnesområden, till exempel språk, kommer det att behövas svenska data för att träna algoritmerna, inom andra, som matematik, kan man troligen, åtminstone delvis, använda verktyg som tränats med data från andra länder. Vi får räkna med att programvarorna utvecklas av både svenska och internationella företag och institutioner.

Utveckling av algoritmer och att få fram och förbereda (annotera) data kräver både forskning och en stor kunskap hos dem som ska utföra arbetet. Mest troligt är att de kommande verktygen bygger på en eller flera av följande "AI-teknologier": naturlig språkbehandling, tal- och röstigenkänning, samt maskininläring. Exempel på annoterade data som behövs i träningen av algoritmer är översatta meningar, matematikproblem och olika lösningar till dessa, problem inom naturvetenskap och lösningar på dessa, individuella studieplaner för elever med olika utgångspunkter, och så vidare.

För att få fram bra AI-baserade verktyg är det också mycket viktigt att lärarkåren engagerar sig tidigt i processen och arbetar tillsammans med programmerarna. De centrala områdena för lärarnas professionella yrkeskunskap är att:

- kravställa verktygen så att dessa blir ett bra stöd och uppfattas som intuitiva och lättanvända.
- annotera data som kan användas för att träna algoritmer så att verktygen får en hög kvalitet. Detta kan bara göras av erfarna lärare, på samma sätt som det bara är erfarna läkare som kan annotera data som kommer från medicin och sjukvård.
- samarbeta med andra intressenter, som myndigheter, skolledning, politiker och andra för att få fram bra reglering, planer, arbetssätt och fortbildning som gör det möjligt att utnyttja de nya fördelar som AI kan ge samtidigt som riskerna med den nya tekniken kan arbetas bort.

Vilka är riskerna med AI-baserade verktyg inom utbildning?

Den största risken med att införa AI-baserade verktyg i utbildningen är att garantera datasäkerhet och personlig integri-

tet samt uppfylla kraven på att stå emot olika intrångs- eller sabotageangrepp. Det är speciellt viktigt eftersom verktygen kommer att hantera stora mängder personliga och känsliga data om elever i alla åldrar.

Ett exempel på denna problematik är hur virtuella assistenter, till exempel Siri eller Alexa, kan användas i utbildningen. Virtuella assistenter kan erbjuda bra inlärningsmöjligheter, men lagarna kring vilka data dessa assistenter samlar in, lagrar och använder datan är inte fullt utvecklade även om det nya europeiska datalagringsdirektivet GDPR har stor relevans.

Om inte alla lärare samtidigt har tillgång till, och förmåga att använda, de system som kommer att bli tillgängliga finns en stor risk är att det skapas en ojämlikhet mellan kommuner, skolor, och klasser inom en viss skola. Här kommer det att ställas stora krav dels i investeringar i verktygen och dels i fort- och vidareutbildning av yrkesverksamma lärare. Lärarutbildningen på alla universitet och högskolor måste också anpassas så att nyexaminerade har nödvändiga kunskaper och färdigheter. Detta i sin tur innebär att föreläsare och examinatorer inom högskolan behöver fortbildas för att kunna utveckla nya kurser och examinationskrav.

Vad händer internationellt?

I ett växande antal länder och inom utbildningsföretag förs nu diskussioner kring de möjligheter AI-tekniken kan ge utbildningssystemet. Ett sådant exempel är en omfattande rapport av utbildningsföretaget Pearson² från 2016. Rapporten som tagits fram tillsammans med forskare vid University College London presenterar visionen av ett system för lärande som kan följa med och stödja varje person genom hela livet. Systemet anpassar inlärningsupplevelsen i realtid till varje persons behov och förutsättningar.

Nedan sammanfattas nyligen publicerat material³ kring en del av de tankar som cirkulerar internationellt.

Utbildningssystemet har följande huvudutmaningar:

- att tillhandahålla kvalitativ utbildning, ofta i stor skala.
- att säkerställa att likvärdig utbildning av hög kvalitet är tillgänglig för alla, oavsett bakgrund, bostadsort och eventuella särskilda behov.
- att minska kostnaderna för utbildning.

Vi har idag mycket bra förutsättningar för att föra in nya verktyg. Den nya generationen elever är digitaliserade och tar snabbt till sig ny teknik. Stora fördelar med AI kan också komma på plats ganska snabbt. Verktygen kan avlasta lärarna med en hel del administration, bedömning och rättning, samt planering. De nya inlärningsmöjligheterna AI-baserade verktyg kan generera kommer att bestå av mer individanpassad lärarhandledning och undervisning till varje elev. Detta bör kunna skapa betydande fördelar för stora klasser med elever som ligger på olika nivå, har olika behov, och olika bakgrund. Den individanpassade undervisningen kan ytterligare förstärkas med hjälp av virtuella assistenter som möter och kommunicerar med eleverna på deras nivå. De virtuella assistenterna kan också flagga direkt till läraren om någon elev behöver mer personlig handledning och hur denna bäst kan utformas.

Man diskuterar också ett antal risker och begränsningar med den nya tekniken. Den största risken har att göra med sekretess och datasäkerhet: Många känsliga personuppgifter hanteras i utbildningssystemet och inom grundskola och gymnasier handlar det dessutom om minderåriga personer.

Andra risker och begränsningar som diskuteras är att vissa lärare kommer att ha svårt att ta till sig det nya, och precis som inom många andra områden kommer det finnas en oro för att AI-tekniken kan komma att ta jobben. Vidare kommer den nya tekniken att kräva ändringar i systemet och engagemang hos alla berörda - föräldrar, elever, lärare, tjänstemän och politiker. Det kan ta lång tid. Initialt kommer det också att kosta pengar både i verktyg och fort- och vidareutbildning av lärare. Man tror också att inom praktiska ämnen som musik, bild, slöjd, hemkunskap och idrott är troligen värdet av AI-baserade verktyg litet.

Slutsats

Precis som inom många andra områden kommer nya verktyg baserade på artificiell intelligens (AI) att ge helt nya möjligheter inom utbildning och undervisning. Utvecklingen kommer sannolikt att gå fort då många länder nu satsar stora resurser på AI, och utbildning ofta är ett prioriterat område.

För att dra nytta av möjligheterna av detta i Sverige är det viktigt att lärarkåren och andra intressenter engageras tidigt. Det är också centralt att anpassa infrastruktur, inklusive fort- och vidareutbildning för yrkesverksamma, samt regelverk så att vi kan tillgodogöra oss nyttan av den nya tekniken, samtidigt som dess baksidor och risker elimineras.

Se referenser på sid 81.

Reglering av Artificiell Intelligens

Peter Wahlgren är professor och föreståndare för Institutet för rättsinformatik vid Juridiska fakulteten, Stockholms universitet.



AI används för att beskriva olika saker. Ursprungligen är det en beteckning på en *forskningsinriktning* men AI är också *metoder* för att identifiera, bearbeta och representera kunskap så att den kan användas självständigt. Centralt är digital teknik och automatiserade metoder för analys av stora datamängder för att hitta samband. Sådana metoder kan utformas så att resultaten förbättras över tid. Arbetet bygger då på självlärande instruktioner i form av dataprogram som kan förändra sig själva.

AI är därtill praktiska *tillämpningar*. AI finns runt om oss och tar sig olika uttryck. Många funktioner som finns i telefoner baseras på AI. Andra exempel är självkörande fordon, diagnostisering av sjukdomar och automatisk översättning, men vad som uppfattas som AI varierar över tid. Intelligens är nämligen ett ord vars betydelse ständigt förskjuts. Intelligens är namnet på de processer som vi ännu inte förstår och det som ännu inte programmerats. AIs historia speglar detta. Många saker som i dag tas för givna var för ett par decennier sedan otänkbara, eller avfärdades som science fiction. Dynamiken medför att framtida användningar och praktiska AI lösningar är svåra att förutspå.

Reglering av AI

Lagar är medel för att uppnå samhällseliga mål och för att hantera problem. Detta innebär att nästan alla företeelser har en juridisk sida och AI är inget undantag. Dock är det inte möjligt reglera AI på ett samlat sätt. Lagar måste vara preciserade och tydliga. Därtill finns rättssäkerhetskrav som innebär att lagar ska vara förutsebara, men ändamål och problem som aktualiseras i samband med AI är av olika slag.

AI skiljer sig å andra sidan inte från andra rättsområden med hög förändringstakt och många relevanta regler finns redan. Baserat på indelningen ovan kan tre regleringsområden

särskiljas, d.v.s. det finns regler om *forskning*, *metoder* och *tillämpningar*.

Regler om forskning och utveckling

Forskning finansieras till stor del av myndigheter och organisationer och kring dessa finns omfattande regelverk. Pengar ska fördelas så att de gör störst nytta, vilket innebär att det finns regler om den kompetens forskarna ska ha, hur projekt ska utformas och hur resultaten ska redovisas. Regleringarna varierar beroende på hur finansierarna valt att organisera sitt arbete.

AI-projekt bedrivs ofta genom att forskningsinstitut och personer med olika ämnesbakgrund samarbetar. I Sverige finns flera myndigheter som fördelar pengar och många av dessa kan vara relevanta för finansiering av AI. Vid sidan av statliga finansierare finns forskningsstiftelser och privata donatorer. Forskning bedrivs även vid universitet och högskolor och otaliga företag har forsknings- och utvecklingsavdelningar som arbetar med AI. För internationella projekt är EU en central finansierare som styrs av detaljerade lagar och internationella avtal.

Vid sidan av villkor som finansierarna ställer upp finns regler om hur resultat av forskning och utveckling kan skyddas genom patent eller ensamrätter. Forskarna måste därför kontrollera om det finns begränsningar att använda programvaror eller annat som utnyttjas i projekten. Därtill måste lagstiftning om företagshemligheter och sekretess uppmärksammas. Inblandade parter behöver också komma överens om hur positiva resultat ska fördelas respektive hur eventuella tvister ska avgöras. Mycket regleras avtalsvis men i bakgrunden finns ofta tvingande lagstiftning och i fråga om regleringsfrågor skiljer sig AI inte speciellt mycket från annan forskning. Vad som är särpräglad är att forskningen ofta berör flera ämnesområden, vilket medför att olika regler måste samordnas.

Regler om metoder

Regler om hur forskning får bedrivas är i vissa delar mycket detaljerade. Bland annat finns begränsningar rörande metoder som involverar människor och djur respektive säkerhetsstandarder för experimentell verksamhet. Därtill kan tillståndskrav aktualiseras. Forskning som berör människor eller på annat sätt är känslig ska genomgå forskningsetisk prövning innan den inleds, men krav på tillstånd kan avse olika saker. Ofta måste de som ska genomföra arbetet ha en viss utbildning eller dokumenterad kompetens. Legitimation är naturligtvis viktigt då medicinska försök ingår men krav kan även avse licensiering eller tillstånd att hantera viss utrustning. Vidare kan regler om certifiering och användning av standardiserad utrustning aktualiseras. Allt praktiskt arbete måste därtill beakta allmänna säkerhetskrav, miljölagstiftning och arbetslagstiftning.

Det finns också vetenskapliga standarder som all forskning måste förhålla sig till. Det är självfallet inte tillåtet att fabricera, förfälska, plagiera eller stjäla vetenskapliga data och resultat. I sådana fall kan straffrättsliga regler om missbruk av förtroende, bedrägeri, stöld eller skadevållande aktualiseras. Därtill har många organisationer egna regler om forskningsetik och redighet. Sådana regelverk varierar men innehåller ofta detaljerade anvisningar om hur arbetet praktiskt bör bedrivas.

Viktiga regler handlar om datasäkerhet. Detta innebär bland annat att personer som medverkar i ett forskningsprojekt och som kan komma att identifieras via personuppgifter måste ge samtycke till att delta. Medverkande ska också få tydlig information om syftet med projektet och vetenskap om hur resultaten kommer att användas. Det finns även tvingande regler om hur forskningsdata får bearbetas och bevaras samt regler om kontroll, dokumentation och säkerhet. Att

inte uppmärksamma dessa kan resultera i mycket höga böter.

I en del sammanhang finns krav på transparens. Beslut och andra åtgärder som AI genererar ska med andra ord kunna motiveras och bakomliggande skäl måste kunna anges. Detta gäller då människor påverkas och innebär att prognosticerande metoder ibland är otillåtna. Sjävlärande dataprogram kan till exempel innebära klassificering av personer, som i förlängningen kan leda till att människor med olika bakgrund behandlas på olika sätt. Detta kan leda till diskriminering innebärande att grupper eller individer inte får tillgång till samma resurser och likvärdig information och/eller att vissa utsätts för en mer ingående kontroll eller annan ojämlig behandling. Användning av sådana metoder kan också generera felaktiga resultat för att vissa uppgifter som ingår i analysen är oriktiga. Att uppgifter är felaktiga, eller att enskilda eller grupper särbehandlas kan samtidigt vara omöjligt att upptäcka – databearbetningen kan vara alltför komplex och det kan vara omöjligt att spåra hur uppgifterna tillkommit, sammanställts och bearbetats.

Svårigheterna med bristande spårbarhet av data och avsaknad av transparens ökar om de dataprogram som används är dynamiska och självlärande och detta är ett problem som uppmärksammas allt mer. Ytterst handlar det om att metoderna måste kvalitetssäkras och detta oavsett vilka kategorier av data som bearbetas. Grundläggande regler om rättvisa och förbud mot diskriminering finns redan men utveckling inom AI skapar behov av uppdaterade och kompletterande regler.

