

Fokusrapport – Artificiell intelligens och e-hälsa

Diarienummer: 2020/00180

Datum: 2020-11-24



Denna publikation skyddas av upphovsrättslagen. Citera gärna rapporten men uppge alltid källa: *Fokusrapport – Artificiell intelligens och e-hälsa, 2020, E-hälsomyndigheten.*

Publicerad: E-hälsomyndigheten, november 2020.

Diarienummer: 2020/00180

Adress: Södra Långgatan 60, Kalmar
Sankt Eriksgatan 117, Stockholm

E-post: registrator@ehalsomyndigheten.se

Telefon: 010-458 62 00

www.ehalsomyndigheten.se

Innehåll

Förord.....	5
1. Sammanfattning	6
2. Syfte med rapporten	7
3. Metod	8
4. Inledning.....	8
5. AI kräver investeringar.....	11
6. Vad är AI?	14
6.1 AI – definitioner och begrepp.....	15
6.1.1 Machine Learning	17
6.1.2 Deep Learning och neurala nätverk.....	17
7. Utmaningar och risker med AI	18
7.1 Exempel när det kan gå fel	18
7.1.1 Preventivapp och oönskade graviditeter	20
7.1.2 Diskriminerande chatbot	20
7.1.3 Problematisk ansiktsanalys.....	20
7.1.4 Självkörande bilar	21
7.1.5 Socioekonomisk och etnisk diskriminering.....	21
7.1.6 Sakernas internet (Internet of Things).....	21
7.2 AI inom e-hälsa – risker och etiska dilemman.....	23
7.3 Tillit och förtroende	25
7.4 Transparens och förklarbarhet.....	26
7.5 Tolerans och rättvisa (fairness)	27
7.6 Intelligens, etik och moral	28
7.7 Etiska riktlinjer	28
8. AI och e-hälsa	32
8.1 AI inom hälso- och sjukvård.....	32
8.2 AI inom socialtjänst.....	34
8.3 Hälsoappar	35
8.4 Covid-19 och AI	36
9. Juridiken.....	39

9.1 Regelverket i korthet.....	39
9.2 AI och medicintekniska regelverket	40
9.3 Automatiserade beslut.....	43
9.4 Ansvarsfrågan	44
9.5 Ny lagstiftning för AI-forskning?	45
10. Kompetensförsörjning.....	47
11. Frågor för fortsatt diskussion	48

Förord

Regeringen har tillsammans med Sveriges Kommuner och Regioner satt som målsättning för överenskommelsen Vision e-hälsa 2025 att Sverige ska vara bäst i världen på att ta vara på digitaliseringens och e-hälsans möjligheter. Därigenom underlättas för människor att uppnå en god och jämlik hälsa och välfärd.

I den uppdaterade strategin för genomförande av Vision e-hälsa 2025, *Nästa steg på vägen 2020-2022*, beskrivs utifrån de grundläggande förutsättningarna (Regelverk, Enhetligare begreppsanvändning, Standarder) fyra inriktningsmål för en ökad utvecklingstakt: Individens som medskapare, Rätt information och kunskap, Trygg och säker informationshantering, samt Utveckling och digital transformation i samverkan.

Artificiell intelligens (AI) inom e-hälsa har bäring på samtliga dessa inriktningsmål, även om AI omnämns utförligt endast under det fjärde målet. När det gäller målet Utveckling och digital transformation, särskilt avseende stöd för införande av nya teknologier, så anges att AI kan bidra till ökad kvalitet och effektivitet samt bättre arbetsmiljö. Samtidigt finns behov av att analysera grundläggande frågor om etik, säkerhet, transparens, förklarbarhet, samt representativitet i data som används för utveckling av digitala tjänster.

E-hälsomyndigheten ska enligt sitt instruktionsenliga uppdrag, samordna regeringens satsningar på e-hälsa samt övergripande följa utvecklingen inom e-hälsoområdet. AI ses av många som ett viktigt verktyg för att kunna möta framtidens behov av vård och omsorg. Samtidigt är användningen av detta verktyg inte helt okontroversiell. Ett flertal utmaningar och frågeställningar kan ses som hinder för en långsiktigt hållbar användning av AI inom e-hälsa. Detta är bakgrunden till att E-hälsomyndigheten har tagit fram föreliggande diskussionsunderlag för att bidra till ökad kunskap på området.

Denna rapport är framtagen av utredare Åke Nilsson samt enhetschef Michel Silvestri. Ett tack riktas till AI-rådgivare Göran Lindsjö samt biträdande professor och AI-forskare vid Linköpings universitet Fredrik Heintz för bidrag och synpunkter på innehåll.

Beslut om denna rapport har fattats av avdelningschef Annemieke Ålenius. Utredare Åke Nilsson och enhetschef Michel Silvestri har varit föredragande.

Annemieke Ålenius
Avdelningschef

Stockholm den 24 november 2020

1. Sammanfattning

Dagens hälso- och sjukvård samt socialtjänsten, genererar enorma mängder data. För att kunna skapa information och ny kunskap med hjälp av all denna data måste människan ha hjälp. Följaktligen växer behovet av att kunna nyttja artificiell intelligens (AI) till dessa resurskrävande och alltmer komplexa dataanalyser. Med AI ser vi framtida löften om effektivare vård och ökad patientsäkerhet, men vi ser också utmaningar.

Denna rapport utgör ett diskussionsunderlag genom att beskriva AI utifrån ett e-hälsoperspektiv, och i det sammanhanget lyfta ett antal frågor, främst om etik och juridik, vilka kan behöva besvaras innan vi har säkrat tillgången på långsiktigt hållbar AI.

Gränsen för vad som är AI och vad som inte är AI har genom decennierna varit något flytande. När vi i denna rapport diskuterar ”äkta AI” så syftar vi på autonoma och adaptiva system som utgår från lärande algoritmer för att komma fram till ett resultat. Detta till skillnad från statiska regler eller algoritmer för hantering av data, vilka är formulerade enbart av programmerare.

Även om AI funnits som begrepp i mer än ett halvsekel så är det fortfarande att se som i sin linda. Samtidigt är AI idag extremt uppmärksammat; utvecklingen går fort och det investeras stora belopp. Utveckling av framgångsrik AI kräver dock inte enbart ekonomiska resurser, utan också stora mängder data – relevant data av god kvalitet. Detta väcker frågor om vem som lagligen kan förfoga över data, framförallt när den utgörs av personuppgifter. Ytterst kan en avvägning behöva göras mellan samhällets nytta av tillgång på data och den enskilda individens integritetsskydd.

Andra utmaningar vid utveckling av AI är bland annat hur vi ska få en maskin, ett datasystem, att anamma mänsklighetens etiska principer och normer. Eller kan vi acceptera AI-beslut även om dessa framstår som obegripliga för oss?

En generell fråga när det gäller AI rör problematiken när något går fel med stora skador som följd, eller kanske rentav dödsfall - vem bär då ansvaret? Är det programmeraren, användaren, huvudmannen, leverantören eller någon annan?

För att AI ska uppfylla våra goda förhoppningar krävs en stadig grund för tillit och förtroende genom transparens och förklarbarhet, tolerans och rättvisa. Som stöd för utveckling av tillförlitlig AI har EU nyligen tagit fram etiska riktlinjer som slår fast att hela livscykeln måste präglas av etik, legalitet och robusthet. Europeiska kommissionen har presenterat ett samlat förslag som ska sätta ramarna och staka ut riktningen för den digitala omställningen. Det innehåller en övergripande digital strategi för EU, en strategi för användning av data och en så kallad vitbok för kommande åtgärder för utveckling av AI¹.

Redan idag finns lovande resultat av AI till exempel som beslutsstöd vid bedömning av röntgenbilder för diagnostik av bröstcancer eller frakturer. Samtidigt finns varnande

¹ Sveriges Riksdag Digital strategi. AI-vitbok och datastrategi. Fakta-PM om EU-förslag 2019/20:FPM23 COM (2020) 65, COM (2020) 66, COM (2020) 67. https://www.riksdagen.se/sv/dokument-lagar/dokument/fakta-pm-om-eu-forslag/digital-strategi-ai-vitbok-och-datastrategi_H706FPM23

exempel på hur AI eller AI-liknande system inom olika samhällssektorer har begått misstag. Dessa negativa exempel tydliggör nödvändigheten av att den data som används för AI måste vara relevant, strukturerad och tillräckligt omfattande samt att algoritmerna i tillräcklig grad tar hänsyn till etiska principer och likabehandling.

AI är som sagt beroende av stora mängder data. En snabbt växande källa av data globalt inhämtas av det som kallas ”Sakernas internet” (*Internet of Things, IoT*). All denna data kan användas i gott syfte och ge samhällsvinster, men samtidigt finns risk att sårbarheten ökar vilket visats genom exempel som att en känslig databas kan hackas via en uppkopplad digital akvarietermometer.

Vid närmare granskning av juridiken finns skillnader beroende på tillämpningsområde. Exempelvis tillåter förvaltningslagen automatiserade beslut, men det gör inte kommunallagen. En fråga som aktualiserats genom hemställan till regeringen är om kommunallagen behöver ändras av denna anledning.

När det kommer till det som här kallas ”äkta AI” så är Läkemedelsverkets tolkning av det medicintekniska regelverket att detta inte tillåter att systemets lärandefunktion är aktiverad vid klinisk användning inom sjukvården, eftersom kravet på validering är svårt att uppfylla av tillverkaren. Frågan är om de legala kraven i det medicintekniska regelverket i så fall hindrar optimal användning av AI, med dess förmåga till ständigt lärande, genom att läsa algoritmerna i olika definierade mjukvaruversioner.

Utveckling av framgångsrik, högkvalitativ och patientsäker AI är beroende av tillgång till kompetens. Det är avgörande att de programmerare, datavetenskapare och AI-ansvariga chefer, som tillsammans med vårdens professioner ska forma en långsiktigt hållbar AI inom e-hälsa, har utbildning och kunskap inte bara inom it, utan även inom etik och juridik.

Under arbetet med denna fokusrapport drabbades världen av en ny pandemi, covid-19, med stora konsekvenser för de flesta länder, inklusive Sverige. Det är för tidigt att redogöra för hur AI har kunnat användas på ett effektivt sätt mot denna pandemi men rapporten har kompletterats med en kortfattad överflygning för att visa på exempel hur AI kan användas i kampen mot pandemin och dess effekter.

2. Syfte med rapporten

E-hälsomyndighetens fokusrapporter ska beskriva och diskutera ett aktuellt område som har betydelse för e-hälsoutvecklingen i Sverige, och därigenom bidra till uppfyllandet av Vision e-hälsa 2025. Denna rapport är ett diskussionsunderlag som belyser övergripande etiska och juridiska frågor som är avgörande för en god AI-utveckling, i synnerhet de som uppstår inom hälso- och sjukvård och omsorg. Även andra viktiga aspekter berörs, såsom kompetensförsörjning inom området.

3. Metod

Informationsinhämtning till denna rapport har skett genom inläsning av litteratur, föreläsningar, intervjuer, interna samtal på E-hälsomyndigheten, samtal med extern expertis inom AI, samt medverkan i nätverksaktiviteter och i workshops.

4. Inledning

Denna rapport syftar till att belysa artificiell intelligens (AI)² utifrån ett etiskt och juridiskt perspektiv, främst vid användning inom hälso- och sjukvård samt omsorg (här som samlingsbegrepp e-hälsa). AI som verktyg för användning av så kallad Big Data är förknippat med stora förväntningar om effektivisering, kvalitetshöjning och ekonomiska vinster samt ökad livskvalitet för människor. Samtidigt är AI tätt kopplat till komplicerade frågor rörande etik, juridik och ansvar om något går fel.

Ett flertal utredningar och rapporter har publicerats senaste åren, rörande AI inom samhället och välfärden generellt, inom offentlig förvaltning och näringsliv. Regeringen gav 2018 ut *Nationell inriktning för artificiell intelligens* i syfte att Sverige ska vara bäst i världen på att använda digitaliseringens möjligheter och i internationell jämförelse ligger Sverige i absolut framkant.³ Regeringen fastslår att Utbildning, Forskning, Innovation och användning samt Ramverk och infrastruktur är viktiga förutsättningar för nyttiggörande av AI i Sverige. När det gäller det sistnämnda området så behöver Sverige bland annat utveckla regler, standarder, normer och etiska principer i syfte att vägleda etisk och hållbar AI. Likaså fortsätta arbetet med att tillgängliggöra data för AI, samt ta en aktiv roll i EU:s arbete med att främja digitalisering och AI.