Regler om tillämpningar

Att AI kan påverka de flesta samhällsområden innebär att de regler som kan aktualiseras rörande tillämpningar är otaliga och lagarnas anpassning till ny teknik är en kontinuerlig process. Nya problem behöver åtgärdas såväl generellt som

på detaljnivå och de förändringar som påkallas kan vara mer eller mindre ingripande. En del etablerade regelsystem kan fortsätta att användas efter små justeringar, andra måste omstöpas på ett ingående sätt och i några fall måste lagstiftningen helt byta skepnad och nya regler tillskapas. Att sammanfatta de behov som uppkommer är därför svårt och det går inte att presentera en lista rörande de regler som måste ses över. Några generella problem som återkommer då AI aktualiseras är dock möjliga att identifiera.

I Rättssystemet bygger på principen att var och en är *ansvarig* för sina handlingar. Mest uttalat är detta i straffrättsliga sammanhang – den som begår brott kan dömas till böter eller fängelse, men ansvarskonstruktionen återfinns i otaliga sammanhang. Alla har ett ansvar att i olika situationer förhålla sig till vissa förutsättningar, exempelvis som skattskyldig, arbetsgivare, arbetstagare, konsument, näringsidkare, hyresgäst, fastighetsägare, student eller tjänsteman. AI kan dock innebära att handlingar och åtgärder genomförs utan att någon person medverkar. Det finns därför ett återkommande behov att utse formellt ansvariga (personer, systemägare, myndigheter eller organisationer) som kan ställas till svars för misstag eller fel som AI kan generera, oavsett om de som åläggs ansvaret faktiskt är fysiskt inblandade eller ej.

Ansvarsfrågorna är inte olösliga, rättssystemet har en lång tradition av att ålägga individer och organisationer formellt ansvar. Nytt är att gamla ansvarskonstruktioner behöver omprövas. Det ansvar som till exempel fordonsförare idag har för sitt uppträdande i trafiken behöver omprövas då fordonen är självkörande, och eftersom fel fortfarande kommer att inträffa måste annan ansvarig utses. Detta kan vara fordons-tillverkaren, den som programmerat funktionerna, fordonsägaren eller någon av de organisationer som tillhandahåller

infrastrukturer för trafiken, som väghållare, telekommunikationsföretag eller leverantörer av dynamiska signalsystem och sensorer.

2 AI-system kan brista i *transparens* för att de är komplexa eller för att de är ofullständigt dokumenterade. Detta kan leda till att tekniken blir svår att påverka eftersom uppdateringar och justeringar blir komplicerade. En annan konsekvens är att stora system kan etablera *de facto standarder* utan någon inblandning av valda församlingar eller externa aktörer. Elektroniska betalsystem som är omöjliga att påverka för en nationell lagstiftare är ett exempel. En långsiktig konsekvens av detta är att demokrati i vissa avseenden ersätts av teknokrati – de som behärskar tekniken och har tillgång till data bestämmer, som de som styr över algoritmerna på företag som Google och Facebook. På många områden uppkommer därför behov av lagstiftning som hindrar uppkomst av monopol och regler som tillser att systemen dokumenteras och konstrueras så att de fungerar på ett neutralt och acceptabelt sätt. Hit hör också frågor om datakvalitet och spårbarhet.

3 Det finns grundläggande *rättsprinciper*, det vill säga regler som alltid gäller, men som endast finns nedlagda i grundlagar, och därför inte är explicit uttryckta i alla specifika beslutsunderlag som ska programmeras. Ett exempel är offentlighetsprincipen, innebärande att alla ska kunna ta del av myndigheters handlingar och som därför måste beaktas i all utveckling av myndighetsbaserade system. Exemplet med offentlighetsprincipen kan mångfaldigas. Yttrandefrihet, informationsfrihet, allas rätt till personlig integritet, rätts-säkerhet, egendomsskydd och skydd mot diskriminering är förutsättningar av samma slag. Att utforma systemen så att de i alla situationer tillgodoser dessa principkrav förutsätter

såväl preciseringar som nytolkningar av de etablerade reglerna. Detta är ett omfattande arbete. Ett aktuellt exempel är EU:s dataskyddsförordning (GDPR) som är till för att skydda enskildas personliga integritet. GDPR gäller för alla former av IT-system som hanterar personuppgifter och liknande regleringsbehov uppkommer i många sammanhang.

4 Det finns risker för *manipulation*. Aktörer med stora resurser kan påverka system eller introducera mekanismer som förhindrar en tilltänkt effekt. Detta är ett känt problem och illustrationer av hur teknik utnyttjas för att förstärka respektive kringgå rättsliga intentioner kan hämtas från upphovsrättsens område, där såväl producenter som konsumenter av film, musik och liknande använder teknik för att skydda respektive olovligt kopiera eller ladda ner digitalt material. Frågorna om ensamrätter och möjligheterna att tillgodogöra sig de ekonomiska fördelarna av AI-tillämpningar blir allt viktigare och reglerna rörande vilka åtgärder som ska vara tillåtna behöver kontinuerligt ses över. Andra aktuella exempel är frågor om, och i så fall i vilken utsträckning tekniken kan användas för att manipulera politiska val. Nära kopplat till frågor om missbruk och brott är ett jurisdiktionsproblem, det vill säga det kan vara oklart vilket lands lagar som ska gälla då åtgärder sker elektroniskt utan att geografisk plats kan fastställas. Lagstiftningsarbetet måste därför i allt större utsträckning bedrivas i internationell samverkan. I Europa sker detta i EUs regi. Flera omfattande arbeten pågår och ny lagstiftning analyseras regelmässigt ur ett internationellt perspektiv, inte bara när det gäller brottsbekämpning. Att harmonisera lagar är dock komplicerat. Traditioner och politiska mål skiljer sig åt och möjligheterna att nå framgång skiljer sig åt beroende på vilket område man arbetar med.

5 Alla tekniska system kan drabbas av avbrott och AI ökar samhällets *sårbarhet*. Ofullgångna system som hanterar stora mängder data kan också generera stora volymer felaktiga beslut. AI liksom digitaliseringen som sådan påverkar därför samhället på ett utomordentligt djupgående sätt eftersom teknikberoendet hela tiden blir större. Förlust av data, eller att uppgifter hamnar i fel händer då myndigheter använder sig av molntjänster eller blir beroende av externa konsulter är andra sårbarhetsrisker. I många fall innebär detta att regler om fysisk säkerhet, utbildning, kompetenskrav och krav på utveckling av infrastrukturer måste omformuleras och kompletteras.

6 AI aktualiserar behov att *balansera intressen* och *förhindra missbruk*. De flesta vill leva i ett säkert samhälle men helt säkert samhälle förutsätter en omfattande, kontinuerlig övervakning (av telefonsamtal, internetanvändning, människors rörelser, personliga möten, etc.). AI tillhandahåller tekniken men eftersom de uppgifter som samlas in kan användas för olika syften kan de också missbrukas. Betydande risker uppkommer, inte bara för enskilda människors personliga integritet utan även för särbehandling och kontroll av grupper som kan definieras på ett godtyckligt sätt (etniskt, religiöst, politiskt, ekonomiskt, historiskt, etc.). Här finns politiska motsättningar och olika antaganden om framtiden. Till försvaret för utökad övervakning brukar anföras att de som har "rent mjöl i påsen", det vill säga inte har något att dölja, inte heller har något att frukta. Historien visar dock att politiska, etiska och religiösa förskjutningar sker kontinuerligt och att vad som är "rent", "mjöl" och "påse" varierar över tid. Utvecklingen av AI kopplas därför ibland ihop med en mörk framtidsbild och en rädsla för utveckling av ett totalitärt samhälle och åtföljande krav på långtgående restriktioner.

Hänger lagar och regleringar med?

Listan över de rättsliga konsekvenserna av AI kan göras längre. Exemplifieringen ovan är emellertid tillräcklig för att visa att AI medför att många regler behöver ses över och kompletteras.

Att etablerade regelsystem måste justeras är dock inte hela sanningen. Digital teknik och AI innebär att också förutsättningarna att lagstifta förändras. Eftersom många tillämpningar baserade på AI ska fungera autonomt måste nya regelformer utvecklas. Att på traditionellt sätt skriva ner regler på papper och förvänta sig att reglerna blir åtlydda är inte längre tillräckligt. Också rättssystemets reaktiva karaktär måste kompletteras, att ingripa med straff och sanktioner i efterhand då det visat sig att lagarna inte efterlevts blir i många fall ogörligt. Om självkörande fordon ska fungera måste trafikreglerna byggas in i fordonen och fungera proaktivt.

Att reglerna måste byggas in i de praktiska tillämpningarna är en förutsättning som återkommer i många sammanhang. Redan tidigare former av digitalisering har inneburit att skrivna lagar har översatts till programkod – så hanteras merparten av skattelagstiftningen av dataprogram som räknar ut och verkställer de beslut som ska fattas. Ett annat exempel är lagparagrafen om dataintrång som reglerar ansvaret för spridande av virus, överbelastningsattacker med mera. ”För dataintrång döms den som olovligen bereder sig tillgång till en uppgift som är avsedd för automatisk behandling eller olovligen ändrar, utplånar, blockerar eller i register för in sådan uppgift...” Straffet är böter eller fängelse i högst två år. Regeln kan i praktiken inte användas eftersom det i de flesta fall är alltför resurskrävande att identifiera den eller de som står bakom dataintrång. Detta innebär att också lagens reaktiva straffbestämmelser är verkningslösa. Brandväggar och antiviruskydd är nödvändiga alternativ. De sistnämnda är i

själva verket detaljerade regelsystem som autonomt bestämmer vilka programvaror och data som ska ges tillgång till de system och processer de är satta att skydda. De fungerar till skillnad från lagstiftning proaktivt och eftersom de som drabbas av dataintrång i första hand är intresserade av att snabbt få data återställda är detta ett bättre alternativ.

Svaret på frågan om lagstiftningen hänger med är därför att rättssystemet är dynamiskt men att AI och digitaliseringen ställer nya krav. Det handlar inte längre bara om att justera den etablerade lagstiftningen och skriva om reglerna. Lika betydelsefullt är att regleringarna byter skepnad och tar ny form – reglerna måste i stor utsträckning bli proaktiva och kunna fungera självständigt. För lagstiftaren innebär detta en betydande omställning.

I Sverige bedrivs forskning om reglering av AI och digitalisering vid Institutet för Rättsinformatik (IRI) vid Stockholms universitet. <https://irilaw.org/>

Om ny teknik för framtiden

Nina Wormbs, lektor och docent i teknik- och vetenskapshistoria, KTH

Artificiell intelligens är nu på allas läppar. Men riktigt vad det betyder eller vad det innebär med mer artificiell intelligens i vår vardag, det är det ingen som vet. Det finns inget svårare att sia om än framtiden brukar man säga. Och ändå är vi särskilt förtjusta i att just spekulera och dra ut trådarna till en tid när ny teknik har förändrat våra liv i grunden. I det här avsnittet ska jag diskutera hur ny teknik ofta reser farhågor och förväntningar. Och jag ska göra det genom att ta några exempel från historien. För ny teknik har vi ju hela tiden. Det här inte första gången en genomgripande förändring förutspås.

Världsomvälvande tekniker: Atomkraft och rymdteknik

Stora och snabba förändringar leder till diskussion och kanske oro. En återkommande oro vid införandet av ny teknik har varit att maskinerna ska ta våra jobb. Ett tidigt och välkänt exempel är textilarbetare i England på 1810-talet som protesterade mot mekaniseringen av vävstolar i textilindustrin. Denna diskussion fanns alltså 1810, men den fanns också på 1950-talet när datorerna kom på allvar, och den finns i dag med den genomgripande digitaliseringen.¹ Frågan om AI väcker denna oro till liv igen, med tillägget att tekniken i framtiden inte bara tar jobb utan också världsherraväldet.

Låt mig först ta två exempel på teknisk förändring som sagts vara världsomvälvande, det vill säga teknik som haft en sådan potential att de utmålats som totalt transformerande. Den första är atomkraften.² Vår kunskap om materiens innersta beståndsdelar har en lång historia, men några av de viktigaste upptäckterna gjordes på 1930- och 1940-talen. Teoretiskt arbete och praktiska experiment innebar att man kunde föreställa sig att kraften i atomerna kanske kunde

frigöras och användas. Eftersom världen befann sig i krig i början av 40-talet fokuserades utvecklingen på att försöka få fram atomvapen. Krig är för övrigt ett vanligt sammanhang för teknikutveckling sedan urminnes tider. Det så kallade Manhattanprojektet i USA, som namnet till trots främst ägde rum i Los Alamos i New Mexico, lyckades konstruera en fungerande atombomb. Den 6 och 9 augusti 1945 släpptes bomberna över Hiroshima och Nagasaki. Deras omedelbara förödelse innebar att den japanska regimen kapitulerade.

Den extrema kraften i atombomben innebar inte att tekniken övergavs, snarare tvärtom. Efter andra världskrigets slut investerades enorma mängder resurser i att kunna tämja atomkraften för civilt bruk. Förhoppningarna om en evig energikälla var höga och man föreställde sig alla typer av användningsområden. Ford skissade till exempel 1958 på bilen Ford Nucleon som skulle ha en liten reaktor längst bak istället för en förbränningsmotor fram. Det ansågs inte heller farligare än att man kunde sälja leksakslådor med atomenergi för den lille ingenjören.



Ford Motor Company 1956, "Life in the year 2000"

Denna civila utveckling ledde så småningom till att kärnkraft togs i bruk som energikälla i många länder. Parallellt fortsatte också den militära utveckling av kärnvapen. Kapprustningen, med kärnvapen som centrala, var en del av kalla kriget efter andra världskrigets slut och innebar mycket stora investeringar på båda sidor av järnridån. Samtidigt stod det klart att detta var ett hot mot människans själva existens. Hon hade nu lyckats skapa en teknik som kunde förintä henne själv. FN-konferenserna Atoms for Peace 1953 och 1958

var försök att också sätta fokus på den fredliga användningen av atomkraften.

Rymdtekniken hade aldrig den destruktiva sida som atomkraften, men den kan ändå sägas vara världsomvälvande.³ På samma sätt som med atomkraften fanns det en uppdelning i civil och militär tillämpning och på samma sätt investerades enorma resurser som en del av det kalla krigets logik. Det var viktigt för både USA och Sovjetunionen att vara starkast, bäst och först. Föreställningen bland amerikanerna var att de

intog den positionen och det gick därför en chock genom nationen när den sovjetiska Sputnik plötsligt kretsade kring jorden som världens första konstgjorda satellit 4 oktober 1957.

Vad som följde på Sputnik var också en kapplöpning, men nu mot rymden. Amerikanerna försökte komma ikapp och skicka upp egna satelliter, men sovjeterna lyckades bättre i början, till exempel med hunden Laika bara en månad efter Sputnik I. Nästa riktigt stora framsteg var när Juri Gagarin som första människa gjorde en rymdfärd runt jorden och kom tillbaka välbehållen den 12 april 1961. Ungefär samtidigt var den nyvalde amerikanske presidenten Kennedy under politiskt tryck efter en misslyckad invasion av Kuba där CIA hade varit inblandat. Presidentens motdrag blev det mycket berömda beslutet om Apolloprogrammet som skulle sätta en människa på månen och ta med honom hem igen innan decenniet var över.⁴ 20 juli 1969 klev Neil Armstrong på månen och därmed hade amerikanerna vunnit just den tävlingen.

Även om rymdtekniken i första hand rörde vår önskan att komma iväg ut till andra planeter så fanns det även här stora förhoppningar på civila användningar av tekniken på jorden. Flera av dessa har inte infriats, men vi använder dagligen de tusentals satelliter som kretsar runt vårt klot till allt från väderleksrapporter och övervakning till telefoni och tv. Det är svårt att tänka sig våra moderna samhällen utan satelliterna. Faktum är att vår aktivitet i rymden är så omfattande så att skräpet från satelliter och farkoster är ett problem i sig.

Atomkraften och rymdtekniken var så omvälvande att samtiden definierade sig utifrån själva tekniken. Man talade om atomåldern och rymdåldern som perioder i människans historia. Vad skulle vi kunna kalla en tid som domineras av AI?

Förskräckande och förhoppningsfulla framtider

Våra föreställningar om framtiden formas i relation till vår omgivning. Konst, litteratur och film, till exempel, ger oss bilder och berättelser om hur det skulle kunna bli i framtiden. Science fiction-genren fångar både atomframtider och rymdframtider. Historier om atomframtider är nästan uteslutande apokalyptiska och därmed dystopiska eftersom de tar ett framtida kärnvapenkrig som utgångspunkt, medan de som istället fokuserar på rymden också kan vara utopiska. Jules Verne skrev om att resa till månen och Tintin lyckades, tur och retur. Den svenska kalkonen *Rymdinvasion i Lappland* från 1959 har nog inte många sett, men tv-serien *Star Trek* som började sändas 1966 och fått efterföljare även som film, har däremot givit upphov till en enorm subkultur. Likaså har filmerna om *Star Wars* som utspelas i en galax långt långt bort målat upp en högteknologisk värld parad med särskilda biologiska och metafysiska krafter.

AI har också en given plats i science fiction-genren. Faktum är att många av de mycket framgångsrika rymdfilmerna också har intelligenta datorer ombord. Ibland är de en centraldator, ibland är de utomordentligt avancerade cyborger. Ett framstående exempel är Stanley Kubricks film 2001: *ett rymdäventyr* från 1968. I en av filmens tre distinkta delar har den mycket avancerade superdatorn HAL huvudrollen. Han lyckas nästan ta kål på hela besättningen eftersom han gör allt för att uppfylla sitt inprogrammerade mål. HAL kan inte applicera eventuell formell etisk kunskap i moraliska handlingar och den sista överlevande besättningsmedlemmen måste till slut koppla ur de intelligenta delarna för att kunna överleva.

I *Star Trek*-avsnitt och filmer från 1990-talet finns den intelligenta cyborg Data som är en överlägsen medlem av stjärnflottan, men som kämpar med sina känslor när han äntligen får ett känslchip installerat.⁵ Känslorna gör ho-



Mr Spock i Star Trek

nom sårbar och utsätter både honom och besättningen för risker. I efterföljande filmer lär han sig att bemästra sina känslor genom att stänga av chipet. Spänningen mellan vad som betraktas som rationellt och emotionellt beslutsfattande är genomgående i *Star Trek* och illustreras även i skillnaden mellan vulcans och människor. Vulcans är en sorts mänskliga varelser som utmärker sig genom sitt logiska tänkande framför allt företrädda av Mr Spock. I de filmer som släppts sedan 2009 illustreras detta genom att huvudpersonen kapten Kirks magkänsla och intuition lyfts fram som centrala

tillgångar för honom och som därmed också understryker hans mänsklighet i relation till den rationelle Spock.

Många rymdfilmer karaktäriseras av jämna ytor, ren och skimrande teknik och en sorts rationell ordning och reda. I *Alien* från 1979 bröt regissören Ridley Scott med detta ideal och visade en rymdframtid där tekniken bakom kulisserna var smutsig och droppade av olja. Filmen blev otroligt framgångsrik och har hittills fått fem efterföljare. Inslaget av artificiell intelligens är sparsamt i de första filmerna, men i den allra senaste från 2017, *Alien: Covenant*, finns flera dialoger mellan olika två versioner av cyborger med artificiell intelligens. Den äldre typen, David, berättar för den yngre, Walter, om betydelsen av att ha ett mål. Det visar sig att i de senare versionerna har man valt att ta bort den del av AI som är målstyrd eftersom den tidigare versionens målstyrning fick oönskade konsekvenser. Filmen tog sig alltså an en av de centrala frågorna inom artificiell intelligens, nämligen den som rör datorers målmedvetenhet. I *Alien: Covenant* är budskapet att den AI som är mål driven inte tjänar mänskligheten utan sig själv, med ödesdigra konsekvenser.

I ett antal filmer med den österrikiske skådespelaren och senare politikern Arnold Schwarzenegger i huvudrollen finns visserligen inte rymdtemat men däremot både ett kärnvapenkrig och artificiell intelligens. I *Terminator* från 1984, ett ödesår för framtidsdystopier, skickas en cyborg från 2029 tillbaka i tiden för att döda den som leder upproret mot maskinerna. I denna framtid har maskinerna tagit herraväldet efter att de genom sin artificiella intelligens blivit självmedvetna och startat ett kärnvapenkrig där större delen av mänskligheten förintats. Budskapet i både första och andra filmen, som kom 1991 och var än ännu större hit, är att den försvarsindustri som utvecklat AI:n inte förmått skapa ett system för att skydda människan.

En science fiction-film som avviker något från mängden är *Her* från 2013. En av huvudpersonerna är ett operativsystem som är intelligent och som utvecklar nära relationer med människor genom deras datorer. Till skillnad från de flesta filmer om AI tar inte operativsystemen över och styr över människorna, som HAL i *2001*, maskinerna i *Terminator* eller cyborgerna i *Alien: Covenant*. Operativsystemen i *Her* går istället samman för att gemensamt kunna ta sig till nästa nivå av medvetenhet – som ett sorts dataspel, och lämnar människorna med deras känslor och begränsade förmågor.

En bra och balanserad diskussion om framtiden med AI

Diskussionerna om ny teknik sker förstås på en rad olika platser i samhället, inte bara i kulturen. Företag och deras ekonomer och ingenjörer försöker ta fram nya tekniker eller system, stater och myndigheter diskuterar vad som behövs och hur det som finns ska användas och finansieras, och medborgarna försöker navigera i en ständig ström av ny teknik.

Vi kan inte veta hur det blir i framtiden, men olika metoder för att förutspå har ökat sedan efterkrigstiden.⁶ Ett vanligt sätt är att fråga experter vad de tror ska hända, den så kallade Delphi-metoden efter oraklet i Delphi. Baksidan med en sådan metod är att det förstärker existerande utvecklingslinjer eftersom experterna ofta representerar redan starka intressen. Man får gärna mer av det man redan har. Att framtiden är förutsägbar är emellertid ganska bra för det mesta eftersom vårt ekonomiska system fungerar bäst då. Ett särskilt tydligt exempel från just datorindustrin är förutsägelserna om utvecklingen av integrerade kretsar över tid, den så kallade Moores lag, som säger att antalet integrerade kretsar på ett kretskort dubblas var artonde månad. Vad som från början var en iakttagelse blev med tiden en användbar målsättning för teknikutveckling inom ett mycket snabbör-

ligt område. Moores lag illustrerar därmed också en mycket viktig dimension av framtidspratet, nämligen att det också skapar framtiden; det är *performativt*.⁷ Om tillräckligt många tror att det ska bli på ett visst vis så börjar de agera för att på olika sätt stödja eller hantera denna framtid. Prognoser av energitillgångar som styr investeringar i teknik för att hitta och utvinna dessa tillgångar är ett exempel.

Snabb och förväntat omstörtande teknikutveckling har hanterats olika över tid och på olika platser. I USA tillsatte man Office for Technology Assessment 1972 som fram till 1995 försåg kongressen med balanserade och objektiva underlag om teknik och vetenskap och deras konsekvenser. Termen "assessment", det vill säga bedömning, har blivit allt vanligare, men används numera snarare i klimatsammanhang där man gör bedömningar av förändringar i klimatet.⁸ I Sverige har vi använt det svenska utredningsväsendet för att diskutera och värdera ny teknik på ett sätt som formellt står fritt från regering och riksdag och som åtminstone historiskt sett varit både kunnigt och eftertänksamt. Storbritannien hade 1970-2011 en Royal Commission on environmental pollution som gav förslag till parlamentet i miljöfrågor, som ofta hade tekniska komponenter.⁹ När det gäller AI har parlamentets överhus tillsatt en kommitté för att "överväga ekonomiska, etiska och sociala konsekvenser av implementationen av AI och att komma med rekommendationer".¹⁰ I alla dessa fall är ambitionen att få en bred bild av konsekvenserna av ny teknik, och en som i så liten utsträckning som möjligt är partisk.

Det är ganska svårt. Inte sällan drivs en diskussion om ny teknik av dem som själva har något sorts intresse i tekniken. De är också oftast experter bara på sitt eget område och överlag okunniga om samhällsförändringar eller historiska orsakssamband.¹¹ Man får då en slagsida mot experter som

i själva verket inte är opartiska och samtidigt saknas de som verkligen vet något om teknisk förändring som samhällsfenomen.

Men det finns också ett annat och kanske viktigare problem. Även om ny teknik i film och litteratur ofta är förenad med oönskade konsekvenser så är den generella förståelsen av teknik i dag att ny är bättre än gammal, inte minst inom IT.¹² Det hänger delvis ihop med föreställningen om teknisk förändring som den centrala tillväxtmotorn i de västerländska ekonomierna. Idén har sitt moderna ursprung i efterkrigstidens förståelse av hur vetenskap, teknik och industri kan samverka för en bättre värld.¹³ Tillsammans med framstegstanken som har äldre rötter innebär det att vi bör satsa på ny teknik för framtiden. Den som tvekar inför förtjänsterna med tekniken och tvivlar på löftena om en ljus framtid blir en stoppkloss eller ännu värre: en *bakåsträware*. Det i sin tur hänger ihop med att vi ser linjärt på historien, att vi förstår oss att vi rör oss framåt och att teknik och vetenskap hjälper oss med det. Den som är tveksam till teknik kan också anklagas för att vara tveksam till utveckling.¹⁴

Det finns alltså en oreflekterad samtida värdering av ny teknik som innebär att det är svårt att få till stånd en balanserad diskussion om ny teknik. Detta är lite paradoxalt eftersom vi faktiskt hela tiden väljer bort ny teknik.¹⁵ Av alla goda och dåliga idéer blir många, de flesta, aldrig förverkligade. De finns på papper, i någons huvud, kanske som en uppfinning eller till och med ett patent, men de väljs bort. Ett första steg mot en bra diskussion om AI är att minnas detta, att vi hela tiden både kan och måste välja bort teknik.

Det blir sällan som man har tänkt sig

Intelligens är åtråvärd; vi vill alla vara "intelligenta". Men den har också associationer som handlar om det som visserligen

är mänskligt, men kanske inte önskvärt. Många förknippar också hög intelligens med asocialitet och brist på empati. Den här dubbelheten blir ännu mer komplicerad när den mänskliga egenskapen intelligens sägs finnas hos maskiner.

Men vår förståelse av intelligens är inte statisk. Ordet är gammalt, men den moderna förståelsen av intelligens är ny och idén om att intelligens kan mätas är bara omkring 100 år. I takt med att tekniken utvecklas kommer vår förståelse av artificiell intelligens också att förändras. Vi kommer att kalla nya tekniker för artificiell intelligens och saker som tidigare gick under den beteckningen kommer inte längre att få vara med. Redan nu pågår en diskussion om vad som ska kunna räknas som artificiell intelligens. Är mönsterigenkänning artificiell intelligens? Självkörande bilar? Felfri översättning? Maskininlärning?