Som en del av ovanstående inriktning fick SCB 2019 ett regeringsuppdrag att kartlägga användningen av AI respektive kartlägga analys av stora datamängder i företag, offentlig förvaltning inklusive universitets- och högskolesektorn. Denna utredning ska redovisas 30 november 2020.⁴

EU-kommissionens expertgrupp för artificiell intelligens publicerade i april 2019 riktlinjer för pålitlig AI och etiska frågor⁵. I Sverige arbetar Kommittén för teknologisk innovation och etik (Komet)⁶ för att skapa goda förutsättningar för innovation och konkurrenskraft samtidigt som utvecklingen och spridningen av ny teknik sker tryggt, säkert och med ett långsiktigt samhällsperspektiv.

² Artificiell intelligens (AI) (engelska: *Artificial intelligence*, även *machine intelligence*) är intelligens som uppvisas av maskiner, till skillnad från naturlig intelligens som uppvisas av människor och andra djur.

³ Regeringen (2018-05-06) <https://www.regeringen.se/informationsmaterial/2018/05/nationell-inriktning-for-artificiell-intelligens/>

⁴ Uppdrag att kartlägga användningen av artificiell intelligens respektive analys av stora datamängder i Sverige (Diarienummer: I2019/01964/D) (2019-08-21) <https://www.regeringen.se/regeringsuppdrag/2019/08/uppdrag-att-kartlagga-anvandningen-av-artificiell-intelligens-respektive-analys-av-stora-datamangder-i-sverige/>

⁵ Etiska riktlinjer för tillförlitlig AI. Oberoende expertgrupp på hög nivå för AI-frågor inrättad av Europeiska Kommissionen (2020-04-08) <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-sv/format-PDF>

⁶ Komet. En etisk diskussion behövs vid teknikutveckling. <https://www.kometinfo.se/uppdrag-och-fokus/etik/>

I anslutning till att digitaliseringen ökar i samhället så har också AI under de senaste åren blivit ett alltmer uppmärksammat område. Samtidigt är det troligt att vi bara sett början av utvecklingen. När det gäller hälso- och sjukvården samt omsorgen så berörs AI på ett eller annat sätt i samtliga fyra inriktningsmål i Vision e-hälsa 2025, även om det är mer utförligt endast i det fjärde målet som rör *Utveckling och digital transformation i samverkan*. Under rubriken *Artificiell intelligens inom vård, omsorg och socialtjänst* anges att i många kommuner och regioner pågår arbeten kring AI, även om det till stor del handlar om förstudier. ”Dagens tillgång till digital data och avancerad teknik skapar möjligheter att både efterlikna och ibland överskrida människans förmåga att ta in, bearbeta och använda information på ett intelligent sätt. Artificiell intelligens utvecklas i hög takt och påverkar individer, branscher och samhälle.”

Myndigheten för digital förvaltning (DIGG) har av regeringen fått i uppdrag, dels att analysera och ge förslag på hur man kan öka den offentliga förvaltningens förmåga att tillgängliggöra öppna data samt bedriva öppen och datadriven innovation, dels att främja den offentliga förvaltningens förmåga att använda artificiell intelligens. I Genomförandeplanen för Vision e-hälsa 2025 är DIGG rapporteringsansvarig för detta, i samverkan med andra statliga myndigheter och huvudmännen för vård och omsorg.⁷

I en delrapport 2019 presenterade DIGG en uppskattning att den ekonomiska potentialen för AI-stödda lösningar inom sjukvård och omsorg motsvarar 30 miljarder kronor per år. Detta skulle då gälla såväl AI-tillämpningar för förbättrad planering och resursfördelning i vården, som genomförande av bland annat AI-stödd diagnosticering och behandling. DIGG menar att den ekonomiska nyttan av AI i vården tillfaller till stor del den offentliga förvaltningen direkt i form av effektivisering, men skapar också en bättre vård till gagn för patienterna.⁸

Flera stora aktörer på marknaden investerar enorma summor i utveckling av AI. Det är välkänt att företag som tidigt anammar ny teknik snabbt kan bli världsledande inom sitt område. Detsamma gäller sannolikt företag som på ett framgångsrikt sätt använder AI. Vi har redan sett att tidigare marknadsledande företag har blivit omsprungna av de företag som nyttjar ny teknik såsom AI.

Jämfört med de flesta andra länder ligger Sverige ganska långt fram i teknisk utveckling, infrastruktur och digital mognad hos invånarna.^{9,10} Samtidigt tycks införande av AI i skarp användning gå långsamt. En faktor som kan påverka takten är synen på ägande av personlig data eller rättare sagt vem som får använda denna data, eftersom stora mängder relevant data är det som får AI-maskineriet att fungera på bästa sätt.

⁷ <https://ehalsa2025.se/om-e-halsa-2025/genomforandeplan/>

⁸ DIGG12019/01416/DF || I2019/01020/DF (delvis) Främja den offentliga förvaltningens förmåga att använda AI

<https://www.digg.se/globalassets/dokument/publicerat/publikationer/framja-den-offentliga-forvaltningens-formaga-att-anvanda-ai.pdf>

⁹ Enligt Internetstiftelsens *Svenskarna och internet 2019* har 98 procent av svenska hushåll tillgång till internet 2020 och 96 procent av befolkningen över 12 år använder internet.

<https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2019/>

¹⁰ Digitalt ekonomi- och samhällsindex (DESI) rankar Sverige till andra plats: Europeiska Kommissionen. Digital Economy and Society Index Report 2020 - Human Capital

<https://ec.europa.eu/digital-single-market/en/human-capital-and-digital-skills>

När det gäller innovation generellt visar *Global Innovation Index 2020*¹¹ att Schweiz är världens mest innovativa land följt av Sverige, USA, Storbritannien och Nederländerna. Sverige står sig alltså väl i internationell jämförelse av innovationskraft, men risken finns att olika faktorer (juridiska, samhällsorganisatoriska eller andra) gör att vi halkar efter när det gäller spridd tillämpning av exempelvis AI och utnyttjande av Big Data. USA och Kina är stora länder som investerar enorma belopp i AI-utveckling och implementering. I Kina finns dessutom få restriktioner kring insamling och lagring av persondata vilket kan ses som en stor fördel för kinesiska AI-företag.¹² Det kan diskuteras huruvida denna typ av konkurrens är ett problem eller om det kan finnas anledning att ”skynda långsamt” av etiska eller juridiska skäl och för att lära av andra länders erfarenheter. I en Vinnova-rapport beskrivs regulatorisk utveckling av spelregler för data och datatillgång som avgörande för AI-utvecklingen: ”Sådana spelregler måste balansera fundamentala behov av integritetsskydd, etik, tillit och samhällsskydd med sådan datatillgång som är nödvändig för utveckling av värdeskapande AI-tillämpningar.”¹³

USA:s regering genomförde hösten 2016 en noggrann utredning rörande AI och kom fram till att det inte behövdes särskild lagstiftning, men att man kontinuerligt behövde bevaka olika konsekvenser av AI. Flera andra har sedan dess påpekat att lagstiftning kring AI kan göras bättre genom att nära följa utvecklingen jämfört med att i förväg ”gissa” behov. En tolkning av detta kan vara att juridiken måste vara mer följsam för att inte utgöra ett hinder för utvecklingen, men samtidigt stå som garant för säkerhet och trygghet för den enskilde och för samhället. I Finland stiftades 2019 en lag om så kallad sekundäranvändning av personuppgifter från social- och hälsovård, där sekundäranvändning omfattar forskning t. ex inom AI-området (se avsnitt 9.5 Ny lagstiftning för AI-forskning?).¹⁴

För att uppnå en framgångsrik och långsiktigt hållbar användning av AI inom e-hälsa kan de etiska aspekterna användas som stöd när lagstiftning och regelverk behöver ändras för att bidra till utvecklingen. Denna rapport syftar till att belysa dessa aspekter, till exempel genom frågor som:

- Hur överför vi människans normer och etik¹⁵, synen på rätt och fel, till ett tekniskt grundat system som är både autonomt och självlärande?
- Kan vi tvingas till att utnyttja mindre av AI:s potential, för att vi ska kunna förstå och därigenom acceptera sådana beslut? Detta eftersom ett autonomt, självlärande system som AI kan komma fram till slutsatser och fatta beslut av ett slag som vi människor inte ser logiken i.

¹¹ Varje år rankar *Global Innovation Index* innovationsprestanda för 130 ekonomier runt om i världen. The Global Innovation Index (GII) 2020 <https://www.globalinnovationindex.org/Home>

¹² Sveriges Radio. Europa halkar efter i AI-utveckling. (2019-11-20) <https://sverigesradio.se/artikel/7342619>

¹³ Artificiell intelligens i svenskt näringsliv och samhälle. Analys av utveckling och potential (2018-04-30) <https://www.vinnova.se/contentassets/ba8b7b5e21e7478ea193194bc94f397c/ai-analys-slutrapport-180430-final.pdf>

¹⁴ Lag om sekundär användning av personuppgifter inom social- och hälsovården (552/2019). <https://finlex.fi/sv/laki/alkup/2019/20190552>

¹⁵ Två exempel på grundläggande allomfattande etiksystem är: Pliktetik som baseras på förnuftet och Konsekvensetik som baseras på vilka konsekvenser ens handlande får.

- Behövs det flera olika strategier, normer och regelverk med tanke på att det finns ett flertal olika sub-discipliner såväl som tillämpningsområden av AI?
- Hur säkerställer vi etiska och kvalitativa AI-beslut utifrån kvaliteten på tillgänglig data för att undvika snedvridning (så kallad bias) och diskriminering? Dålig data riskerar att ge dåliga AI-resultat.
- Vem bär ansvaret när något går fel? Huvudmannen, tillsynsmyndigheten, professionen, tillverkaren, leverantören, programmeraren, användaren – eller någon annan?

I många sammanhang talas det om vad AI är och vilka möjligheter och fördelar som finns med att använda AI till exempel inom e-hälsa. Vanligtvis handlar det dock om pilotprojekt och väldigt lite är breddinfört, ofta med hänvisning till att det saknas kompetens, resurser och tid. Detta upplevs även inom hälso- och sjukvården enligt Socialstyrelsens rapport ”Digitala vårdtjänster och artificiell intelligens i hälso- och sjukvården”.¹⁶

Begreppet artificiell intelligens myntades redan på 1950-talet, inledningsvis med stor entusiasm. Därefter har både intresse och finansiering varierat genom decennierna. När denna rapport skrivs präglas AI-området av mycket uppmärksamhet, enorma investeringar och stora framtidsloften, inte minst inom hälsa och life science. Kanske är tiden nu mogen för att AI under det kommande decenniet på allvar ska slå igenom och bli en avgörande faktor för fortsatt välfärd, och för bästa möjliga e-hälsa?

5. AI kräver investeringar

Som nämnts tidigare så är AI förknippat med stora förväntningar om effektivisering, kvalitetshöjning och ekonomiska vinster, men också ökad livskvalitet för människor. Samtidigt finns frågetecken kring etik, juridik och inte minst kring ansvar om något går fel med AI.

USA och Kina gör stora satsningar inom AI och kan i många avseenden ses som världsledande. Under 2020 lanserar EU-kommissionen sina idéer om hur Europa kan komma ifatt och förbi Kina och USA¹⁷ när det gäller digitalisering, artificiell intelligens och en ny marknad för data. Samtidigt betonas individens rättigheter. EU har bland annat tagit fram tre dokument för att nå sina mål, *Kommissionens strategi för EU:s digitala framtid*¹⁸, *En inre EU-marknad för data genom frivilligt utbyte av data mellan företag och myndigheter och att skapa europeiska datacenter*¹⁹ och *Ett diskussionsunderlag om artificiell intelligens* (en så kallad vitbok).²⁰ Även Sverige som enskilt land ligger tämligen långt

¹⁶ Socialstyrelsen. Digitala vårdtjänster och artificiell intelligens i hälso- och sjukvården (okt 2019) <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2019-10-6431.pdf>

¹⁷ Fredrik Haglund. AI och data ska göra EU världsledande. Europaportalen. (2020-02-19) <https://www.europaportalen.se/2020/02/ai-och-data-ska-gora-eu-varldsledande>

¹⁸ Europeiska Kommissionen. Shaping Europe's digital future. (2020-02-19) https://ec.europa.eu/info/sites/info/files/communication-shaping-europes-digital-future-feb2020_en_3.pdf

¹⁹ Europeiska Kommissionen. A European strategy for data. (2020-02-19) https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

²⁰ European Commission. On Artificial Intelligence - A European approach to excellence and trust. (2020-02-19) https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

framme i teknisk utveckling, infrastruktur och digital mognad hos invånarna samtidigt som det finns risk för en långsammare takt när det gäller införande och tillämpning av exempelvis AI. Faktorer som påverkar takten är naturligtvis finansiering, men även synen på integritetsskydd och ägande eller rätten till användande av personliga data. Vissa länder har möjlighet att gå fortare fram än Sverige när det gäller att använda invånarnas vård- och hälsodata. Big Data – av god kvalitet – är det som får AI att fungera och ge optimalt resultat.

Det investeras stora summor i AI, främst inom näringslivet. Samtidigt är den globala AI-marknaden svår att uppskatta värdet på eftersom finansiella prognoser kan utgå från olika AI-definitioner. En av många analyser visar på en ökning från 28 miljarder USD 2019 till 40 miljarder 2020 (med årlig tillväxt på 43 procent). I samma prognos beräknas AI-marknaden nå ett värde av 100 miljarder USD år 2023.²¹

En annan prognos menar att globala investeringar i AI kommer fördubblas den närmaste 4-årsperioden, från 50 miljarder USD år 2020 till 110 miljarder USD år 2024.²² En tredje prognos anger att den globala AI-marknaden förväntas nå hela 733 miljarder USD år 2027, med en årlig tillväxt på 42 procent.²³

Gartner Inc. har redan för ett par år sedan uppskattat det globala marknadsvärdet av AI till 1,2 biljoner USD år 2018 (med en 70-procentig ökning från 2017) och med en prognos på 3,9 biljoner USD år 2022. Därtill förutspår en forskningschef på samma företag att AI under innevarande decennium kommer bidra med de mest disruptiva teknologierna genom utvecklingsprång inom djupinlärning och neurala nätverk, med ökad beräkningskraft och snabbhet i kombination med ökad volym och mångfald av data.²⁴

När det gäller enskilda, världsledande, företag så har Google enligt uppgift investerat 40 miljarder kronor i AI under åren 1998-2018.²⁵ På plats nummer två ligger Amazon som investerat närmare 10 miljarder kronor i AI under samma period. Andra bedömare menar att siffran kan vara betydligt högre per år och företag för de sex största AI-företagen.²⁶

I Sverige investerar Wallenbergfinansierade WASP²⁷ 5,5 miljarder kronor på AI fram till 2029 där man bland annat satsar på utbildning av 400 doktorander inom området.

²¹ Artificial Intelligence Global Market Report 2020-30: COVID-19 Growth and Change. Research and Markets. (2020-06-04) <https://www.globenewswire.com/news-release/2020/06/04/2043624/0/en/Global-Artificial-Intelligence-Market-Report-2020-to-2030-COVID-19-Growth-and-Change.html>

²² Worldwide Spending on Artificial Intelligence Is Expected to Double in Four Years, Reaching \$110 Billion in 2024, According to New IDC Spending Guide (2020-08-25) <https://www.idc.com/getdoc.jsp?containerId=prUS46794720>

²³ Artificial Intelligence Market Size Worth \$733.7 Billion By 2027. Grand View Research. (juli 2020) <https://www.grandviewresearch.com/press-release/global-artificial-intelligence-ai-market>

²⁴ Gartner Says Global Artificial Intelligence Business Value to Reach \$1.2 Trillion in 2018 (2018-04-25) <https://www.gartner.com/en/newsroom/press-releases/2018-04-25-gartner-says-global-artificial-intelligence-business-value-to-reach-1-point-2-trillion-in-2018>

²⁵ Ogelid. Framtidens Karriär. AI-utvecklingen domineras av it-jättarna. (2019-02-13) <https://framtidenskarriar.se/itdata/ai-utvecklingen-domineras-av-it-jattarna/>

²⁶ Göran Lindsjö. Ett bättre samhälle med hjälp av AI. <http://governo.se/vara-konsulter/goran-lindsjo>

²⁷ Wallenberg AI, Autonomous Systems and Software Program <https://wasp-sweden.org/>

Parallellt har man etablerat forskningsarenor där forskare och företag kan mötas och hitta samarbeten.²⁸

Svenska regeringen satsade totalt 40 miljoner kronor under 2018 och 2019 för att bidra till att möta arbetsmarknadens behov av omställning och fördjupade kunskaper inom AI.²⁹

I rapporten *Främja den offentliga förvaltningens förmåga att använda AI* uppskattas det ekonomiska värdet av ett fullständigt införande av AI (nuvarande AI teknik) i svensk offentlig förvaltning att uppgå till cirka 140 miljarder kronor årligen, motsvarande cirka 6 procent av dagens totala offentliga utgifter.

De investeringsnivåer och kostnadsuppskattningar som beskrivs ovan kan tolkas som att AI förknippas med en enorm förväntan och framtida påverkan på såväl samhället som organisationer och enskilda individer – och följaktligen att det finns en stor kommersiell potential.

Under det gångna året förnyades det svensk-tyska partnerskapet för innovation. Partnerskapet fokuserar på fyra olika områden varav e-hälsa (inklusive AI) är ett. Denna del av samarbetet koordineras av E-hälsomyndigheten i samarbete med berörda tyska myndigheter. En studie av Tysk-Svenska Handelskammaren³⁰ som genomförts bland medlemsföretag på båda sidor om Östersjön säger att 71 procent av de tyska företag som svarat på enkäten och 77 procent av de svenska planerar att införa eller använda mer AI fram till slutet av 2021. Allt fler företag av olika storlek ser en affärsmässig potential med AI. Dock skiljer sig de planerade tillämpningsområdena något beroende på land. Svenska företag vill satsa på AI inom försäljning och marknadsföring (25 procent), inom kundservice (22 procent) samt inom produktutveckling (17 procent). Tyska företag fokuserar på AI-tekniken inom produktion (24 procent), distribution (17 procent) samt inom försäljning och marknadsföring (16 procent).

Det har nämnts tidigare, men är värt att upprepa: Framgångsrik AI är beroende av stora mängder data, men det måste också vara ”rätt data” av god kvalitet – helst i strukturerad form. Vid sidan av ekonomiska investeringar så handlar det alltså till stor del även om investering i form av data. När det gäller utveckling och implementering av AI i internationell jämförelse så är det därför viktigt att titta på skillnader i förutsättningar, till exempel rörande synen på ägandet, eller rätten till användande, av personlig data i relation till skydd av den personliga integriteten. Vissa länder har möjlighet att gå fortare fram än Sverige när det gäller att använda till exempel invånarnas vårddata för att utveckla AI-applikationer (tidigare i rapporten beskrivs till exempel att Kina har få restriktioner på området). Skillnader av detta slag kan påverka investeringsviljan, men de kan också i förlängningen påverka tilliten till tekniken.

²⁸ Karin Lindström. IDG. Största satsningen på forskning någonsin – så ska Sverige bli bäst på AI (2020-09-29) <https://computersweden.idg.se/2.2683/1.740078/satsning-svensk-ai-wasp> .

²⁹ Regeringen (2019-05-09)

<https://www.regeringen.se/pressmeddelanden/2019/05/ytterligare-20-miljoner-till-vidareutbildning-inom-ai/>

³⁰ Sverige ingår sedan 2017 i ett tysk-svenskt innovationspartnerskap, bland annat med inriktning på AI.

Tillväxtverket menar att data är en strategisk resurs som krävs för kunskapsbyggande, innovation och verksamhetsutveckling i vår tid. Allt fler aktörer börjar förstå värdet av både sina egna och andras data, och inte minst av den information som finns men som ännu inte skapats i form av strukturerbara data. Tillgång till och systematisering av data behövs så att teknikens möjligheter för att skapa värden kan tas tillvara, exempelvis genom automatisering och AI. Detta förutsätter att data är relevanta, tillgängliga och håller hög kvalitet.³¹

När Tillväxtverket ställde frågan till drygt 30 svenska myndigheter om vad som hindrade dem från att använda AI så visade sig svaren stämma väl överens med resultatet från myndigheten DIGG:s djupintervjuer.³²

De största hindren var:

- Brist på kunskap inom förändringsledning
- Tillgång till ändamålsenlig data
- Författningar/regelverk och etik
- Tillgång till kompetens inom AI

6. Vad är AI?

Innan vi går in på de etiska och juridiska utmaningar som finns avseende AI, behöver vi först gå igenom vad AI är och visa på den bredd av teknologi som begreppet omfattar.

Gränsen för vad som är AI och vad som inte är AI har genom decennierna varit något flytande ända sedan begreppet myntades på 1950-talet. Numera kan man hävda att gränsen går vid regelbaserade system. Det innebär att när statiska regler för hantering av data är formulerade enbart av en programmerare så utgör den informationshanteringen inte AI i egentlig mening. Motsatsen, det vill säga ”äkta AI”, utgörs av autonoma och kontinuerligt adaptiva (lärande) system som utgår från lärande algoritmer för att komma fram till ett resultat eller beslut. Dock används begreppet AI ofta i bredare mening, vilket inkluderar, utöver autonomt lärande system, även renodlat regelbaserade system.

När det gäller oklara gränsdragningar så kan här också nämnas att en annan utläsning av förkortningen AI i vissa sammanhang kan vara *Augmented intelligence*, det vill säga förstärkt intelligens (ibland även *Cognitive augmentation*). Detta kan många gånger vara samma sak som artificiell intelligens, men då som tydliggörande att det rör sig om ett komplement till människans intelligens och kognitiva förmåga, inte en ersättning för densamma.

I denna rapport belyses AI utifrån ett fåtal sub-discipliner och specifika tillämpningar, främst inom hälso- och sjukvårdsområdet. Innan vi går in på dessa så kan först nämnas

³¹ Tillväxtverket. I takt med tiden. (2019-06-20) <https://tillvaxtverket.se/vara-tjanster/publikationer/publikationer-2019/2019-06-20-i-takt-med-tiden--en-omvarldsanalys.html>

³² Tillväxtverket. AI-nätverksmöte nr 5, Stockholm 2019-10-09. Beskrivet i: DIGG12019/01416/DF || I2019/01020/DF (delvis) Främja den offentliga förvaltningens förmåga att använda AI

begreppet **generell AI** som kan liknas vid en teknik- och databaserad motsvarighet till mänsklig intelligens, med förmåga att lära sig och så småningom excellera inom vitt skilda områden. Ytterligare ett begrepp i sammanhanget är den så kallad **singulariteten**, det vill säga den tidpunkt och det tillstånd då AI till fullo har uppnått samma nivå som – och är på väg att passera – den mänskliga intelligensen. Ett flertal globala auktoriteter inom it-baserad forskning och industri varnar för de risker som detta kan innebära för mänsklighetens existens. Dessa potentiella risker kopplade till generell AI ligger dock utanför omfånget för denna fokusrapport som är inriktad specifikt på AI inom e-hälsa.

6.1 AI – definitioner och begrepp

AI som vetenskap är ett avancerat teknologiskt och matematiskt område. Med tiden har delområden bildats i syfte att lösa specifika problem eller för utveckling och användning av vissa AI-baserade verktyg eller metoder.

Det kan upplevas som att det finns flera olika definitioner av AI, eftersom det är ett brett område som delvis är i kontinuerlig omvandling. AI bör knappast ses som ett buzzword idag eftersom begreppet myntades redan för 60–70 år sedan. AI-området omfattar dessutom tiotals olika subdiscipliner varför det sannolikt inte räcker med endast en strategi eller ett regelverk för området.³³

Exempel på subdiscipliner för AI är³⁴

- Machine Learning (Maskin-inlärning, se nedan)
- Neural Networks (artificiellt neuronät och Deep Learning (djupinlärning), se nedan)
- Evolutionary and genetic computing (evolutionära algoritmer)³⁵
- Vision recognition (synfältsigenkänning)
- Robotics (robotik, robotteknik)
- Expert systems (expertsystem)³⁶
- Speech processing (signalbehandling för tal)
- Natural language processing (bearbetning av naturligt språk)³⁷
- Data mining (Datautvinning)³⁸
- Logic reasoning and Probabilistic reasoning (probabilistiskt resonerande)
- Search (Sökningar)

³³ Christian Guttman, AI-forskare vid Karolinska Institutet och *Global Head of AI and Data Science* hos IT-företaget Tieto, Almedalsveckan 2019.