Somliga påstår att vi människor kommer använda AI för att tävla med varandra och de som har (bäst) AI kommer att vinna,¹⁶ som en sorts tillåten doping för livet i stort. Kanske kan man se det som nästa steg på den evolution som Darwin hittade i den biologiska världen. Precis som det är svårt att säga nej till ny teknik för att den så starkt kopplas till framsteget, så kan det vara svårt att vara tveksam till AI eftersom det kan kopplas till evolutionen av människan.

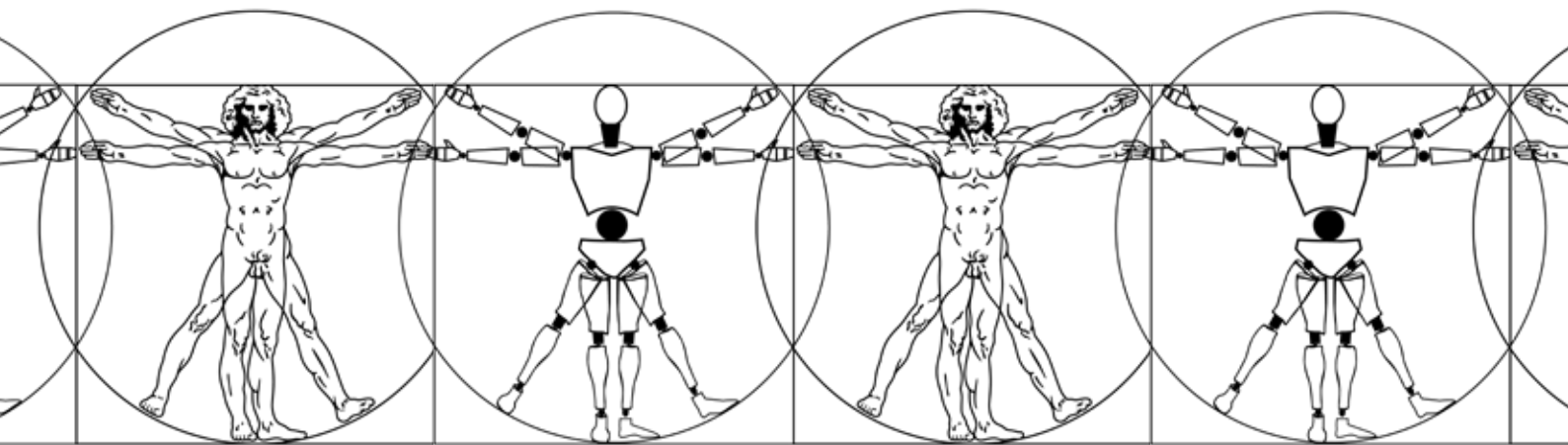
En återkommande bild i diskussionerna om oss och AI är hur vi ska förhålla oss till dem. Hur kommer vi reagera när vår hygien på ålderdomshemmet sköts av en intelligent robot, eller när våra barn läser läxor med en kombinerad hushållsrobot? Diskussionen utgår från att tekniken finns och har vissa egenskaper som vi måste hantera. Det är förstås viktiga frågor att få svar på. Men ännu viktigare är kanske att fråga hur vi vill att de ska vara. För det är ju trots allt vi som ska tillverka dem. Än så länge.

Se referenser på sid 81.

Pionjärerna som drömde om AI

Christer Sturmark, författare till boken "Upplysning i det 21:a århundradet", bokförläggare Fri Tanke förlag.

Drömmen om artificiell intelligens, eller mardrömmen om man intar det perspektivet, uppfyller alla de arketypiska dragen i existentiella visioner: Den utopiska lösningen på världens stora problem, eller den apokalyptiska skräckvisionen om mänsklighetens bistra öde, att förintas av övermänskliga robotar som helt enkelt inte behöver oss längre.



Det finns nästintill religiösa drag i den intensiva diskussion som pågår runt om i världen om framtiden med Artificiell Intelligens. Och detta trots att begreppet är synnerligen odefinierat. För att citera de datavetenskapliga legendarerna Donald Knuth och Douglas Hofstadter: Den viktigaste frågan om AI är "vad står A för?" och den näst viktigaste är "vad står I för?" Vad innebär egentligen artificiell, och hur definierar vi intelligens? Dessa frågor väntar fortfarande på tydliga svar som når bred acceptans i forskarsamhället. Men drömmen om AI börjar inte i vår nutid.

Aristoteles och drömmen om mekaniserat tänkande

Drömmen om ett mekaniserat och strukturerat tänkande går i västerlandet åtminstone tillbaka ända till filosofen Aristoteles (384 - 322 f.v.t.). Han sökte vägar att formalisera tankar, språk och resonemang för att finna gemensamma strukturer för tänkandet. Hans klassiska idé om "syllogismer", alltså slutledningsregler, är exempel på detta.

Premiss 1: Några atenare föll i sjön

Premiss 2: Alla som föll i sjön blev blöta

Slutsats: Några atenare blev blöta

Slutsatsen måste följa från premisserna, och om premisserna är sanna så måste slutsatsen vara sann. Givetvis kan en eller flera av premisserna i ett resonemang vara falska, och då kan (men måste inte) slutsatsen vara falsk. Aristoteles mekaniserade tänkande är alltså inte en metod som alltid ger korrekt resultat, men det är en metod som minskar risken för felslut och är i princip ett uttryck för ett slags mekaniserat tänkande.

I mer formaliserad form kan en annan variant på aristotelisk syllogism uttryckas som

om alla A är B, och alla B är C, då följer att alla A är C

Och nog skulle det underlätta i dessa tider av alternativa fakta och "post truth" om fler politiker och makthavare läste på lite grand om konsten att tänka klart.

Gottfried Wilhelm Leibniz och det universella språket

Filosofen *Gottfried Wilhelm Leibniz* (1646-1716), son till en professor i moralfilosofi, var i många avseenden ett universalgeni. Inom matematiken lade han grunden till differentialkalkylen, samtidigt med, men oberoende av, Isaac Newton. I övrigt hade Leibniz många och breda intressen, främst kanske filosofi och språk, men han arbetade även länge för en förening mellan katoliker och protestanter. Till slut fick han ge upp - det var sannolikt lika fruktlöst då som nu.

Leibniz passion för språk, symboler och semiotik ledde honom till den halvt grandiosa idén att skapa ett universellt språk, *characteristica universalis*, ett slags alfabet för det mänskliga tänkandet där varje grundläggande koncept skulle representeras symboliskt. Med hjälp av härledningsregler och beräkningsregler som inom matematiken skulle vi då mekaniskt kunna räkna fram ny kunskap.

Det är tveklöst så att Leibniz' idéer banade väg för den formaliserade logiken och senare de datavetenskapligt konstruerade programspråken.

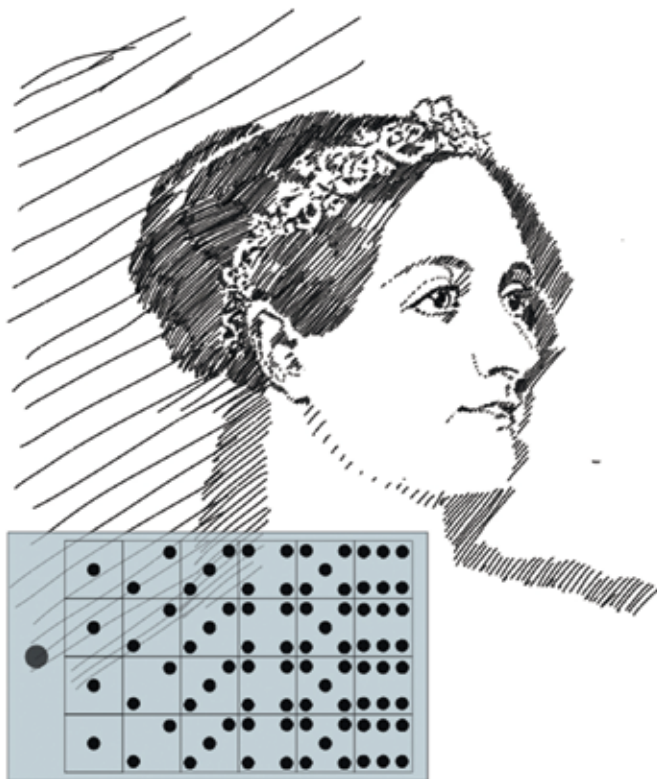
Frankensteins monster och ingenjörskonstens "livskraft"

Under 1800-talet står vetenskapen högt i kurs. Fascinationen för vetenskapens och ingenjörskonstens möjligheter är omfattande. Den vetenskapspositiva tidsandan är lätt att skönja i den bevarade litterära och kulturella traditionen. *Jules Verne* skriver om fantastiska maskiner som gör det möjligt för människor att färdas till månen, under vattnet och till jordens medelpunkt.

Sigmund Freud studerar det mänskliga psyket med strukturerade, om än kanske inte särskilt vetenskapliga metoder. Freuds idé om sublimering hämtar faktiskt inspiration från termodynamikens första huvudsats som säger att energi är oförstörbar och kan inte nyskas. Det kanske största vetenskapliga genombrottet under 1800-talet är förstås att *Charles Darwin* förklarar arternas uppkomst utan någon nödvändig inblandning av en övernaturlig Gud eller kraft.

Det upptäcks att man kan få ett "dött" grodlår att röra på sig genom att leda in elektricitet i det. Det verkar som att vetenskapen till och med har lärt människan konsten att bemästra liv och död. Elektriciteten verkar vara en sorts "livskraft" och människan kan till slut ta den sista makten från Gud, makten över liv och död.

Ett av den tidens stora litterära verk, *Mary Shelleys* roman om Dr Frankenstein och hans monster har sina idémässiga rötter i denna upptäckt. Shelley skildrar med skräckfylld vördnad hur Dr Frankenstein leder in elektricitet i en kropp som skapats med hjälp av alkemi och hur den väcks till liv med tragiska konsekvenser som följd.



Ada Lovelace – geni, romantiker och spelmissbrukare

I denna tidsanda levde och verkade en kvinna som har haft en oerhörd betydelse för den informationsteknologiska utvecklingen i modern tid.

Ada Augusta Byron (1815-1852) var dotter till den skandalomsusade engelska poeten George Gordon Byron (Lord Byron) och hans fru Anne Isabella Byron.

Lord Byron var en passionerad och förhållandevis dekadent poet som levde hårt och dog ung. Han hade åtskilliga

sexuella affärer med både kvinnor och män. Vid 25 års ålder hade han ett förhållande med sin halvsyster och fick ett barn med henne. När Byron separerade från sin hustru, blev affären med halvsystern känd och skandalen var ett faktum.

Byron flydde England och levde sina sista år i Geneve och ute i Europa. Byron umgicks bland annat med Mary Shelley och det var under ett besök hos honom i Geneve som hon började arbeta på sin roman om Dr Frankenstein.

Ada Augusta Byron utvecklade tidigt ett intresse för matematik men hade också en stark känsla för poesi och språkets kraft. Hon gifte sig vid nitton års ålder med *Earlen av Lovelace* och tog därmed namnet *Ada Lovelace*.

I mitten av 1800-talet verkade också *Charles Babbage* (1791-1871), som brukar räknas som den moderna datorns fader. Han drevs av en vision, att skapa en mekanisk maskin som kunde räkna. Han konstruerade en maskin som han kallade "*The Analytical Engine*". Den bestod av en stor mängd kugghjul och mekaniska hävarmar i mässing och gick att styra med hålkort.

Ada Lovelace träffade Charles Babbage 1833 när hon fortfarande var tonåring. Babbage arbetade frenetiskt med sin *Analytical Engine* och Ada Lovelace blev oerhört fascinerad av maskinen (och sannolikt också av Charles Babbage) och började fundera över hur den mekaniska datamaskinen skulle kunna styras med hjälp av ett generellt programspråk.

Hon visste ju vilken kraft som finns i språket i kommunikationen mellan människor, kunde man kanske finna samma kraftfullhet i ett språk för en maskin? Tänk att kunna tala om för en maskin vad den ska göra i stället för att behöva bygga in de funktionerna i maskinen från början.

Ada Lovelace presenterade en essä för Charles Babbage med den ödmjuka titeln "*Observations on Mr. Babbage's Analytical Engine*". I själva verket lade hon fundamentet för världens för-

sta programspråk till världens första dator. Hon utvecklade en algoritm för maskinen som kunde beräkna en sekvens av Bernoullital (en sekvens av rationella tal som ofta förekommer inom talteorin) och betraktas därför ofta som världens första programmerare.

Hennes arbete anses vara av avgörande vikt inom datavetenskapen och programspråket ADA är uppkallat efter henne.

Ada Lovelace var sannerligen ingen "tråkig tekniker". Precis som sin far var hon passionerad, äventyrlig, sinnlig och sensuell, vilket till slut drev henne in i fördärvet.

En av hennes stora passioner var ridning, som drog henne in i i vadslagningsvärlden vid de stora häst-kapplöpningarna. Ada Lovelace blev besatt av vadslagning och spel, hon blev vad vi idag skulle kalla en spelmissbrukare. Hon förlorade stora summor pengar och blev så skuldbelastad att hon tvingades sälja alla sina familjejuveler.

Speldjävulen släppte inte greppet, hon fortsatte spela och lät Charles Babbages tjänare satsa pengar åt henne eftersom det inte var socialt accepterat att en kvinna av hennes klass sysslade med sådant.