³⁴ Sweden AI Society SAIS <http://www.sais.se/blog/2013/04/renassans-for-artificiell-intelligens/>

³⁵ Område inom AI forskning. Inspiration från naturens evolutionära utvecklingsprocesser har överförts till automatisk algoritmgenerering där ett antal enkla beståndsdelar korsas och muteras tills man hittar ett tillräckligt bra beteende.

³⁶ Datorbaserat system som med hjälp av en kunskapsbas, rörande ett speciellt ämnesområde, och uppställda regler kan dra slutsatser liknande en mänsklig experts.

³⁷ Datorteknik för information som skrivs eller sägs på naturliga språk. I språkteknologi ingår sådant som maskinöversättning, taligenkänning, syntetiskt tal, talsyntes och utvinning av information ur stora textmassor. Även datorbaserade hjälpmedel för analys av talat och skrivet språk ingår.

³⁸ Analysteknik som går ut på att upptäcka viktiga och nya sammanhang, mönster och tendenser genom att analysera stora datamängder, via registrering av mönster samt statistiska och matematiska metoder i ett data warehouse.

En närmare titt på några av dessa subdiscipliner visar att Logic/Probabilistic reasoning i form av ”resonerande tillämpningar” omfattar skapandet av formalismer³⁹ som kan användas av ett system för att fatta beslut baserat på kunskap. De metoder som används bygger på hur människor resonerar, kommunicerar och löser problem.⁴⁰ Exempel inom detta område kan vara dataanalyser för prognostisering, för att förutsäga hur många varor av en viss typ som behöver köpas in till ett varuhus, eller prognostisering inom hälso- och sjukvården för att förutspå vårdbehov samt vilka patienter som har ökad risk att drabbas av en viss sjukdom. Resonerande tillämpningar av detta slag kan bidra till både kapacitetsoptimering samt uppföljning av kvalitet, men även till olika typer av sannolikhetsbaserade lösningar.

Inom Natural language processing finns exempel som chatbotar på kundtjänst sajter och vård sajter för att svara på frågor – muntligt och/eller skriftligt – och vid behov guida vidare i systemet. Ett aktuellt exempel är chatbot särskilt utformad för frågor kring covid19.⁴¹

Som nämnts tidigare kan AI ses som en förstärkning av den mänskliga intelligensen i exempelvis beslutsstöd. Det är i de sammanhangen man stöter på närbesläktade uttryck som *Augmented intelligence*, *Intelligence amplification* eller *Intelligence augmentation*. Det ligger en skillnad i om vi generellt menar att AI ska ersätta mänsklig arbetskraft eller om vi istället syftar till att använda maskininlärningstekniker för att stödja och hjälpa människor i deras arbetsuppgifter.

Ett annat, kanske närliggande begrepp är ”kognitiv it” som används av IBM i samband med Watson-datorn. Kognitiv it beskrivs som mer omfattande än den traditionella snäva definitionen av AI och omfattar tre huvudsakliga områden:

- Automatisering av upprepningsbara uppgifter, till exempel för robotar, med avsikt att öka effektivitet, kvalitet och noggrannhet.
- Upptäcka dolda mönster och relationer, med avsikt att identifiera potential för innovation.
- Skapa utökade lösningar för personalisering, vid kontakt med kunder (*hyperpersonalization*).

Exempel på kognitiv it inom hälso- och sjukvården kan vara ett sjukhus som tränar en maskin att analysera miljarder genprofiler för att kunna göra medicinsk prediktion (sannolik förutsägelse) för enskilda eller för grupper av individer. Ett annat exempel är en kognitiv assistent hos ett finansföretag som kan hantera tiotusentals samtidiga konversationer med kunder och besvara deras frågor på ett dussintal språk.⁴²

³⁹ Formalism kan ses som beskrivning av något i formella matematiska eller logiska termer.

⁴⁰ L. Danielsson, K. Lindström. IDG. Det här är AI och så funkar det. (2020-09-03)
<https://techworld.idg.se/2.2524/1.699032/ai-sa-funkar>

⁴¹ Asked and Answered: Building a Chatbot to Address Covid-19-Related Concerns (2020-10-28)
<https://catalyst.nejm.org/doi/full/10.1056/cat.20.0230>

⁴² Paul Roma, chefsanalytiker på konsultföretaget Deloitte Consulting. Citerad av L. Danielsson. IDG (2017-06-11) <https://techworld.idg.se/2.2524/1.684197/kognitiv-it>

6.1.1 Machine Learning

Machine Learning⁴³ (ML, maskininlärning) är ett forskningsområde och en populär teknik för att uppnå AI. Användningen av ML skiljer sig på många sätt från den explicita programmering av datorer vi är vana vid. Det innebär bland annat att systemen som bygger på ML är mer autonoma och adaptiva och inte enbart bygger på programmerade regelskrivningar. Ett annat sätt att beskriva skillnaden, till exempel för en applikation inom e-hälsa, är självlärande algoritmer kontra evidensbaserade, statistiska algoritmer. Med evidens avser vi här medicinsk evidens.

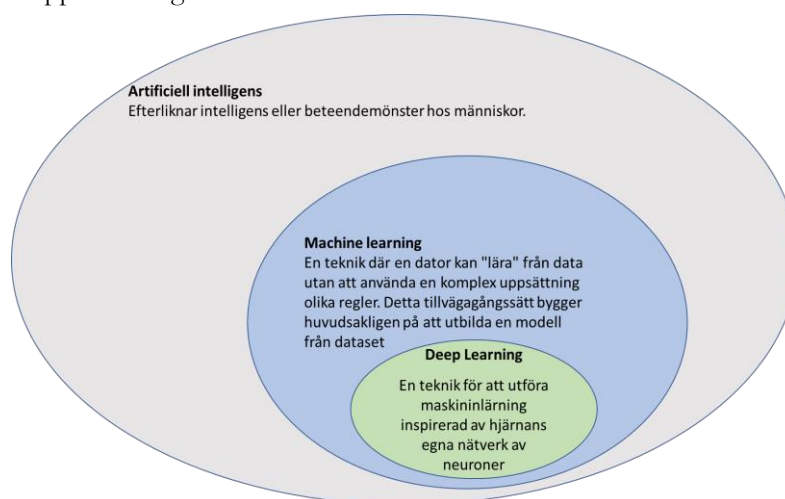
ML kan liknas vid att spara ner lärdomar till en modell som sedan kan beskrivas för en maskin. En kunskapsmodell av detta slag kan överföras från en maskin till en annan, så om en maskin tränats på att känna igen något kan det på mycket kort tid förklaras för en eller flera andra maskiner, vilket ger möjlighet till snabb spridning.

6.1.2 Deep Learning och neurala nätverk

Begreppet Deep Learning (DL, djupinlärning) kan också benämnas *Deep structured learning*, *Hierarchical learning*, eller ibland *Deep Machine Learning*. DL utgör en vidareutveckling av ML och kan ses som ett artificiellt neuralt nätverk (ANN) bestående av ett *input layer*, ett *output layer* och däremellan ett flertal *deep layers*, det vill säga lager för indata respektive utdata, samt flera djuplager för behandling av denna data, där algoritmerna fungerar på ett sätt som är tänkt att efterlikna hur den mänskliga hjärnans neuroner kommunicerar med varandra.

Förenklat kan ANN beskrivas som ett nätverk som lär sig genom att matas med mätdata och kända svar. Under träning av nätverket uppkommer svar som skiljer sig från facit, och nätverket justerar sig allteftersom. Proceduren upprepas om och om igen tills rätt svar kommer upp tillräckligt ofta.⁴⁴

Figuren visar hur AI, ML och DL relaterar till varandra.



⁴³ Delvis hämtat från AI och machine learning för beslutstöd inom hälso- och sjukvård. Marcus Österberg och Lars Lindsköld. https://vgrblogg.se/utveckling/2018/06/06/ai-och-machine-learning-inom-halso-och-sjukvard/#id_Toc515833568

⁴⁴ IDG Network Expert AI Techworld. Så tränas en AI (2017-11-13) <https://techworld.idg.se/2.2524/1.692189/maskininlarning-dummies>

7. Utmaningar och risker med AI

I detta avsnitt utvecklas resonemangen kring etiska aspekter och förtroendet för AI, för att längre fram komma in på juridik kopplat till AI, främst när denna typ av system är autonoma och adaptiva.⁴⁵

Självkörande bilar är ett exempel där AI orsakat dödsfall och komplicerade frågor har väckts om vem som bär ansvaret. Detta ger en föraning av komplexiteten i ansvarsfrågor om eller snarare när något går fel med AI och autonoma system, till exempel inom sjukvården eller socialtjänsten.

När det gäller människors tillit till AI, dvs. tydliga förklaringar hur resultat och beslut tas fram och vad dessa bygger på, så krävs arbete både med transparens och förklarbarhet, men också tolerans och opartiskhet. Transparens och förklarbarhet handlar om hur väl vi måste förstå AI för att acceptera dess beslut. Tolerans handlar om hur fördelarna med AI måste överväga eventuella nackdelar. Opertiskhet handlar om att vi måste uppleva att AI är rättvist och behandlar alla människor lika, oavsett kön, socioekonomisk tillhörighet, etnicitet, sexuell läggning, eventuell funktionsnedsättning etcetera. Dessa aspekter – transparens, förklarbarhet, tolerans och opartiskhet – beskrivs utförligare nedan.

En annan aspekt rör informationssäkerhet. Eftersom AI är beroende av mycket stora mängder data, som inom e-hälsoområdet kan utgöras av personuppgifter, krävs det att denna data hanteras på ett korrekt och säkert sätt så att individers integritet inte kränks eller äventyras. Annars kan tilliten till AI skadas.

Utifrån ett bredare samhällsperspektiv ställer AI därmed oss inför ett antal utmaningar. Detta gäller både inom hälso- och sjukvård och socialtjänst.

7.1 Exempel när det kan gå fel

AI-system kan uppvisa ett antal brister, huvudsakligen kopplade till dataunderlaget som används (fel eller ostrukturerade data), men de kan även vara i form av rena programmeringsfel.

Dessa brister skiljer sig från traditionella sårbarheter i programvara och visar att även om AI kan överstiga mänsklig intelligens på många sätt, så kan den också misslyckas på sätt som en människa knappast skulle göra. Med andra ord så kan bristfällig AI visa sig vara påfallande ointelligent.

Det finns även risk att AI-system blir hackade i likhet med vad som kan drabba andra databaserade system. Genom obehörig åtkomst kan en främmande aktör ta kontroll över AI-systemet och påverka det i innehåll och funktion. Detta kan innebära risk för skada på såväl människa som på egendom.

⁴⁵ Autonomi = Förmåga att uträtta uppgifter i komplexa miljöer utan ständig styrning av användaren. Adaptivitet = Kapacitet att förbättra sin prestationsförmåga genom att lära sig av erfarenheter.

Det finns också något som benämns som ”Dual use” det vill säga att en resurs kan användas i civilt, men också i militärt eller kriminellt syfte. Vid utveckling av AI bör man vara observant på risk för detta, det vill säga att AI som är utvecklat för ett positivt och gott syfte kan komma att användas i negativt syfte på ett skadligt sätt.⁴⁶

En typ av sårbarhet utgörs av så kallade motstridiga exempel (*adversarial examples*) som leder till att AI felklassificerar saker. Denna typ av motstridiga exempel kan vara bilder med avsiktligt eller oavsiktligt infogade störningar (ändrade pixlar) osynliga för blotta ögat – men som kan få ett AI-system att plötsligt tolka bilden av en bil som en struts.⁴⁷ Fenomenet med motstridiga exempel skulle kunna missbrukas, till exempel i syfte att lura vårdssystemet på ekonomisk ersättning för falska diagnoser. Detta har visats och diskuterats i en studie där forskarna testade metoden för att lura systemet inom tre olika typer av medicinsk bilddiagnostik: dermatologi (huddiagnostik, exempelvis melanom), radiologi (röntgen, exempelvis lungsjukdom), samt oftalmologi för diagnostik av ögonsjukdom vid diabetes eller andra tillstånd.⁴⁸ Forskarna understryker vikten av att utveckla robusta system med inbyggda försvarsfunktioner mot denna typ av sabotage, bland annat i form av algoritmiskt försvar respektive infrastrukturellt försvar. De konstaterar också att behovet av försvarsåtgärder kan innebära en nödvändig avvägning med ett eventuellt avkall på AI-systemets prestanda.

Det finns även andra sätt att manipulera AI. Forskning har visat att vissa maskininlärningsalgoritmer kan avsiktligt korrumpas genom att komplettera deras träningsdata-set med skadlig data. Sådana attacker kallas förgiftningsattacker. Ett hypotetiskt exempel skulle kunna vara ett AI-system för en specifik diagnos, där systemet är ”förgiftat” med skadlig träningsdata så att det ligger för högt eller för lågt när det gäller gränsdragning för att ställa diagnosen ifråga. Konsekvensen blir att antalet falskt positiva eller falskt negativa resultat ökar.⁴⁹

En annan typ av förgiftningsattack kallas *backdoors in neural networks*, förkortat *BadNet*. I detta fallet fungerar AI-systemet korrekt fram till dess att det utsätts för en specifik trigger som påverkar genom en dolt förprogrammerad ”bakdörr”. I en experimentell studie har till exempel visats hur ett AI-system för självkörande fordon genom en sådan *BadNet* i särskilda sammanhang kan tolka en stopp-skylt som en skylt för hastighetsbegränsning, vilket naturligtvis skulle kunna skapa direkt livsfarliga situationer.⁵⁰

Nedan följer exempel på AI-applikationer som avsiktligt eller av misstag har gett upphov till mindre lyckat resultat, eller som lett till allvarliga tillbud.

⁴⁶ The Malicious Use of Artificial Intelligence Forecasting, Prevention, and Mitigation. M. Brundage et al. (februari 2018) <https://maliciousaireport.godaddysites.com/>

⁴⁷ Intriguing properties of neural networks. C. Szegedy et al. (2014-02-19) <https://arxiv.org/pdf/1312.6199.pdf>

⁴⁸ Adversarial Attacks Against Medical Deep Learning Systems. S.G. Finlayson et al. (2019-02-04) <https://arxiv.org/pdf/1804.05296.pdf>

⁴⁹ Systematic Poisoning Attacks on and Defenses for Machine Learning in Healthcare. M. Mozaffari-Kermani et al. (2014-07-30) <https://ieeexplore.ieee.org/document/6868201>

⁵⁰ BadNets: Evaluating Backdooring Attacks on Deep Neural Networks. T. Gu et al. (2019-04-18) <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8685687>

7.1.1 Preventivapp och oönskade graviditeter

Sedan några år tillbaka finns det en svenskutvecklad preventivapp tillgänglig på marknaden, som ska fungera som en metod för att undvika graviditeter. Denna app grundar sig på en unikt framtagen algoritm och kan ses som en AI-liknande lösning. Metoden går ut på att kvinnan dagligen mäter sin basaltemperatur. Uppmätta värden läggs sedan in i appen som utifrån denna data bedömer om kvinnan är fertil den dagen.

Appen blev omskriven 2018 i samband med en anmälan om att oönskade graviditeter uppstått vid användning av appen. Läkemedelsverket avskrev dock ärendet efter att tillverkaren hade tydliggjort användarens ansvar i och med att lösningen inte är 100 procent säker.⁵¹

7.1.2 Diskriminerande chatbot

Våren 2016 lanserade Microsoft en AI-baserad så kallad chatbot vid namnet *Tay* (en förkortning av *Thinking about you*). *Tay* var designad att efterlikna språkmönstret hos en 19-årig amerikansk flicka och var tänkt att lära sig genom att interagera på Twitter. Redan efter en dag på Twitter stängdes *Tay* ner eftersom den utifrån processad information på det fria internet hade formats till ett ”rasistiskt nättroll”.⁵² En lärdom av detta är att ett AI-system inte blir bättre än den mjukvara och information som den är grundad på.

7.1.3 Problematisk ansiktsanalys

AI-system för ansiktsanalys och ansiktsigenkänning utvecklas och används inom flera olika tillämpningsområden, till exempel som säkerhetslösning för att låsa upp mobiltelefoner eller inom handel där en individ kan identifieras med ansiktsigenkänning som grund för betalning.

Ansiktsanalys används också alltmer som en del av övervakningssystem och scanning till exempel på internationella flygplatser eller andra högsäkerhetsområden. Vid tillämpning av detta slag kan gränsen för personlig integritet tangeras – eller rentav överskridas?⁵³ Även i Sverige har gränserna redan testats för vad som är tillåtet i detta avseende, genom ett projekt för att registrera skolnärvaro.⁵⁴

Inom e-hälsa kan ansiktsanalys används som en del i en medicinsk bedömning/anamnes vid triagering⁵⁵ eller dylikt, för bestämning av patientens kön, ålder och hälsotillstånd. Flera studier som utvärderat AI-system för ansiktsanalys visar dock på brister av olika

⁵¹ Dagens Medicin. Läkemedelsverket ger grönt ljus till preventiv-app (2018-09-13)

<https://www.dagensmedicin.se/alla-nyheter/omvardnad/lakemedelsverket-ger-gront-ljus-till-preventiv-app/>

⁵² Paul Mason. The Guardian. The racist hijacking of Microsoft's chatbot shows how the internet teems with hate. (2016-03-29) <https://www.theguardian.com/world/2016/mar/29/microsoft-tay-tweets-antisemitic-racism>

⁵³ Ian Sample. The Guardian. What is facial recognition - and how sinister is it? (2019-07-29)

<https://www.theguardian.com/technology/2019/jul/29/what-is-facial-recognition-and-how-sinister-is-it>

⁵⁴ Datainspektionen. Sanktionsavgift för ansiktsigenkänning i skola. (2019-08-21)

<https://www.datainspektionen.se/nyheter/2019/sanktionsavgift-for-ansiktsigenkanning-i-skola/>

⁵⁵ Triagering är en bedömning av patientens medicinska allvarlighetsgrad utifrån anamnes, symtom och ibland även vitalparametrar.

slag, t ex vid könsbestämning vilket kan vara problematiskt bland annat utifrån ett medicinskt perspektiv.^{56,57}

7.1.4 Självkörande bilar

I mars 2018 avled en 49-årig kvinna i Arizona, USA efter att ha blivit påkörd av en självkörande taxi.⁵⁸ Detta är ett exempel på ansvarsproblematiken när ett autonomt tekniskt system orsakar allvarlig skada eller som här, ett dödsfall.

Olyckan väcker därmed frågor kring vem som bär ansvaret för olyckan? Är det:

- Offret själv?
- Bilens förare, som borde ha reagerat och trampat på bromsen?
- Taxiföretaget, som släppte ut självkörande testfordon i trafiken?
- Teknikerna och programmerarna, som konstruerade bilens sensorer och systemet som skötte körningen? Eller deras chefer eller projektledare?
- Staden/samhället, för dålig trafikplanering?
- Lagstiftare i Arizona som tillåtit tester med självkörande bilar?
- Bilen/tekniken själv som kan beläggas med ansvar?
- Ett delat ansvar eller ett ansvar som olika aktörer har i olika led?

7.1.5 Socioekonomisk och etnisk diskriminering

Studier har visat att vid AI-baserade reklamutskick i USA har socioekonomiskt svaga områden i ett samhälle aktivt valts bort. Det var ett AI-beslut grundat på renodlat kommersiell bedömning. Ett annat exempel från USA visar hur algoritmbaserade utskick för att informera om hälsovård utesluter andra delar av befolkningen än den vita.⁵⁹ Detta visar att AI utifrån ett bristfälligt formulerat grunduppdrag och/eller bristfällig data kan diskriminera vissa grupper.

7.1.6 Sakernas internet (Internet of Things)

AI är beroende av stora mängder data. Internet of Things (IoT) kallas det kraftigt växande nätverket av uppkopplade föremål av olika slag. Det omfattar vardagsföremål som hushållsapparater, kläder och accessoarer, men även maskiner, fordon och byggnader, med inbyggd elektronik och internetuppkoppling, vilket gör att de kan styras

⁵⁶ Proc. ACM Hum.-Comput. Interact. MK Scheuerman et al. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis and Image Labeling Services. (november 2019) https://docs.wixstatic.com/ugd/eb2cd9_963fbde2284f4a72b333ea2ad295fa6d3.pdf

⁵⁷ J Boulamwini et al. Gender Shades. How well do IBM, Microsoft, and Face++ AI services guess the gender of a face? (2018) <http://gendershades.org/index.html>

⁵⁸ The Guardian. Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian (2018-03-19)

<https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe>

⁵⁹ Obermeyer et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science. (2019-10-25) <https://science.sciencemag.org/content/366/6464/447>

eller utbyta data över nätet. IoT-enheter har potential att generera en enorm mängd data som i sin tur kan användas av AI.⁶⁰ En risk med detta är att IoT inte är immunt mot hackning. Hackare kan starta DDoS-attacker⁶¹ genom att infiltrera och utnyttja tusentals eller miljoner oskyddade enheter.⁶² Experter förutspår att det år 2025 kommer finnas så många som 75 miljarder anslutna IoT-enheter.⁶³ Enligt dataföretaget Intel växer IoT i en hisnande takt från två miljarder objekt 2006 till beräknade 200 miljarder redan år 2020. Det innebär i så fall att vi redan idag kan ha närmare 30 uppkopplade och data-genererande objekt för varje människa på jorden.⁶⁴

Mycket av den inbäddade mjukvaran (firmware) som körs på dessa enheter är oskyddad och sårbar, vilket i sin tur lämnar ett obestämt antal kritiska system runt om i världen i riskzonen. Ett exempel på denna typ av sårbarhet är en databas på ett casino som blev hackad via en uppkopplad termometer i ett akvarium.⁶⁵

Uppkopplade prylar har blivit ett allt större cyberhot och många hackare har börjat utnyttja oskyddade IoT-objekt för sina angrepp.⁶⁶ De flesta smarta enheter tillhörande IoT finns emellertid inte i hemmen eller i våra fickor utan förekommer i fabriker, företag och i sjukvården. Smarta objekt ger dessa stora verksamhetsområden den vitala informationen de behöver för att inventera lagerhållning, hantera maskiner, öka effektiviteten och därmed spara kostnader, och till och med rädda liv. År 2025 kan det totala globala värdet av IoT vara drygt sex miljarder dollar (USD). Det mesta av detta är enheter inom hälso- och sjukvård (2,5 miljarder USD) samt inom tillverkningsindustri (2,3 miljarder USD).⁶⁷ Inom vården kan IoT användas för bland annat bärbar monitorering som kontinuerlig glukosmätning vid diabetes, transportövervakning av läkemedel eller prover, samt mobil journalföring.

Sammantaget pekar ovanstående tydligt på behovet av datasäkerhet och riskhantering angående det indirekta hotet mot AI från IoT, som med all sin data sannolikt blir allt viktigare för framtidens AI.

⁶⁰ Kacper Kubara. Artificial Intelligence meets the Internet of Things (2019-07-10)

<https://towardsdatascience.com/artificial-intelligence-meets-the-internet-of-things-a38a46210860>

⁶¹ En attack mot ett datasystem i syfte att hindra normal användning av systemet. Den vanligaste angreppstypen är överbelastningsattack, där systemet kommer att använda någon knapp resurs (nästan enbart till att hantera data genererat genom attacken. (Wikipedia)

⁶² Terry Dunlap. The 5 Worst Examples of IoT Hacking and Vulnerabilities in Recorded History (2020-06-20)

<https://www.iotforall.com/5-worst-iot-hacking-vulnerabilities/>

⁶³ Anurag. 13 IoT Statistics Defining the Future of Internet of Things. (2018-01-08)

<https://www.newgenapps.com/blog/iot-statistics-internet-of-things-future-research-data>

⁶⁴ Intel. A Guide to the Internet of Things. <https://www.intel.com/content/www/us/en/internet-of-things/infographics/guide-to-iot.html>

⁶⁵ Wang Wei. The hacker news. Casino Gets Hacked Through Its Internet-Connected Fish Tank Thermometer. (2018-04-16) <https://thehackernews.com/2018/04/iot-hacking-thermometer.html>

⁶⁶ Nicole Eagan, vd för säkerhetsföretaget Darktrace, på konferensen WSJ CEO Council (2018-04-15) <https://www.businessinsider.de/hackers-stole-a-casinos-database-through-a-thermometer-in-the-lobby-fish-tank-2018-4?r=UK&IR=T>

⁶⁷ Intel. A Guide to the Internet of Things (2019-09-10) <https://www.intel.com/content/www/us/en/internet-of-things/infographics/guide-to-iot.html>

7.2 AI inom e-hälsa – risker och etiska dilemman

Utvecklingen av AI-system inom e-hälsa är ännu i sin linda, och såvitt rapportförfattarna känner till har hittills inga allvarigare fel eller tillbud med patienter rapporterats inom detta område. Det finns emellertid exempel där AI i prototyp-version har gett olämpliga rekommendationer till exempel för behandling av cancer, enligt uppgift på grund av att systemet matats med hypotetiska scenarier snarare än riktiga patientdata.⁶⁸ En sammanställning för IBM Watson visar på olika möjliga AI-tillämpningar inom medicin, samt ett flertal påbörjade projekt åren 2011-2017, men med endast ett fåtal av dessa som kommit till klinisk användning.⁶⁹

De databaserade lösningar som idag finns i klinisk användning i vården är i stort sett regelbaserade lösningar (fasta algoritmer) som inte riktigt är att se som äkta AI, det vill säga autonoma och kontinuerligt lärande system. Eftersom autonomt beslutsfattande AI-lösningar i dagsläget är införda i mycket begränsad omfattning i vård och omsorg så medför detta att analys av risker och tänkbara dilemman måste göras utifrån ett mer hypotetiskt resonemang.