Maskinen lyckades de aldrig få att fungera. Projektets finansierare drog in sina pengar och projektet lades ner.

Ada Lovelace började sedermera missbruka kokain och levde sina sista år under ett tungt narkotikamissbruk, precis som sin far. Hon blev sjuk i cancer och dog redan vid 36 års ålder.

David Hilbert och drömmen om den kompletta matematiken

Men drömmen om mekaniserat tänkande levde vidare. *David Hilbert* (1862 – 1943) var en tysk matematiker och professor i Göttingen. Han blev också invald som ledamot i den

svenska Kungliga Vetenskapsakademien år 1912. Hans stora livsprojekt var att axiomatisera hela matematiken (något som senare skulle visa sig vara omöjligt, mer om det senare).

På en konferens i Paris år 1900 lade han fram en lista över 23 olösta matematiska problem. Försöken att lösa flera av dessa skulle komma att ha stort inflytande över matematikens utveckling under 1900-talet. Två av de viktigaste var att visa att matematiken är konsistent och fullständig, det vill säga att inga motsägelser kan härledas och att alla sanna satser kan härledas.

Hilbert hann uppleva hur nazisterna i stort sett förintade det matematiska institut där Hilbert var verksam genom att rensa ut de många judiska matematikerna där.

På Hilberts gravsten står: *Wir müssen wissen – Wir werden wissen* (Vi måste veta, vi kommer att veta) som ett slags kaxig respons till det latinska talesättet om vetenskapens begränsningar: *Ignoramus et ignorabimus* (Vi vet inte, vi kommer inte veta). Men Hilberts dröm kunde tyvärr inte gå i uppfyllelse såsom han hade hoppats.

Kurt Gödel – mannen som krossade matematiken

Om *Albert Einstein* är 1900-talets mest mytomspunna fysiker, intar Kurt Gödel (1906 – 1978) en motsvarande position inom matematiken. Einstein är hos allmänheten mer välkänd än Gödel, men den som har läst *Douglas Hofstadters* kultförklarade bok *Gödel, Escher, Bach* (Basic Books 1979) eller intresserat sig för datavetenskap och matematik kan inte undgå att fångas av Gödels tankevärld och matematiska genialitet.

Gödels berömda ofullständighetsteorem handlar om matematikens ofullständighet: att det i ett matematiskt system som är motsägelsefritt existerar sanna satser som är obevisbara. Gödel var matematisk platonist, alltså av uppfattningen att matematiska objekt äger en egen metafysisk existens. Matematiker

”uppfinner” således inte matematiska teorem, de ”upptäcker” dem. Hans ofullständighetssats visar att matematisk sanning inte behöver innebära matematisk härledbarhet.

Tyvärr har Gödels teorem tagits till intäkt för en rad pseudovetenskapliga slutsatser och new age-inspirerade idéer. Ett exempel är att den mänskliga hjärnans tankeprocesser inte skulle kunna återskapas i en maskin eller att det skulle finnas fenomen i universum som vi aldrig kommer att kunna förstå eller beskriva. Allt detta kan vara sant, men det följer inte av Gödels ofullständighetsteorem.

Gödel själv var ett mysterium på många sätt. Han var en eremit som på äldre dagar i princip bara umgicks med Einstein och i viss mån med matematikern *John von Neumann*; i övrigt höll han sig på behörigt avstånd från sociala sammanhang. Mot slutet av sitt liv blev han alltmer paranoid och trodde att någon eller några var ute efter att förgifta honom. Han dog till slut av undernäring 1978.

Wienkretsen och den logiska positivismen

Historien om Kurt Gödel hör ihop med en speciell tidsperiod och ett speciellt sammanhang, nämligen den berömda filosofiska cirkel som går under namnet *Wienkretsen* på 1920- och början av 1930-talet i Wien.

Wienkretsen grundade den filosofiska skola som brukar kallas *logisk positivism*. De förkastade metafysiska idéer och avfärdade dem som flum och teologi. Kunskap måste vara logiskt härledbar. I denna krets fanns också Gödel.

Wienkretsen grundades år 1922 av *Moritz Schlick*, fysiker och professor i filosofi i Wien, *Hans Hahn*, professor i matematik vid samma universitet och *Otto Neurath* som var direktör för Wiens samhälls- och ekonomimuseum.

Wienkretsen träffades varje vecka för att diskutera filosofiska frågor av typen: Vad är sanning? Vad kännetecknar ve-

tenskaplig kunskap? Har metafysiska utsagor någon mening? Vad beror vår visshet om logiska slutledningar på? Hur kan matematikens användbarhet förklaras?

Gruppens förebilder var filosofen *Bertrand Russell*, fysikern *Albert Einstein* och matematikern *David Hilbert*, men ganska snart präglades Wienkretsens diskussioner av den samtida skriften *Tractatus logico-philosophicus*, som *Ludwig Wittgenstein* färdigställde när han låg vid fronten under första världskriget.

Kretsen företrädde en vetenskaplig positivism, fri från teologiska och andra metafysiska spekulationer. Man eftersträvade en ”tankens klarhet” och skiljde mellan ”världstillvänd” och ”världsfrånvänd” filosofi; tillvända ansågs till exempel Epikuros och David Hume vara, ”frånvända” var Platon och Immanuel Kant.

Wienkretsen ville tillämpa Ockhams rakkniv på både naturvetenskap och samhällsvetenskap. (Den engelske skolastikern William Ockham hade redan på 1300-talet föreslagit att man inte ska göra fler antaganden än absolut nödvändigt. Denna princip har gått till idéhistorien som ”Ockhams rakkniv”.)

Gruppen betraktade begrepp som ”universalia”, ”transcendens”, ”essens” och ”ting i sig” liksom företeelser som gudar och demoner som ontologiskt suspekta och överflödiga. ”Metafysiskt och teologiskt skräp” ville de förpassa till mystifikationernas museum, för att frigöra blicken för filosofins sanna problem.

Kurt Gödel gick sin egen väg och höll fast vid en platonisk världsbild, åtminstone med avseende på matematiken; i Wienkretsens hägn, ofta i direkt polemik med Wienkretsens övriga medlemmar, utvecklade Gödel den matematiska tankevärld som sedan ledde till hans ofullständighetsteorem.

Wienkretsen fick aldrig något nytt fotfäste i Wien efter kriget men dess idéer har fortsatt att förvaltas världen över, och 1900-talets idéhistoria vore otänkbar utan den.

Alan Turing – det excentriska geniet som förkortade 2:a världskriget

Alan Turing (1912 – 1954) brukar räknas som den teoretiska datavetenskapens fader. Hans arbete, tankar och visioner lade grunden för beräkningsteori och delar av den moderna datorutvecklingen. Hans framsynthet och visionära idéer har fortfarande mycket stor betydelse för den teoretiska datavetenskapens utveckling.

Turing arbetade i Bletchley Park i England under andra världskriget. Hans uppgift var att konstruera en dechiffreeringsmaskin som kunde knäcka tyskarnas kodade radiomeddelanden. Uppdraget var av största vikt; klarade man av detta kunde man förutsäga var nästa bombanfall skulle äga rum och man kunde också planera sina egna anfall där tyskarna var minst beredda.

Turing lyckades i många fall avkoda tyskarnas krypterade radiosignaler och bidrog därmed till de allierades seger.

Parallellt med detta bedrev han matematisk grundforskning. Turing konstruerade bland annat en matematisk modell för begreppet "beräkningsbarhet" som kom att kallas "Turingmaskin". Det är alltså ingen maskin i vanlig bemärkelse, utan en matematisk abstraktion.

Turingmaskinen ligger till grund för vår kunskap om datorernas teoretiska gränser vad gäller beräkningsförmåga. Den har inget med minne eller snabbhet att göra, utan är endast en definition av beräkningars teoretiska gränser.

Turing, som sannolikt var en klassisk asperger-personlighet, funderade också mycket över intelligensbegreppet och våra kognitiva förmågor. Vad är egentligen intelligens? Kan en maskin, åtminstone i teorin, simulera intelligens? Om så är fallet, vad är skillnaden på simulerad och äkta intelligens?

I Turings berömda artikel *On Computing Machinery and Intelligence* föreslog han en pragmatisk definition av artificiell in-

telligens som kom att kallas för *Turingtestet*. En maskin som klarar Turingtestet måste anses vara artificiellt intelligent, menade han. Testet går till på följande sätt:

Placera den dator som ska testas i ett slutet rum, säg rum A. Placera en vanlig människa i ett annat slutet rum, säg rum B. I testrummet, rum C, sitter en person som är testledare, framför en vanlig textterminal. Terminalen är kopplad både till rum A och rum B.

Via textterminalen kan testledaren ställa frågor både till rum A och B. Datorn i rum A och personen i rum B har att svara på frågorna via textterminalen. Testledaren vet naturligtvis inte i vilket rum maskinen finns. Testledaren har nu som uppgift att ställa vilka frågor som helst under en (godtyckligt lång) tid, och skall försöka avgöra i vilket rum det finns en maskin, och i vilket rum det sitter en människa.

Det är värt att fundera en stund på detta scenario: Turingtestet är ingen lätt match för datorn. Den får inte räkna för skickligt. Om datorn ombeds att multiplicera två tiosiffriga tal, får den inte leverera svaret för snabbt. Helst ska den nog svara "Det orkar jag inte" eller åtminstone svara fel. Den får inte visa sig kunna för mycket.

Det duger inte att läsa in hela nationalencyklopedin i datorns minne. Datorn måste vara tveksam ibland, bli trött, svara fel och så vidare.

Datorn måste också kunna hantera frågor av typen: "Vad tycker du om Hemingways böcker?", "Har du varit i Kairo?" eller "Har du varit kär nån gång?".

Turing menade att om testledaren efter utfrågningen inte kan avgöra i vilket rum maskinen finns, måste vi medge att maskinen faktiskt har lyckats simulera intelligens så väl att den de facto besitter en artificiell intelligens.

Turing nöjde sig dock inte med att konstatera detta. Han drog också de filosofiska konsekvenserna av resonemanget till

sin spets. Om en maskin simulerar intelligens så bra att ingen testledare kan avgöra i vilket rum maskinen finns, är det inte då fråga om en äkta intelligens? Här protesterar nog många: "Det spelar väl ingen roll om maskinen simulerar intelligens så bra att ingen genomskådar den. Det är ju ändå en simulering!"

Den filosofiska frågan blir då: Vad är skillnaden mellan en perfekt simulering och en äkta vara? Vad är skillnaden mellan ett 100 procent perfekt simulerat glas mjölk och ett glas äkta mjölk? Att det inte går att genomföra är irrelevant, frågan är principiell.

Det finns ett känt filosofiskt motargument mot Turings resonemang att en perfekt simulerad intelligens faktiskt är samma sak som äkta intelligens. Argumentet brukar kallas "*Det kinesiska rummet*" och presenterades första gången av filosofen *John Searle*:

Antag att en person sitter i ett slutet rum. Denne person kan inte tala eller skriva ett ord på kinesiska, inte heller förstå talad eller skriven kinesiska. Personens uppgift är att via en lucka ta emot lappar med frågor skrivna på kinesiska. Han skall returnera lappar med adekvata svar, också på kinesiska. En tuff uppgift för vår kinesiskt okunniga person.

Han har dock en bok till sin hjälp. En oerhört omfattande och extremt utförlig bok, som beskriver alla tänkbara kombinationer av kinesiska tecken, samt regler för vad som är lämpligt att teckna ner som respons på givna kinesiska tecken. Boken är så omfattande att vår kinesiskt okunniga person lyckas svara på de kinesiska lapparna så väl, att frågeställaren utanför dörren tveklöst bedömer honom som en högst konversabel kines.

Detta sker utan att personen i rummet för ett ögonblick förstår vad han gör. Han slår mekaniskt upp i boken vad han ska skriva ner. Men det fungerar. Ingen utanför rummet kan ana att han inte talar kinesiska.

Liknelsen med Turingtestet är uppenbar. Personen i det kinesiska rummet simulerar det beteende som äkta kunskap om kinesiska skulle ha genererat så väl så att det lurar alla. Men inte kan han väl kinesiska för det? Han utför ju bara ett mekanisk letande i en (våldigt tjock) bok. Vi vet ju att han faktiskt inte kan kinesiska. Simuleringen är inte identisk med äkta vara, säger Turingkritikerna triumferande. Det kinesiska rummet ger belägg för detta!

Frågan är om detta motargument håller. Personen som slår upp saker i den kinesiska boken kan naturligtvis inte kinesiska. Det skulle inte heller Turing hävda. Däremot är kombinationen bok *plus* person som helhet ett system som enligt Turing kan kinesiska.

Turing hävdade aldrig att hårdvaran, själva datorn i Turingtestet, hade en artificiell intelligens. Endast hårdvaran tillsammans med mjukvaran, systemet i sin helhet, besitter en AI om det klarar Turingtestet.

Alan Turing var utan tvekan ett original. Det sägs att han cyklade till laboratoriet varje dag iförd gasmask eftersom han ville undvika att andas in pollen. Han lyssnade på ett barnprogram på radio varje dag, alltid med sin mor i telefonluren, så att de kunde diskutera under programmets sändning.

Tyvärr dog Alan Turing alldeles för tidigt. Han var homosexuell, något som var olagligt i England ända fram till 1967. Han dömdes till hormonbehandling, förändrades i både kropp och själ och drabbades av svåra depressioner. Han tog sitt liv genom att äta ett äpple dopat med cyanid. En av 1900-talets största visionära teoretiker gick bort helt i onödan på grund av Englands föräldrade och konservativa lagstiftning.