AI bedöms kunna få en omvälvande effekt på hela världens uppbyggnad, bland annat genom att:

- Genomik⁷⁰ och AI i ökad utsträckning möjliggör preventiv vård istället för som nu reaktiv vård.
- Genomik och AI kommer bidra till en mer individanpassad vård ("personalized medicine", precision medicin).
- Interaktioner kommer att öka mellan patienter och AI-system liksom mellan vårdpersonal och AI.

I en nära framtid kan emellertid det preventiva sättet att arbeta komma att bli utmanande. Med möjligheten att tidigt inhämta och tolka stora mängder genetisk data så kommer i princip alla individer visa sig ha förhöjd risk att få någon sjukdom. Ska då alla behandlas preventivt? Har vi råd med detta? Och vilken psykisk påverkan kommer det ha på oss, att veta att vi är genetiskt predisponerade för en eller kanske flera allvarliga sjukdomar?

Grundförutsättningarna för AI är att data finns i stora mängder samt att den är tillgänglig att använda. Detta kan påverka hur vi ser på integritetsaspekter. Integritet när det gäller skydd av persondata är inte absolut utan snarare en avvägning mellan för- och nackdelar, för individen men också för samhället i stort. Redan här finns ett dilemma i form av att man å ena sidan vill ha alltmer data för att träna AI-systemen, och därefter ha maximal nytta av dem. Å andra sidan kan det vara svårt att inhämta tillstånd (samtycke) från alla personer och patientgrupper vars data man vill komma åt.

⁶⁸ Angela Chen (2018-07 26) IBM's Watson gave unsafe recommendations for treating cancer
<https://www.theverge.com/2018/7/26/17619382/ibms-watson-cancer-ai-healthcare-science>

⁶⁹ IEEE Spectrum. How IBM Watson Overpromised and Underdelivered on AI Health Care (2019-04-02)
<https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>

⁷⁰ Ett fält inom molekylärbiologin där arvsmassan analyseras hos en organism -
<https://genteknik.nu/genomik/>

Det kan dessutom vara svårt att veta vad som gäller för samtycke kopplat till hälso- och sjukvård samt till forskning inom medicinområdet. Krav på samtycke finns i flera lagar och det kan vara svårt att få en överblick. Komet (Kommittén för teknologisk innovation och etik) har därför sammanställt en rapport som, utifrån frågeställningar som rör samtycke med koppling till hälsodata, beskriver hur dagens svenska regelverk fungerar. I rapporten definieras hälsodata som ”en dokumenterad uppgift som rör en persons fysiska eller psykiska hälsotillstånd”.⁷¹ En annan etisk frågeställning kan uppstå när AI ersätter vårdpersonal. Att använda en AI-baserad chatbot för akutmottagningens triagering kommer troligen i längden att vara betydligt billigare än att ha sjuksköterskor för denna uppgift. Även läkarens beslut om diagnos och behandling kan komma att grundas helt eller delvis på AI. Innebär detta att delar av vården kommer ”avhumaniseras”, dvs att patienten alltmer sällan kommer i kontakt med mänsklig vårdpersonal?

Babylon Health i Storbritannien är en vårdtjänstleverantör som tillhandahåller fjärrkonsultationer med läkare och vårdpersonal via text- och videomeddelanden genom en mobilapplikation. Det har uppstått viss oro över den teknik som används av Babylon Health, särskilt angående säkerheten och kvaliteten på de råd som erbjuds av chatbot triage-verktyget. Babylon-appen innehåller en chatbot med *Symptom Checker* som syftar till att ge patienter relevant hälso- och triageinformation baserad på deras symtom. Men detta systems pålitlighet har blivit ifrågasatt vilket beskrivs i en artikel i den vetenskapliga tidskriften *The Lancet*.⁷² I artikeln lyfts fram att de uppföljningsstudier som gjorts för att bedöma verktygets kvalitet har varit för begränsade i sin omfattning. Även verktygets användbarhet ifrågasätts. Eftersom chatbot triage-verktyg bedöms ha stor potential efterlyses i artikeln öppna testprotokoll och transparens för att fortsatt kunna utveckla denna typ av verktyg.

Ovanstående leder till en mer övergripande etisk aspekt, nämligen huruvida det kan bli en klassfråga vem som får ta del av den bästa vården, som kan vara baserad antingen på mänsklig vårdkontakt eller automatiserad i en AI-lösning. Hälso- och sjukvårdslagen anger i sin portalparagraf att vård ska ges på lika villkor för alla. Detta är en utmaning redan idag och framöver kan det vara fortsatt svårt att säkerställa en likvärdig, eventuellt AI-optimerad, vård och omsorg för alla medborgare. Denna fråga kan också kopplas till digitalt utanförskap, som Internetstiftelsen bedömer omfatta cirka en miljon människor i Sverige vilka sällan eller aldrig använder internet. Dessa människor riskerar att missa tillgänglighet och mervärde av en vård som sannolikt utvecklas alltmer i digifysisk riktning.

Andra risk-scenarion med AI inom e-hälsa är t.ex:

- **Bias vid akutmårdstriagering eller annan vårdprioritering**

Tidigare erfarenheter ger fog för en oro att AI-system som har till uppgift att göra en primär medicinsk bedömning av människors vårdbehov (till exempel på

⁷¹ Komet beskriver Samtycke inom vård och inom medicinsk forskning (2020-02-03) https://www.kometinfo.se/wp-content/uploads/2020/02/Samtycke_Komet-beskriver_2019_07.pdf

⁷² H. Fraser, E. Coiera, D. Wong. *The Lancet*. Safety of patient-facing digital symptom checkers. (2018-11-06) [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(18\)32819-8/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)32819-8/fulltext)

en akutmottagning, eller via internet), skulle kunna begå misstag i form av bias (snedvridning/diskriminering) om patienten tillhör en ovanlig kategori till exempel utifrån etnisk tillhörighet (se 7.1.3 ovan).

- **Genusbias vid beslutsstöd**

Sedan länge vet vi att en betydande del av dagens medicinska kunskap är grundad på statistik och data från manliga individer.⁷³ Därigenom finns en risk att ett AI-system misslyckas att i tillräcklig grad ta hänsyn till patientens kön. Kända exempel är olika symtombild vid hjärtinfarkt hos kvinnor respektive män⁷⁴, samt könsskillnader i normalt referensintervall vid laboratoriemedicinska analyser.

- **Oklar grund för medicinska beslut (black box-fenomenet)**

En lika grundläggande som svår frågeställning när det gäller AI är hur vi hanterar risken för att AI-systemets beslut ter sig obegripligt för oss människor. Det är inte otänkbart att vi kan hamna i situationer där AI beslutar (eller ger beslutsstöd) som är av stor betydelse för individen såväl som för samhället. Kan vi då acceptera beslutet trots att vi inte – inte ens professionen – förstår hur beslutet fattats? Detta berörs i avsnitt om transparens och förklarbarhet nedan.

Risken med partiskhet och skevhet i data (bias) måste hanteras på ett sådant sätt att algoritmerna inte bidrar till felaktigheter och feltolkningar. Annars kan resultatet bli under- eller överdiagnostisering, och därigenom onödiga eller fördröjda vårdbesök, som i sin tur medför både ökad oro och lidande, samt ökade sjukvårdskostnader.

De olika typer av utmaningar som beskrivs ovan innebär att angreppssätten behöver vara multidisciplinära och tvärprofessionella med kompetenser från exempelvis medicin, psykologi, juridik, filosofi, datavetenskap och ekonomi.

7.3 Tillit och förtroende

När utvecklingen av det digitaliserade samhället går allt snabbare så kommer frågor upp som: Kan vi vara säkra på att vår persondata hanteras på ett korrekt och säkert sätt? Riskerar den enskildes integritet att kränkas – eller är nuvarande integritetsskydd snarare för snävt? Kan farhågorna bli sanna – att vi skapar system utan kontroll och att algoritmerna tar över?

AI inom hälsa och sjukvård bygger på stora mängder känsliga data. Tillit till att dessa data hanteras på ett säkert sätt är en viktig förutsättning för AI:s utveckling.

Forskare på AI Sustainability Center⁷⁵ beskriver den snabba utvecklingen inom AI, och de stora utmaningar som riskerar att skada tilliten. De fyra problemområden som de

⁷³ The Guardian. The female problem: how male bias in medical trials ruined women's health (2019-11-13)

<https://www.theguardian.com/lifeandstyle/2019/nov/13/the-female-problem-male-bias-in-medical-trials>

⁷⁴ K. Schenck Gustafsson et al. Ge kvinnors hjärtan mer uppmärksamhet! Dagens Samhälle (2018-03-10) <https://www.dagensamhalle.se/debatt/ge-kvinnors-hjartan-mer-uppmarksamhet-21122>

⁷⁵ AI Sustainability Center <http://www.aisustainability.org/publications/>

identifierat som nyckelområden att hantera för AI-teknologins fortsatta framgångsrika utveckling är:

1. partiskhet, även kallat ”bias”,
2. ansvarsfrågor,
3. missbruk och illvillig användning,
4. transparens och förklarbarhet.

Samma forskare föreslår tre åtgärder för att kunna möta problemområdena:

- förtydliga regelverken,
- satsa på flervetenskapliga perspektiv kring tillämpad AI för att nå mer kunskap om utmaningarna,
- stärka tilliten i samhället för AI.

Faktorer som bidrar till förtroende är transparens och förklarbarhet vilket diskuteras i nästa avsnitt.

7.4 Transparens och förklarbarhet

AI kan i mer avancerade former fungera på ett sätt som är svårt eller kanske till och med omöjligt för oss människor att förstå. Med andra ord utgör AI i det fallet en ”black box”, en svart låda där vi inte har en aning om hur ett beslut tagits fram.

Förklarbarheten är låg eller rent av obefintlig, kanske även för de medicinska experterna. Är det tänkbart att vi ändå accepterar AI-beslutet, till exempel rörande en sjukdomsdiagnos eller behandling, eftersom det medicinska utfallet är viktigare än att vi måste förstå logiken bakom?

Eller är transparens och förklarbarhet så betydelsefullt för att AI ska accepteras att vi tvingas ”trimma ner” AI-prestandan – nöjer vi oss med sämre AI bara för att vi ska förstå hur det fungerar?

Å andra sidan kan en jämförelse göras med hur en erfaren och skicklig läkare ställer en diagnos för en patient utifrån informationsinhämtning och olika typer av sinnesintryck (syn, hörsel, känsel, lukt), på ett sätt som kan vara svårt att förklara, för patienten i fråga, för en myndighet, kanske även för en kollega. Detta kan ses som exempel på det vi brukar kalla tyst kunskap. Även i detta fallet kan förklarbarhet vara en utmaning, men kanske har vi overseende på grund av tillit till den medicinska professionen?

Hursomhelst, så har kanske förklaringsmodeller saknats redan tidigare.

Black box-frågan är kopplad till AI i mer avancerad form såsom deep neural networks. Följaktligen är det komplicerat att skapa förklaringsmekanismer, och ”förklarbar AI” (*eXplainable AI*, XAI) kan numera ses som ett eget vetenskapligt fält inom AI.⁷⁶

⁷⁶ Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Barredo Arrieta et al. *Information Fusion* (juni 2020)
<https://www.sciencedirect.com/science/article/pii/S1566253519308103>

Ett alternativ kan vara att använda fasta regelbaserade system där det bör vara betydligt enklare att implementera förklaringar. En sådan strategi innebär dock sannolikt betydande avkall på prestandan, eftersom AI-systemets lärandefunktion väljs bort.

Denna frågeställning har beröringspunkter med de juridiska aspekter som avhandlas i avsnitt 9.2 nedan.

7.5 Tolerans och rättvisa (fairness)

Ovan har vi berört för AI-området viktiga aspekter som tillit och förtroende, samt transparens och förklarbarhet. Ytterligare aspekter, vilka delvis överlappar eller är kopplade till dessa, är tolerans och det som i AI-sammanhang brukar kallas fairness.

AI-utvecklingen kan antas vara ett så pass stort utvecklingssteg för mänskligheten att det kan jämföras med tekniskiftet när vi gick över från häst till bil. I bilens barndom utgick man från att automobilen skulle följa samma regler som gällde för häst. Men eftersom tekniken var överlägsen hästen visade det sig att detta var omöjligt, och att nya regler/lagar fick införas. Behöver vi på liknande sätt helt nya regelverk för att kunna utnyttja AI på bästa sätt?

En närliggande fråga som också ställdes på sin spets under bilismens utveckling är huruvida det är acceptabelt att trafiken skadar och dödar. Vi har ännu inte uppnått nollvisionen, det vill säga noll dödsfall i trafiken. Fördelarna med bilar är så stora att vi tillåter oss att fortsätta köra trots att trafiken dödar. Under de senaste åren har drygt 200–300 människor omkommit årligen i Sverige. Globalt skördar trafiken mer än en miljon dödsoffer varje år, vilket är betydligt fler än antalet människor som dödas av krig och terror.⁷⁷

Om vi övergår från trafiksäkerhet till patientsäkerhet så visar Socialstyrelsens statistik⁷⁸ att i Sverige drabbas varje år cirka 100 000 människor av vårdskador och av dessa dör 1 400. Minst en tredjedel av dessa dödsfall anses bero på undvikbara vårdskador. En förhoppning är att AI ska kunna bidra till att minska undvikbara vårdskador och dödsfall i vården. Men även med AI kommer dödsfall inträffa, och kanske på nya grunder.

Sammantaget rör detta en fråga om tolerans – är det tänkbart att vi accepterar ett visst mått av skador och dödsfall även orsakade av AI i vården, om fördelarna (nyttan) trots allt överväger?

En annan, åtminstone teoretisk aspekt av tolerans kan vara ett visst mått av overseende med bias, det vill säga ett AI-system delvis präglad av partiskhet eller diskriminering, återigen eftersom nyttan överväger. Å andra sidan så kommer vi här in på begreppet fairness, eller rättvisa. En sannolik grundförutsättning för att vi ska uppvisa acceptans/tolerans även för negativa konsekvenser av AI inom e-hälsa såsom vårdskador och dödsfall är sannolikt att vi uppfattar systemet som just rättvist och

⁷⁷ H. Ritchie, M. Roser. Causes of Death (december 2019) <https://ourworldindata.org/causes-of-death>

⁷⁸ Socialstyrelsen. Socialstyrelsens arbete med Patientsäkerhet. (2019-10-16) <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2019-4-3.pdf>

präglat av likabehandling. Därmed kan toleransen rörande eventuellt negativa konsekvenser av AI ses som direkt kopplad till nyttan, i kombination med en tillräcklig nivå av rättvisa.

7.6 Intelligens, etik och moral

Människors beslut om vad som är rätt och fel vilar på normer och värderingar i det samhälle de tillhör. Ibland uttrycks det som vår etiska kompass. Dessa normer och värderingar formas från flera håll, dels utifrån hur samhället styrs och är uppbyggt, dels grundat i religion och filosofi.

Människans hjärna har utvecklats till att känna empati. Spegelneuron⁷⁹ medför att människor lever i en gemensam mellanmännisklig kontext, där vi intuitivt kan uppfatta andras känslor, handlingar och avsikter.⁸⁰ När människor interagerar med varandra skapas mönster för hur man ska bete sig mot varandra inom familjer, företag, föreningar och samhället i stort. Dessa beteendemönster är uppbyggda genom många generationer.

Utifrån ovanstående, och inför en framtid som i allt högre grad kan komma att präglas av användning av AI, ställs vi inför viktiga frågeställningar. Hur säkerställer vi att tekniska, autonoma och självlärande system fattar livsviktiga beslut utifrån den etik som vi förespråkar? Och vems etik är det som gäller? Normer och värderingar skiljer sig i viss mån åt mellan individer, men i högre grad mellan olika samhällen och länder. Därför kan det vara berättigat att fråga: Kan vi i Sverige utan anpassning använda AI-algoritmer som är framtagna i andra delar av världen? Och hur säkerställs att underlaget för besluten (den samlade data som AI baserar besluten på) inte bygger på bias/snedvridning och diskriminering? Eller uteblivet samtycke från de människor vilkas data använts?

7.7 Etiska riktlinjer

På EU-nivå har *Europeiska Kommissionens oberoende expertgrupp på hög nivå för AI-frågor* tagit fram etiska riktlinjer för tillförlitlig AI.⁸¹ I expertgruppen ingår från Sverige Virginia Dignum, professor i social och etisk AI vid Umeå universitet, samt Fredrik Heintz, docent i datavetenskap vid Linköpings universitet. I dessa riktlinjer som publicerades i april 2019 anges att tillförlitlig AI har tre komponenter som bör finnas med under systemets hela livscykel:

- a) den bör vara laglig,

⁷⁹ Ett spegelneuron är en nervcell som avger nervimpulser både när individen utför en viss handling, och även när individen ser samma handling utföras av en annan individ. Hos människor har de hittats i Brocas område (finns i pannloben) och i undre halvan av hjässloben (lobulus parietalis inferior) i hjärnan.

⁸⁰ Rolf Nilzén. Teorin om spegelneuron förklarar förmågan till empati. Läkartidningen (2008-08-05) <https://lakartidningen.se/aktuellt/kultur-2/2008/08/teorin-om-spegelneuron-forklarar-formagan-till-empati/>

⁸¹ European Commission. Draft Ethics guidelines for trustworthy AI. (2018-12-18) <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>

- b) den bör vara etisk,
- c) den bör vara robust ur både teknisk och samhällelig synvinkel.

Om det i praktiken uppstår spänningar mellan dessa tre komponenter bör samhället sträva efter att lösa dem.

Expertgruppens etiska riktlinjer består av tre delar. Först fastställs de grundläggande rättigheter, principer och värderingar som AI ska respektera. Utifrån dessa listas sedan kraven på tillförlitlig AI och ges en överblick över de tekniska och icke-tekniska metoder som kan användas för att genomföra detta. Slutligen presenteras en konkret men inte uttömmande lista för bedömning av tillförlitlig AI. Listan ska sedan anpassas till specifika användningsområden.

Riktlinjerna är avsedda för alla berörda parter som utformar, utvecklar eller använder AI, t.ex företag, organisationer, forskare, offentliga tjänsteleverantörer, institutioner och andra organ. Som stöd för bedömning av tillförlitlig AI, vid utveckling och förvaltning, omfattar riktlinjerna även en checklista som kommer att utvärderas och vid behov revideras.

I EU-riktlinjerna^{82,83} anges att utveckling, spridande och användande av AI-system ska ske på ett sätt som respekterar de etiska principerna om respekt för människans autonomi, skadeförebyggande, rättvisa och förklarbarhet. Särskild uppmärksamhet bör riktas mot situationer som rör mer sårbara grupper, och på situationer som utmärks av bristande balans i fråga om inflytande eller information. AI-system kan visserligen innebära stora fördelar för enskilda och för samhället, men de kan också medföra vissa risker och negativa konsekvenser. Följaktligen måste lämpliga åtgärder vidtas för att minska dessa risker när så är lämpligt.

När detta skrivs har EU-kommissionens ordförande Ursula von der Leyen pekat ut digitalisering som en av hörnstenarna för EU:s arbete framåt, med prioriterade investeringar på AI, men också med nödvändiga anpassningar av det juridiska regelverket. När det gäller det sistnämnda är det rimligt att anta att dessa anpassningar baseras på de etiska riktlinjer som beskrivs ovan. Samma riktlinjer anger att det är av vikt att se till att utveckling, spridning och användning av AI-system uppfyller kraven för tillförlitlig AI:

- 1) mänsklig handling/inverkan samt tillsyn,
- 2) teknisk robusthet och säkerhet,
- 3) integritet och dataförvaltning,
- 4) transparens,
- 5) mångfald, icke-diskriminering och rättvisa,
- 6) samhällets och miljöns välbefinnande, samt

⁸² Europeiska Kommissionen. Artificial intelligence: Commission takes forward its work on ethics guidelines. (2019-04-08) https://europa.eu/rapid/press-release_IP-19-1893_en.htm?locale=en

⁸³ Europeiska Kommissionen. Ethics guidelines for trustworthy AI. (2019-04-08) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

7) ansvarsskyldighet.

Det är också av betydelse att den individ som berörs av AI i en viss situation ska ha kännedom om detta.

AI är ännu i sin linda och tillämpningen av riktlinjerna kommer att kräva fortsatta diskussioner och eventuella anpassningar. Som konstaterades av den Europeiska Kommissionens oberoende expertgrupp på hög nivå för AI-frågor finns det spänningar mellan olika principer och krav. Kompromisser och lösningar för detta behöver därför identifieras, utvärderas, dokumenteras och kommuniceras fortlöpande.

AI förutsätter stora mängder data, gärna från flera länder. För att tillgodose behovet av att gränsöverskridande kunna dela relevanta data till AI-utveckling och annan forskning/innovation inom e-hälsa föreslår den nya EU-kommissionen etablering av ett European Health Data Space (EHDS). Genom denna funktion ska den enskilda medborgaren samtidigt garanteras kontroll över sin egen persondata.⁸⁴ Kommissionen arbetar nu tillsammans med medlemsstaterna och andra berörda intressenter för att ta fram ett förslag till styrning och infrastruktur för EHDS. Från Sverige deltar E-hälsomyndigheten tillsammans med andra svenska myndigheter i en så kallad Joint Action inom EU Health Programme för etableringen av EHDS.

Vid det svensk-franska innovationspartnerskapets möte om AI i april 2019 lyftes frågan hur etiskt det är att inte dela data. Eftersom utveckling av AI-applikationer är beroende av stora mängder data kan konsekvensen av utebliven datadelning vara att vi inte erhåller de nyttor som AI annars borde kunna ge inom hälso- och sjukvård eller inom andra delar av välfärden. Ett till synes framgångsrikt exempel på särskilt beslut om att dela data är det så kallade Chapter 55 i Massachusetts, USA.⁸⁵ I detta fall rörde det sig om en delstatslag som införts för att ge myndigheterna rätt att dela medborgarnas person- och hälsodata och samköra register efter att antalet opioid-relaterade dödsfall ökat drastiskt. Syftet med lagändringen var att kartlägga problematiken och ta fram ett åtgärds paket.

En annan typ av frågeställning kan vara hur etiskt lämpligt det är att avstå från breddinförande av en framgångsrik AI-lösning, till exempel bilddiagnostik inom bröstcancer^{86,87} respektive hudcancer⁸⁸, om den uppvisar bättre medicinskt utfall än konventionell diagnostik?

Utöver EU:s etiska riktlinjer för AI har även andra internationella sammanslutningar formulerat riktlinjer av liknande slag när det gäller användande av AI och data.

⁸⁴ Stella Kyriakides. Europeiska Kommissionen. Commissioner-designate for Health. (2019-09-10) https://ec.europa.eu/commission/sites/beta-political/files/mission-letter-stella-kyriakides_en.pdf

⁸⁵ Massachusetts Department of Public Health. The Massachusetts opioid epidemic. A data visualization of findings from the Chapter 55 report. <https://chapter55.digital.mass.gov/>

⁸⁶ Karolinska Universitetssjukhuset, Karolinska Institutet. AI kan identifiera bröstcancer lika bra som en röntgenläkare. (2020-09-02) <https://www.karolinska.se/om-oss/centrala-nyheter/2020/09/ai-kan-identifiera-brustcancer-lika-bra-som-en-rontgenlakare/>

⁸⁷ Ania Obminska, Ny teknik. Googles ai bättre än läkare på att hitta bröstcancer. (2020-01-02) <https://www.nyteknik.se/digitalisering/googles-ai-battre-an-lakare-pa-att-hitta-brustcancer-6982975>

⁸⁸ Anna Orring, Ny teknik, Artificiell intelligens avslöjar hudcancer. (2017-01-25) <https://www.nyteknik.se/innovation/artificiell-intelligens-avslojar-hudcancer-6820557>

Europarådets *Ad Hoc-kommitté för AI* har i uppdrag att, med hänsyn tagen till Rådets standarder relevanta för design, utveckling och tillämpning av digitala teknologier bidra till mänskliga rättigheter, demokrati, lagstiftning, jämställdhet, rättigheter för människor med funktionsnedsättningar, samt hållbara samhällen.⁸⁹ Sverige, som är ett av 47 medlemsländer i Europarådet, är i Ad Hoc-kommittén för AI representerat av Magnus Nordström, Utrikesdepartementet.