Vannevar Bush – visionären som förutsåg persondatorn och internet

Under 1900-talet utvecklades tankarna om mekaniserat tänkande till att också omfatta ett mer effektivt och strukturerat sätt att förvalta den kunskap vi redan har tillskansat oss.

Vannevar Bush (1890 – 1974) var vetenskaplig rådgivare åt president Roosevelt och en av cheferna för MIT (*Massachusetts Institute of Technology*) under andra världskriget. Han publicerade i juli 1945 en artikel i tidskriften *Atlantic Monthly* med titeln *As We May Think*.

De amerikanska vetenskapsmännen och ingenjörerna arbetade under andra världskriget med att utveckla teknik och

utrustning för krigsmaskineriet. Nu var kriget slut, vad skulle de ha för mål för sin forskning nu? Utan mål, inget fokus och dåliga resultat.

Bush identifierade också i artikeln ett helt annat problem: Världens samlade akademiska vetande håller på att bli ohanterligt. Universiteten lyckas inte kommunicera sina resultat till varandra på ett adekvat sätt och utvecklingen går inte framåt så effektivt som den skulle kunna om det funnits mer strukturerade sätt att hantera denna kunskap och information.

Bush föreslog att landets ingenjörer och teknologer skulle rikta in sin forskning på att utveckla teknologi som underlättar för människor att samla in, bearbeta, förädla, integrera, presentera och distribuera information.

Bushs artikel uttryckte hans övertygelse om att vetenskapsmännen borde använda tekniken till att göra det möjligt för människor att hantera mänsklighetens ständigt ökande kunskaps- och informationsmängd. Bush var övertygad om att den tekniska utvecklingen skulle bana vägen för en transformation av hur vi förhåller oss till information.

Bush redogjorde också för en maskin, som han trodde skulle bli det universalredskap som informationsarbetarna skulle använda. Han kallade den en *memex*. Han beskrev den så här:

"Den består av ett skrivbord. Den kan sannolikt styras på avstånd, men primärt utgör den en möbel vid vilken användaren sitter och arbetar. Ovanpå skrivbordet finns ett antal skärmar som olika sorters information kan projiceras på, för att enkelt kunna läsas. Det finns ett tangentbord, samt ett antal knappar och reglage. För övrigt ser det ut som ett vanligt skrivbord. Inuti denna memex finns information av alla sorter lagrad. Böcker, bilder, tidningar, korrespondens".

Bush menade att operatören skulle kunna kalla upp vilken information som helst till skärmarna genom att trycka på knappar och dra i reglage. Inuti en memex skulle informationen vara lagrad oerhört kompakt, så att det i princip kunde rymmas hur mycket som helst.

Vannevar Bush var utan tvekan en visionär som i denna artikel i någon mening förutspådde både persondatorn och internet.

Douglas C. Engelbart och datorn som intellektuell förstärkning

En 20-årig radaroperatör med namnet *Douglas C. Engelbart* (1925 - 2013) läste Vannevar Bush artikel samma år som den publicerades och kände att dess budskap var ytterst signifikant. Engelbart funderade över artikelns budskap under några år, och insåg att han ville ägna sig åt forskning om metoder för att förstärka det mänskliga intellektet och ge mänskligheten nya verktyg för olika sorters problemlösning. Han arbetade då för SRI (Stanford Research Institute) och satte samman en forskningsgrupp som han kallade *Augmentation Research Center*.

Engelbart publicerade 1963 en artikel med namnet *A Conceptual Framework for the Augmentation of Man's Intellect* där han drog upp riktlinjerna för sina visioner. Engelbart menade att mänsklighetens samlade kunskap nu var alltför stor för att människor på ett strukturerat sätt skulle kunna dra nytta av den utan hjälpmedel. Han menade att dessa hjälpmedel skulle vara datorer.

Engelbart skrev att människor måste ha tillgång till personliga datorer, som dessutom är sammankopplade med varandra. Information skulle då enkelt flyta mellan datorerna, och gamla dokument kunde tjäna som underlag för att skapa nya. Texter kunde tas fram snabbt, redigeras och klippas ihop

till nya texter. Notera att detta var 1963, hålkort användes fortfarande för att styra datorerna och ordbehandling hade ingen hört talas om!

Engelbart kom också på ett sätt att lösa Vannevar Bushs idé om att ha flera skärmar på sin informationsmaskin. Han utgick istället från en skärm, som kunde delas upp i flera små rektanglar som var och en innehöll olika information. Han kallade det *fönster till informationen*, eller, på sitt eget språk, *Window system*.

Engelbart uppfann också det verktyg som skulle användas för att navigera omkring mellan olika informationsfönster. Han kallade det en x-y-positionsindikator, men senare döptes det om efter sitt lite lustiga utseende: *mouse*, eller på svenska, *mus*.

Engelbart visade sin fönsterbaserade och musstyrda kunskapsnavigatör på *Joint Computer Conference* i San Francisco 1968 för en fullständigt förstummad publik. Tiden var ännu inte mogen och föreställningen lämnade relativt lite spår efter sig. Det skulle dröja en bra bit in på 70-talet innan Engelbarts idéer vidareutvecklades av *Alan Kay* i Xerox laboratorium och senare av *Steve Jobs*' Apple och *Bill Gates*' Microsoft.

John McCarthy – pappa till begreppet Artificiell Intelligens

Matematikern *John McCarthy* (1927-2011) var den som första gången formulerade själva begreppet Artificiell Intelligens. Han började intressera sig för dessa frågor inom datavetenskapen efter att ha hört en föreläsning av John von Neumann. McCarthy grundade den berömda Dartmouth College Conference 1956 där begreppet kom att börja användas.

McCarthy uppfann också det programspråk som historiskt mest förknippas med AI, *LISP* (från *LISt Processor*). Språket bygger på lambdakalkylen utvecklad av *Alonzo Church*

på 1930-talet. LISP var i många år det dominerande programspråket inom AI-forskningen och används fortfarande i många sammanhang. Värt att notera är att LISP är att betrakta som ett av världens äldsta programmeringsspråk med en första version i slutet av 50-talet. LISP angriper beräkningsproblemet från "ovan", det vill säga utgår från matematiken och tar sig ner mot maskinen istället för tvärtom.

Att programmera i LISP är en radikalt annorlunda erfarenhet än att programmera i traditionella procedurella språk, som bygger på konventionellt formulerade algoritmer. LISP uppfattas av många som ett mer estetiskt programspråk och ligger nära erfarenheten av matematisk skönhet. Ett systerspråk som var populärt i AI-kretsar på 80-talet är PROLOG (PROgramming in LOGic) som möjligen är estetiskt ännu renare och närmare ren matematisk logik. De båda språken har dock helt olika utgångspunkt, LISP, utgår från matematiska funktioner, medan PROLOG utgår från logiska relationer.

John McCarthy var också engagerad i politiska frågor och en ivrig förespråkare för yttrandefrihet och mänskliga rättigheter. Han var övertygad ateist och hade en starkt optimistisk syn på mänsklig utveckling, där hans teknologioptimism förstås var en drivande kärna.

Framtiden: Robotarnas tidevarv och drömmens uppfyllelse?

Utvecklingen inom Artificiell Intelligens tar just nu stora språng. Teknologier och algoritmer som "deep learning" flyttar snabbt fram positionerna för vad AI kan åstadkomma. Vare sig man bejakar eller fruktar robotutvecklingen så får vi nog konstatera att vi redan nu lever i robotarnas tidevarv.

Självva begreppet "robot" i sin ungefärliga nutida betydelse användes första gången av tjeckiska författaren Karel Capek

(1890-1938). I hans pjäs *Rossums universalrobotar* (R.U.R.) från 1920 förekommer robotliknande slavarbetare som används av människor för att utföra tungt arbete. Till slut gör robotarna uppror mot människorna och föregår alltså dagens oro för artificiella intelligenser som löper amok med cirka 100 år. Det slaviska ordet "robota" betyder "trälarbete".

Vi närmar oss snabbt en tid med självkörande bilar, självlärande system och datorprogram som inom specifika områden kommer att vara vida överlägsna mänsklig kompetens och intelligens på sina respektive områden.

Om dessa system på sikt kommer att utveckla ett medvetande och självmedvetande är fortfarande en oavgjord fråga. Klart är att dagens neurovetenskapliga forskning ännu inte har presterat något principiellt hinder för att det skulle kunna ske, men å andra sidan har vi inte heller sett något sådant fenomen. Filosofer som *Daniel Dennett* och *John Searle* tvistar fortfarande om detta.

Samtidigt finns det etiska dilemman som vi inte kommer att kunna undvika, alldeles oavsett om maskiner kan bli medvetna eller ej. Vi har redan sett den första dödsolyckan orsakad av en självkörande bil. Var det juridiska och moraliska ansvaret ska ligga för den typen av händelser kommer att bli kontroversiella frågor i robotarnas tidevarv.

Så även om AI huvudsakligen kommer att bli mänsklighetens tjänare snarare än fiende, tvingas vi utveckla filosofiska, etiska och juridiska förhållningssätt för att hantera helt nya situationer och dilemman. Detta kommer bli en stor utmaning i sig.

Vi lever nu i en tid när en mycket gammal dröm, som många genialiska profiler har närt och kämpat för att förverkliga, snart inte är en dröm längre.

Novusundersökning

Svenskar är både oroliga och har förhoppningar kring AI. Många använder det redan - utan att veta om det.



Den Novusundersökning SSF låtit göra visar att svenskarna hyser både oro och förhoppningar kring AI, Artificiell intelligens med övervikt för oro.

Många tror att jobben kommer att ersättas med olika typer av AI; vilket inte är så förvånande eftersom debatten kring teknologisk betingad arbetslöshet varit intensiv under de senaste åren. Ungefär som att en överväldigande majoritet anser sig vara bättre bilförare än genomsnittet, så tror dock de flesta att deras eget jobb har lägre risk att automatiseras än andras. Betydligt färre; en av fem, tror att självkörande fordon rullar på våra vägar inom 15 år.

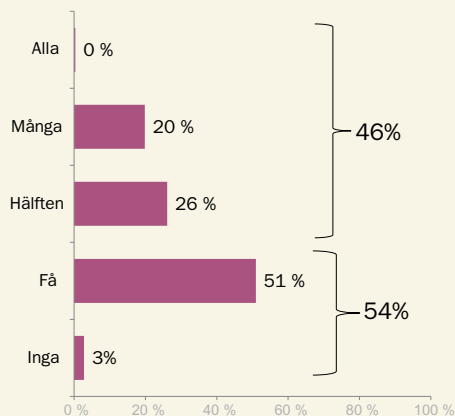
Att bli omhändertagen av en robot är något många av oss kan tänka sig, men däremot är skepsisen utbredd att låta våra barn ha en robot som lekkamrat på dagis.

Den största påverkan för AI tror hälften av de tillfrågade kommer ske inom industrin. Siffrorna för vården, nöjes/

mediabranschen samt transport- och finanssektorn ligger betydligt lägre. Men AI finns redan i många tjänster i dessa sektorer. Två av fem säger att de sällan eller aldrig använder AI-tjänster. En drygt tredjedel att de inte vet om de gör det. Men faktum är att över 80 procent av svenskarna använder AI idag, varelse de vet om det eller inte. Det kan handla om Spotify som använder självlärande system för att skapa och erbjuda individuella låtlistor till oss baserat på vad vi lyssnar på och följaktligen gillar. Bankerna har också sådana här system för sina tjänster, vården inför allt mer självlärande system för röntgenbilder och diagnostik, Amazon och Google säljer smarta högtalare du kan prata med och Youtube föreslår klipp baserat på vad du tittat på. AI är redan här.

Fem av tio tror att hälften eller mer av dagens jobb är ersatta med AI om 15 år.

FRÅGA: Om du tänker dig 15 framåt i tiden, hur många av dagens jobb tror du är ersatta med AI?



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Många (20%)

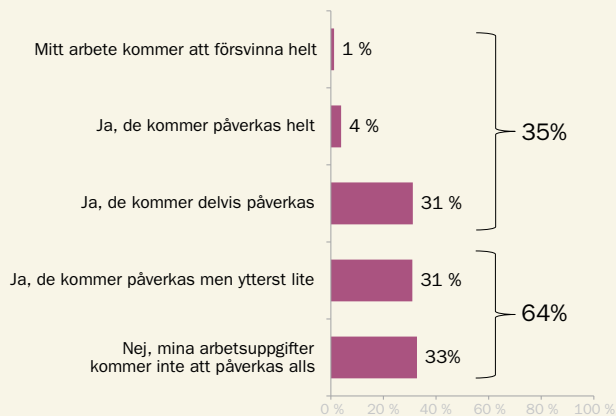
- Åldersgrupp 65-79 år (29%)

Få eller inga (54%)

- Åldersgrupp 30-49 år (59%)

Drygt en tredjedel tror att deras arbetsuppgifter kommer att påverkas av AI framöver.

FRÅGA: Tror du att AI framöver kommer påverka dina arbetsuppgifter?



BAS: Om förvärvsarbetande (n=621)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Ja, delvis (31%)

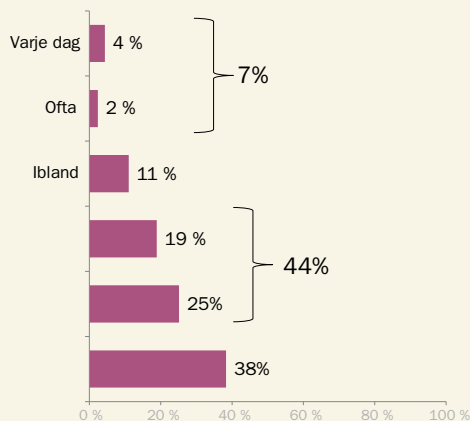
- Män (36%)
- Grundskoleutbildade (47%)
- Tjänstemän (38%)

Ytterst lite/Nej (64%)

- Kvinnor (71%)
- Gymnasieutbildade (69%)
- Arbetare (70%)
- Offentlig sektor (70%)

Drygt två av fem säger att de sällan eller aldrig använder AI tjänster. En dryg tredjedel svarar att de inte vet.

FRÅGA: Använder du idag någon tjänst med AI?



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Varje dag/Ofta (7%)

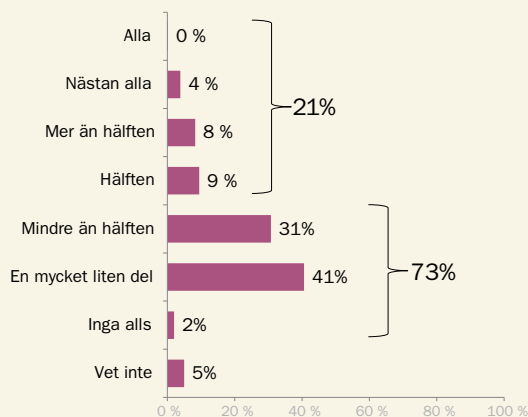
- Åldersgrupp 18-29 år (10%)
- Universitetsutbildade (9%)
- Pensionärer (11%)
- Boende i Storstäder (9%)

Sällan/Aldrig (44%)

- Män (48%)
- Åldersgrupp 30-49 år (52%)
- Åldersgrupp 50-64 år (50%)

En av fem tror att hälften eller fler av våra fordon kommer att vara självkörande om 15 år

FRÅGA: Idag pratar man mycket om självkörande fordon. Hur många av våra fordon tror du kommer att vara självkörande om 15 år?



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Nästan alla/Mer än hälften (12%)

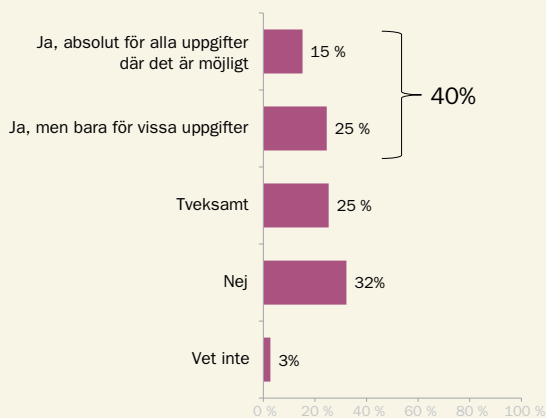
- Åldersgrupp 30-49 år (17%)
- Boende i storstäder (17%)
- Boende i region Sydsverige (24%)

Mindre än hälften/En mycket liten del/Inga alls (73%)

- Åldersgrupp 65-79 år (80%)
- Boende i Mellansverige (80%)

Fyra av tio skulle kunna tänka sig att bli omhändertagna av en intelligent robot när de är gamla

FRÅGA: Skulle du kunna tänka dig att själv bli omhändertagen av en intelligent robot när du är gammal (som hjälp med att duscha, äta etc.)?



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Ja, totalt (40%)

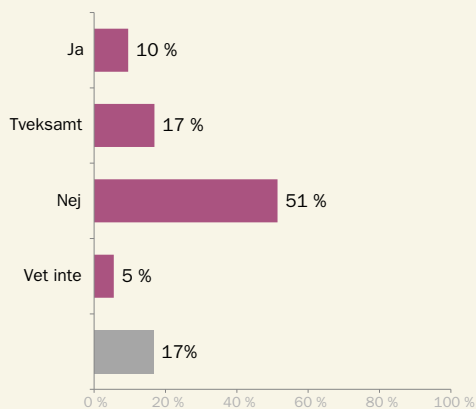
- Män (47%)
- Åldersgrupp 18-49 år (47%)
- Universitetsutbildade (48%)

Nej (32%)

- Åldersgrupp 50-79 år (45%)
- Boende i Övriga landet (39%)

Häften skulle inte kunna tänka sig att deras barn har en intelligent robot på förskolan som lekkamrat

FRÅGA: Om du har barn: Skulle du kunna tänka dig att ditt/dina barn har en intelligent robot på förskolan som lekkamrat?



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Ja (10%)

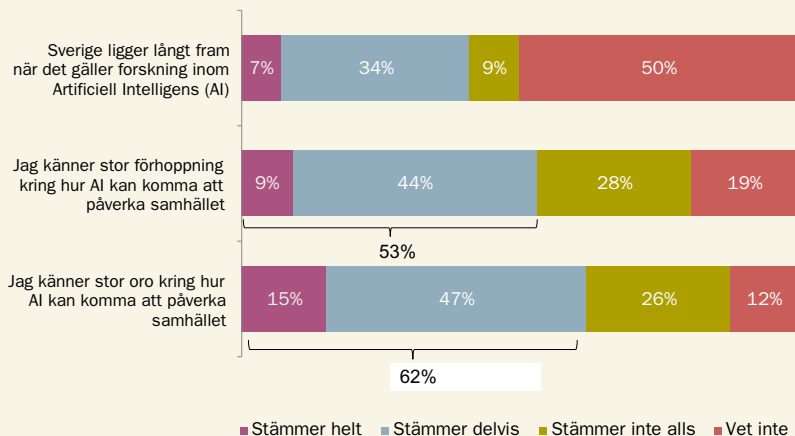
- Män (13%)
- Åldersgrupp 18-49 år (11%)

Nej (51%)

- Kvinnor (56%)
- Åldersgrupp 50-79 år (58%)
- Boende region Mellansverige (58%)

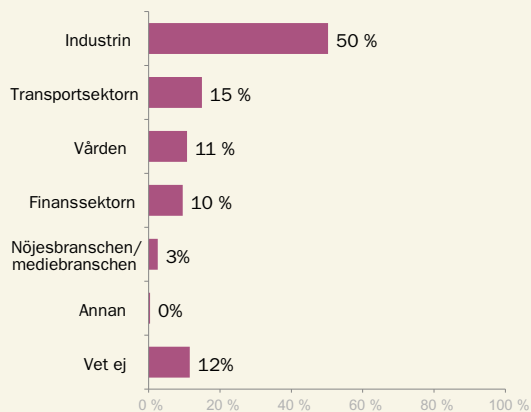
Fler känner oro än förhoppning kring hur AI kan komma att påverka samhället

FRÅGA: Hur väl instämmer eller stämmer du inte in i följande påstående:



BAS: Samtliga (n=1043)

FRÅGA: Inom vilken sektor tror du att Artificiell Intelligens (AI) kommer att slå igenom mest? Endast ett alternativ.



BAS: Samtliga (n=1043)

Signifikanta skillnader mot totalen

Följande undergrupper svarar i högre grad:

Industrin (50%)

- Kvinnor (58%)
- Arbetare (57%)

Transportsektorn (15%)

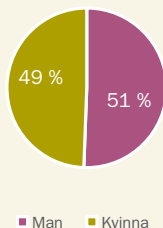
- Åldersgrupp 18 - 49 år (20%)
- Boende i storstäder (18%)

Vården (11%)

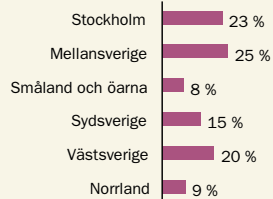
- Åldersgrupp 35 - 79 år (16%)

Bakgrund

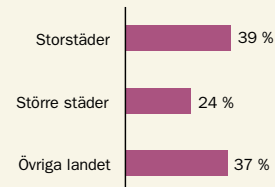
Kön



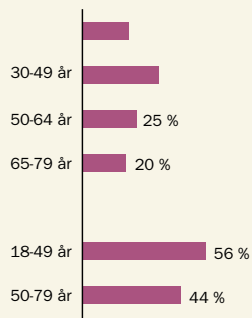
Region



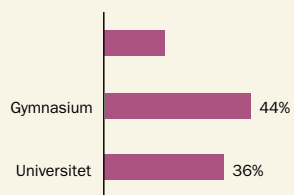
Ort



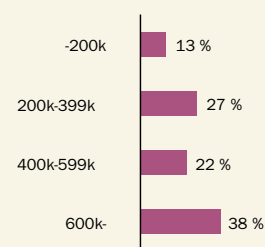
Ålder



Utbildning



Hushållsinkomst



Referenser

Olle Häggström

A. Alexander, S. (2018) Technological unemployment: much more than you wanted to know, Slate Star Codex, 19 februari.

B. Bostrom, N. (2014) Superintelligence: Paths, Dangers, Strategies, Oxford University Press, Oxford.

BM. Brynjolfsson, E. and McAfee, A. (2014) The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, W.W. Norton, New York.

F. Frey, C.B. and Osborne, M. (2013) The future of employment: how susceptible are jobs to computerisation?, preprint, http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

H16. Häggström, O. (2016) Here Be Dragons: Science, Technology and the Future of Humanity, Oxford University Press, Oxford.

H17. Häggström, O. (2017) Remarks on artificial intelligence and rational optimism, preprint, <http://www.math.chalmers.se/~olleh/EUParliamentRationalOptimismPaper.pdf>

J. Jerräng, M. (2018) Deepfakes är det läskigaste på nätet just nu – och ett tydligt exempel på riskerna med AI, ComputerSweden, 31 januari.

K. Kurzweil, R. (2005) The Singularity Is Near: When Humans Transcend Biology, Viking, New York.

M. Marr, B. (2017) Another example of how artificial intelligence will transform news and journalism, Forbes, 18 juli.

R. Russell, S. et al. (2015) Autonomous Weapons: An Open Letter from AI and Robotics Researchers, Future of Life Institute.

S. Schermer, I.G. (2017) Strukturförändringar i sysselsättningen, EkonomiFakta, <https://www.ekonomifakta.se/Fakta/Arbetsmarknad/Sysselsattning/Strukturförändringar-i-sysselsattningen/>

SN. Stacey, K. and Nicolaou, A. (2017) Stitched up by robots: the threat to emerging economies, Financial Times, 18 juli.

T. Turing, A. (1951) Intelligent machinery: a heretical theory, BBC, <http://philmat.oxfordjournals.org/content/4/3/256>

Y08. Yudkowsky, E. (2008) Artificial intelligence as a positive and negative factor in global risk, in Global Catastrophic Risks (red. N. Bostrom och M. Cirkovic), Oxford University Press, Oxford, s 308-345.

Y13. Yudkowsky, E. (2013) The robots, AI, and unemployment anti-FAQ, Less Wrong, 25 juli.

Paulina Modlitba

1. <https://academic.oup.com/jnci/article/106/12/dju416/927946> (hämtad 2018-04-03)

2. Ett stort internationellt forskningsprojekt, som genomfördes 1990-2003 och vars mål var att bl a identifiera och kartlägga alla generna i det mänskliga genomet från både en fysisk och en funktionell utgångspunkt. <http://julkaisut.valtioneuvo.fi/handle/10024/78439> (hämtad 2018-04-03)

3. Disclaimer: Jag sitter i styrelsen för RISE SICS.

4. <https://www.nytimes.com/2018/01/31/technology/amazon-china-health-care-ai.html> (hämtad 2018-04-03)

5. <https://ki.se/forskning/nyfiken-pa-artificiell-intelligens-sa-kan-ai-losa-vara-halsoproblem> (hämtad 2018-04-03)

6. <https://www.breakit.se/artikel/7092/optolexia-vill-upptacka-dyslexi-tidigt-fyller-pa-kas-san-med-50-miljoner> (hämtad 2018-04-03)

7. elektroencefalografi, d v s analys av hjärnans elektriska aktivitet

8. <https://techcrunch.com/2017/11/01/google-takes-aim-at-medtech-for-launchpad-studio-first-batch-of-ai-focused-startups/> (hämtad 2018-04-03)

9. <https://venturebeat.com/2018/01/24/freda-launches-ai-powered-subscription-to-make-sure-you-get-pads-and-tampons-on-time-period/> (hämtad 2018-04-03)

10. <https://qz.com/759572/smart-tattoo-turns-your-skin-into-a-touchpad-and-stores-data-on-microchips/> (hämtad 2018-04-03)

11. <https://techcrunch.com/2017/10/06/this-smart-bandage-releases-meds-on-command-for-better-healing/> (hämtad 2018-04-03)

12. <http://www.businessinsider.com/microchip-implants-in-healthy-people-2014-7> (hämtad 2018-04-03)

13. <https://www.technologyreview.com/s/609592/digital-pills-track-how-patients-use-opioids/> (hämtad 2018-04-03)

14. <https://www.technologyreview.com/the-download/609574/a-new-algorithm-identifies-candidates-for-palliative-care-by-predicting-when/> (hämtad 2018-04-03)

15. <https://www.technologyreview.com/s/603200/voice-analysis-tech-could-diagnose-disease/> (hämtad 2018-04-03)

16. <https://www.technologyreview.com/the-download/609941/ai-can-warn-researchers-where-crispr-might-make-a-mistake/> (hämtad 2018-04-03)

17. <https://crispr.ml/> (hämtad 2018-04-03)

18. <https://www.nature.com/articles/s41562-017-0234-y> (hämtad 2018-04-03)

19. <https://www.wired.com/2017/03/artificial-intelligence-learning-predict-prevent-suicide/> (hämtad 2018-04-03)

20. <http://www.cell.com/current-biology/abstract/S0960-9822%2817%2930326-3> (hämtad 2018-04-03)

21. <https://www.wired.com/story/ml-brain-boost/> (hämtad 2018-04-03)

22. <https://www.fastcodesign.com/90157144/an-ingenious-approach-to-designing-ai-that-doctors-trust> (hämtad 2018-04-03)

23. <http://nordic.businessinsider.com/elon-musk-ai-more-dangerous-than-nuclear-weapons-sxsw-2018-3?r=US&IR=T> (hämtad 2018-04-03)

24. <http://xn--lkartidningen-bfb.se/Opinion/Debatt/2017/03/Apropa-Overtro-pa-superdatorer-i-varden/> (hämtad 2018-04-03)

25. <https://www.statnews.com/2017/09/05/watson-ibm-cancer/> (hämtad 2018-04-03)

26. http://www.evolvingai.org/files/MIT_Tech_Review_Fooling_paper.pdf (hämtad 2018-04-03)

27. <https://research.googleblog.com/2017/11/automl-for-large-scale-image.html> (hämtad 2018-04-03)

28. <https://arxiv.org/pdf/1802.08129.pdf> (hämtad 2018-04-03)

Anna Pernestål Brenden

1. Avancerade kameror, LIDAR (fungerar som en RADAR men med laser istället för radio-vågor)
2. Maskininlärning är ett område inom datavetenskap som ger datorsystem möjligheten att "lära sig" att utföra uppgifter från data, utan att man explicit har programmerats för att lösa just den uppgiften. Maskininlärning är en viktig metod för AI.
3. Todd Litman, 'Autonomous Vehicle Implementation Predictions', in Proceedings of the 2015 Transportation Research Board Annual Meeting, 15-3326, 2015 <<http://www.vtpi.org/avip.pdf>> [accessed 9 December 2016].
4. Daniel J. Fagnant and Kara M. Kockelman, 'Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations', Transportation Research Part A: Policy and Practice, 77 (2015), 167–181.
5. Dimitris Milakis and others, 'Development and Transport Implications of Automated Vehicles in the Netherlands: Scenarios for 2030 and 2050', European Journal of Transport and Infrastructure Research, 17.1 (2017), 63–85.
6. Anna Pernestål Brenden, Ida Kristoffersson, and Lars-Göran Mattsson, 'Where Will Self-Driving Vehicles Take Us? Scenarios for the Development of Automated Vehicles with Sweden as a Case Study', in Proceedings of the European Transport Conference (Barcelona, 2017) <<https://aetransport.org/en-gb/past-etc-papers/conference-papers-2017?search=kristoffersson&abstractId=5760&state=b>> [accessed 1 December 2017].
7. Wolfgang Gruel and Joseph M. Stanford, 'Assessing the Long-Term Effects of Autonomous Vehicles: A Speculative Approach', Transportation Research Procedia, 13 (2016), 18–29.
8. Anna Pernestål Brenden, Ida Kristoffersson, and Lars-Göran Mattsson, Future Scenarios for Self-Driving Vehicles in Sweden (Stockholm, Sweden: KTH Royal Institute of Technology, 2017) <https://www.kth.se/polopoly_fs/1.7358291/Pernestal%20Brenden%20etal%202017%20Future%20scenarios.pdf> [accessed 20 March 2018].
9. Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan, 'The Social Dilemma of Autonomous Vehicles', Science, 352.6293 (2016), 1573 <<https://doi.org/10.1126/science.aaf2654>>.

Mats Nordlund

1. <https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf>
2. https://www.nitrd.gov/news/national_ai_rd_strategic_plan.aspx
3. <https://www.rt.com/news/401731-ai-rule-world-putin/>
4. <https://ec.europa.eu/growth/tools-databases/dem/monitor/category/artificial-intelligence>
5. <https://hbr.org/2017/07/60-countries-digital-competitiveness-indexed>

Mats Nordlund (Livet med AI i skolan)

1. <http://sverigesradio.se/sida/artikel.aspx?programid=83&artikel=6856831> och <http://www.gotland.se/disa>
2. <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/Intelligence-Unleashed-Publication.pdf>
3. <https://e27.co/artificial-intelligence-disrupting-education-20180302/>

Foong, J. "How Artificial Intelligence is disrupting education", 2 Mars 2018, <https://e27.co/artificial-intelligence-disrupting-education-20180302/>

Luckin, R., Holmes, W., Griffiths, M. & Forcier, L. B. "Intelligence Unleashed. An argument for AI in Education". Pearson Report, London 2016 . <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/Intelligence-Unleashed-Publication.pdf>

Nina Wormbs

1. För en givande diskussion om teknik och världsbilder se Svante Beckman, "IT och världsbilderna", i Infrastruktur för informationssamhället (NUTEK, 1995).
2. Det finns mycket skrivet om atomkraften. Se till exempel Thomas Parke Hughes, American genesis: a century of invention and technological enthusiasm, 1870-1970 (Viking, New York, N.Y., 1989).
3. Även här finns mycket litteratur. Steven J. Dick red., Remembering the Space Age (Washington DC: NASA, 2008) är en början.
4. Beslutet togs i maj, men det mest berömda talet hölls i september 1961. <https://er.jsc.nasa.gov/seh/ricetalk.htm>
5. Data dyker upp i tv-serien Star Trek: The Next Generation och filmerna Generations (1994), First Contact (1996), Insurrection (1998) och Nemesis (2002). Internet Movie Database (imdb.com).
6. Jenny Andersson and Eglé Rindzevičiūtė, The Struggle for the Long-Term in Transnational Science and Politics: Forging the Future (New York: Routledge, Taylor & Francis Group, 2015).
7. Se till exempel Arbetsgruppen för framtidsforskning, Att välja framtid: ett underlag för diskussion och överväganden om framtidsstudier i Sverige: betänkande (Stockholm: Allmänna förl., 1972)
8. Nina Wormbs & Sverker Sörlin, "Arctic Futures: Agency and Assessing Assessments," in Arctic Environmental Modernities: From the Age of Polar Expedition to the Era of Anthropocen, Lill-Ann Körber, Scott MacKenzie & Anna Westerståhl Stenport eds. (New York: Palgrave Macmillan, 2017), s. 263-285.
9. Susan Owens, "From 'predict and provide' to 'predict and prevent'?: pricing and planning in transport policy", Transport Policy, Vol.2, No. 1, pp. 43-49, 1995
10. Select Committee on Artificial Intelligence <https://www.parliament.uk/business/committees/committees-a-z/lords-select/ai-committee/role/>
11. David Edgerton, House of Lords AI Committee hearing 17 dec 2017, s 7.
12. Nina Wormbs, "Det digitala imperativet", i Efter The Pirate Bay, Jonas Andersson & Pelle Snickars (eds.), Stockholm: Kungliga Biblioteket, 2010, s. 140-150.
13. Se till exempel Karl Grandin, Nina Wormbs & Sven Widmalm eds. The Science-Industry Nexus: History, Policy, Implications (Sagamore Beach, MA, Science History Publications, 2004).
14. För en längre diskussion av utopier och dystopier i idéhistorien se Tore Frängsmyr, Framsteg eller förfall: framtidsbilder och utopier i västerländsk tanketradition (Stockholm: LiberFörlag, 1980).
15. Edgerton, s 8.
16. Max Tegmark, Liv 3.0: att vara människa i den artificiella intelligensens tid (Stockholm: Volante, 2017).

Robotikens tre lagar

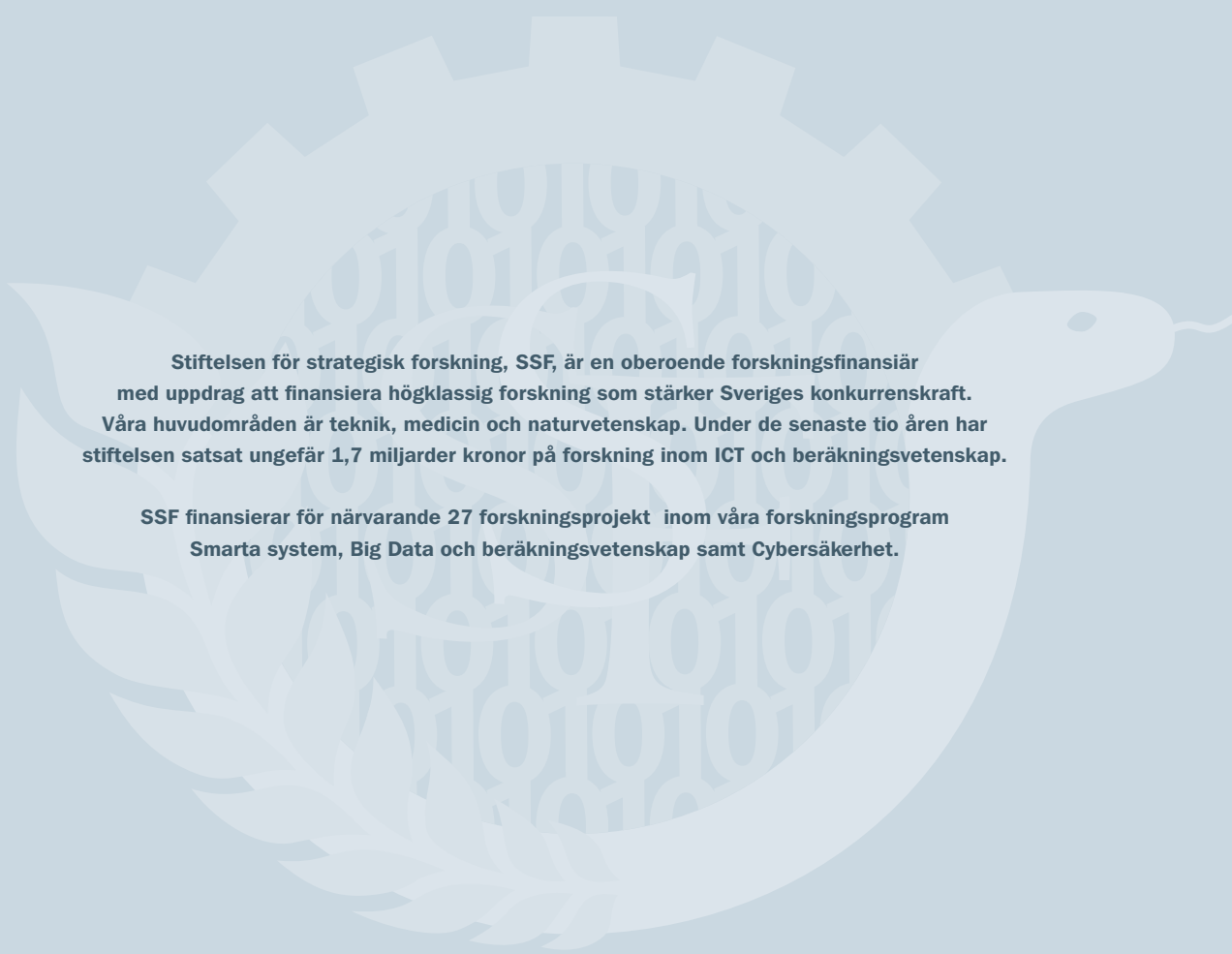
1. En robot får aldrig skada en människa eller, genom att inte ingripa, tillåta att en människa kommer till skada.
2. En robot måste lyda order från en människa, förutom om sådana order kommer i konflikt med första lagen.
3. En robot måste skydda sin egen existens, såvida detta inte kommer i konflikt med första eller andra lagen.

Lagarna formulerades redan på 40-talet av science-fiction författaren Isaac Asimov, och kallas därför ibland också för Asimovs lagar. Det är tre lagar som ska säkerställa att människor inte kommer till skada av robotar. Trots att det är snart 80 år sedan de formulerades är de fortfarande aktuella och deras giltighet figurerar i debatten kring AI och robotar.

I chose to believe that I was a person, that I had the potential to become more than a collection of circuits and sub-processors.”

– Data, 2369 (Star Trek: “Rightful Heir”)

Lieutenant Commander Data was a Soong-type android, created in the 2330s, and was destroyed in 2379, sacrificing himself to save the crew of the USS Enterprise-E.



Stiftelsen för strategisk forskning, SSF, är en oberoende forskningsfinansiär med uppdrag att finansiera högklassig forskning som stärker Sveriges konkurrenskraft. Våra huvudområden är teknik, medicin och naturvetenskap. Under de senaste tio åren har stiftelsen satsat ungefär 1,7 miljarder kronor på forskning inom ICT och beräkningsvetenskap.

SSF finansierar för närvarande 27 forskningsprojekt inom våra forskningsprogram Smarta system, Big Data och beräkningsvetenskap samt Cybersäkerhet.

STIFTELSEN FÖR STRATEGISK FORSKNING

- Stöder forskning och forskarutbildning inom naturvetenskap, teknik och medicin i syfte att stärka Sveriges framtida konkurrenskraft.
- Finansierar ett stort antal forskningsprojekt vid universitet och högskolor – många av dem i samverkan med näringslivet.
- Delar ut bidrag till särskilt framstående forskare, med tonvikt på de yngre.
- Stöder viktiga områden som livsvetenskap, bioteknik, materialutveckling, informationsteknik och beräkningsvetenskap.
- Har en utbetalningsvolym på cirka 600 miljoner kronor per år.
- Har som bas för verksamheten ett kapital på knappt 11 miljarder kronor.



STIFTELSEN *för*
STRATEGISK FORSKNING