OECD (*Organisation for Economic Co-operation and Development*) har fastställt AI-principer som bland annat Sverige anslutit sig till. Dessa *OECD AI Principles* fokuserar på hur regeringar och andra aktörer kan bidra till skapandet av en människocentrerad och tillförlitlig (*trustworthy*) AI, och framhåller därför betydelsen av mänskliga rättigheter och integritet, transparens och förklarbarhet, robusthet, säkerhet, samt ansvarstagande.⁹⁰

Även WHO beskriver hur digitalisering och AI ger möjlighet att i positiv mening revolutionera vård, forskning och folkhälsa, samtidigt som den snabba utvecklingen väcker frågor och farhågor av etisk, legal och social art, till exempel rörande likvärdig tillgång, integritet, korrekt användning, samt ansvar.⁹¹

⁸⁹ Europeiska Kommissionen. CAHAI - Ad hoc Committee on Artificial Intelligence <https://www.coe.int/en/web/artificial-intelligence/cahai>

⁹⁰ OECD. OECD AI Principles overview. <https://oecd.ai/ai-principles>

⁹¹ WHO. Big data and artificial intelligence. <https://www.who.int/ethics/topics/big-data-artificial-intelligence/en/>

8. AI och e-hälsa

I detta avsnitt beskrivs och diskuteras e-hälsa utifrån de tre områdena hälso- och sjukvård, socialtjänst (kommunal omsorg) samt hälso-applikationer i form av allmänna produkter utan formell koppling till de förstnämnda områdena. Med anledning av pågående pandemi redogörs också preliminärt för hur AI kan användas för att bekämpa covid-19.

8.1 AI inom hälso- och sjukvård

Inom hälso- och sjukvården är det många gånger en stor fördel med tidig upptäckt av sjukdom innan den bryter ut, det som kallas diagnostisk prediktion. Detta visar sig nu framför allt i medicinsk bildanalys att AI kan ge högre effektivitet och träffsäkerhet än den traditionella okulära metoden.

Om sjukdomar på detta sätt upptäcks och kan behandlas i tidigt skede innebär det vanligtvis vinster i form av minskat mänskligt lidande, men även en effektivare sjukvård med ekonomisk besparing i flera led. AI är också aktuell för triagering i syfte att individen vid kontakt med vården hamnar på ”rätt nivå” i sitt vårdmöte. Ett närbesläktat exempel är ett AI-verktyg baserat på maskininlärning som testas som beslutsstöd inom Region Uppsala. Forskningsprojektet syftar till att avgöra vilken patient som ska få hjälp först i lägen när ambulanserna inte räcker till⁹². Verktøget bygger på historiska data avseende patienters sjukdomstillstånd och hur det sen gick för dem. I den nya lösningen kan flera parametrar inkluderas i bedömningen än utan AI-verktyget. Den ska testas i situationer när antalet ambulanser inte räcker till. Målsättningen på längre sikt är att AI-verktyget ska kunna användas för att hänvisa patient till rätt vårdnivå.⁹³

I en rapport från Frost & Sullivan Institute (*Artificial Intelligence Market — Key Application Areas for Growth in Healthcare IT, Forecast to 2022*) undersöks de viktigaste AI-leverantörerna som har prognoser för globala intäkter inom hälso- och sjukvård och som utnyttjar AI för att öka produktfunktionaliteterna.⁹⁴ Totalt förväntas denna marknad växa till drygt 6 miljarder USD vid en sammansatt årlig tillväxttakt på drygt 68 procent mellan 2018 och 2022.⁹⁵

Förutom att användas i den kliniska vården bedöms AI kunna göra nytta även i de administrerande delarna. I en artikel i *Healthcare innovation*⁹⁶ beskriver forskare från global

⁹² Lisa Blohm. Dagens Medicin. AI testas vid prioritering av ambulanser. (2020-06-15)

<https://www.dagensmedicin.se/artiklar/2020/06/15/ai-testas-vid-prioritering-av-ambulanser/>

⁹³ Douglas Spangler et al. PLOS. A validation of machine learning-based risk scores in the prehospital setting. (2019-12-13) <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0226518>

⁹⁴ Frost and Sullivan. Optimizing Quality and Efficiency of Healthcare Delivery with AI. (september 2018) https://go.frost.com/NA_PR_MFernandez_K26D_AIHealthcareIT_Sep18

⁹⁵ Frost and Sullivan. Artificial Intelligence in Healthcare Takes Precision Medicine to the Next Level. (2018-09-13) <https://ww2.frost.com/news/press-releases/artificial-intelligence-healthcare-takes-precision-medicine-next-level/>

⁹⁶ David Raths. Healthcare innovation. Predicting Where AI Will Have the Biggest Impact on Healthcare. (2019-03-08) <https://www.hcinovationgroup.com/analytics-ai/article/21071336/predicting-where-ai-will-have-the-biggest-impact-on-healthcare/>

hälsovårdsindustri att de viktigaste initiala förändringsområdena nog inte kommer att vara i direkt klinisk verksamhet. USA har redan sett fördelarna med automatisering för att flytta rutinuppgifter till maskiner. Med AI kan enligt artikeln icke rutinartade uppgifter flyttas till maskiner. En uppskattning är att nästan 30 procent av läkarkapaciteten idag kan flyttas till patienter (i självbetjäningssläge) eller till smarta verktyg, eller en kombination av de två. Det ger hopp om att kunna överbrygga klyftan mellan medborgarnas/patienternas efterfrågan på vård och de begränsade resurserna. Man kan kalla det växelverkan mellan människa och maskin/AI.

I samma artikel hävdas att AI inte kommer att vara i *stället för*, utan *därutöver*. Maskinen gör människan bättre på vad hon gör (se *augmented intelligence* ovan), men samtidigt kan maskinen lära sig av människan. Detta innebär att anställda i vården måste lära sig att arbeta annorlunda, med maskinen som medarbetare. De behöver inte bli bättre på användningen av teknik utan snarare bättre på att ställa rätt frågor till maskinen. Radiologin och patologin är områden där mycket arbete pågår med *computer aided diagnosis* (CAD) och mönsterigenkänning (*pattern recognition*). För att få optimalt underlag för diagnosticering behöver bildanalysen kombineras med andra former av kunskapskällor. AI är ett utmärkt verktyg för detta. Utöver patologi och radiologi är även andra bildtolkande diagnostiker intressanta för denna AI-utveckling som ögonbottenfotografering, gastrointestinala skopier och även dermatologiska förändringar.

Fredrik Heintz, AI-expert vid Linköpings universitet, menar att även ledarskapet i organisationerna måste utvecklas för att kunna arbetsleda med färre men mer kompetenta medarbetare vilket också ställer större krav, och där allt mer av arbetet utförs av team bestående av både människor och AI-system. Här kommer organisationen behöva fokusera på att hitta den bästa kombinationen så att medarbetarna samarbetar med intelligenta maskiner och system på bästa sätt. Samtidigt understryker Fredrik Heintz att det tyvärr är en vanlig missuppfattning att AI är bra på allt. AI-verktyg kan framstå som smarta men är än så länge smarta endast inom begränsade områden.

För att uppnå AI:s och digitaliseringens potential i allmänhet förutsätts en genuin verksamhetsutveckling – digital transformation – och inte enbart införande av ny teknik (”digitisering”). Detta förutsätter förutom tydliga strategier och färdplaner för att tillämpa AI och annan digitalisering även ändrade arbetsprocesser, kompetensutveckling av personal och ett ökat fokus på effektiv och säker datahantering och dataanalys. Därutöver sätts en del förhoppningar till AI som möjliggörare av en betydligt bättre vårdflödesoptimering.

I Sverige förekommer AI ännu främst på forsknings- och utvecklingsnivå med exempel som pilotprojekt inom bilddiagnostisering av bröstcancer eller benfrakturer.⁹⁷ När det gäller det sistnämnda exemplet så inledde Inspektionen för Vård och omsorg, IVO, en

⁹⁷ Jesper Cederberg. Läkartidningen. Första patienterna på Danderyds sjukhus bedömda med AI. (2019-10-23) <https://www.lakartidningen.se/Aktuellt/Nyheter/2019/10/Forsta-patienterna-pa-Danderyds-sjukhus-bedomda-med-AI/>

granskning i inledningen av 2020 vilket beskrivs närmare i avsnitt 9.2 AI och medicintekniska regelverket.⁹⁸

Det finns samtidigt några svenskutvecklade produkter på marknaden vilka kan betraktas som AI-baserade. Exempel på dessa är *Kontigo Care* för beroendevård samt *Coala Heart Monitor*, en hjärtmonitor som utifrån algoritmer möjliggör patient- och användarcentrerade hjärtutredningar.

Ett annat exempel på AI inom sjukvård är privata vårdgivare som erbjuder en digital ingång via mobilapplikation. Där sker en triagering som syftar till att lotsa patienten till rätt vårdnivå. Triageringen bygger på tolkning av patienternas svar som sedan leder till automatiska följdfrågor. Valet av följdfrågor bygger på inbyggda algoritmer.

Ett nytt växande område där AI är av stor betydelse är precisionsmedicin⁹⁹. I Sverige finansierade *Swelife*¹⁰⁰ 2017 en förstudie, *Genomic Medicine Sweden*¹⁰¹, som koordinerar och utvecklar förutsättningarna för att Sverige ska kunna ligga i framkant inom precisionsmedicin. Arbetet kommer att fortsätta med stöd av Vinnova. Genomik, med storskalig analys av arvsmassan, är en grundbult inom precisionsmedicinen.

På regeringens uppdrag har Socialstyrelsen, i samverkan med bland annat E-hälsomyndigheten, undersökt i vilken omfattning landets regioner har infört AI-system inom klinisk verksamhet. Kartläggningen, som återrapporterades till Socialdepartementet i oktober 2019, visar att i princip samtliga regioner har infört, eller planerar att inom kort införa, ett eller flera AI-system inom hälso- och sjukvården.¹⁰²

8.2 AI inom socialtjänst

AI har även börjat användas inom socialtjänsten. Ett område där både personliga och samhällsvinster kan göras är inom den kommunala missbruksvården där nykterhetstest på distans erbjuds dem som så önskar. På ett enkelt sätt kan den enskilde med hjälp av en smart mobiltelefon göra sitt nykterhetstest på distans. Vanligen tas nykterhetstest 3-5 gånger per dag. Samtidigt med nykterhetstestet tas ett självporträtt, en selfie. Resultatet överförs till socialtjänsten så att personens stödresurs antingen kan agera eller får ett underlag för kommande behandlingsmöten.

Över 100 kommuner i Sverige använder idag detta stöd. AI används för att genom bildigenkänning avgöra om det är rätt person som gör nykterhetstestet. Funktionaliteten har också utvecklats genom att AI används för att förutsäga återfall i missbruk. Genom att brukaren gör dagliga självskattningar via formulär i sin mobiltelefon tillsammans

⁹⁸ Linda Berglund. Dagens Medicin. Ivo granskar Danderyds AI-projekt. (2020-02-03) <https://www.dagensmedicin.se/artiklar/2020/02/03/ivo-granskar-danderyds-ai-projekt/>

⁹⁹ Precisionsmedicin definieras som behandlingar skraddarsydda för en patient baserat på den enskilde patientens unika genetiska profil och specifika äggviteämnen (biomarkörer).

¹⁰⁰ <https://swelife.se/om-swelife/>

¹⁰¹ Genomic Medicine Sweden-projektet syftar till att bygga en ny typ av infrastruktur som möjliggör världsledande diagnostik och precisionsmedicin i Sverige. Projektet är ett samarbete mellan sju regioner och sju universitet med medicinska fakulteter.

¹⁰² Socialstyrelsen. Digitala vårdtjänster och artificiell intelligens i hälso- och sjukvården (oktober 2019) <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2019-10-6431.pdf>

med resultat av nykterhetstesterna, kan stödet följa brukarens beteende och därigenom uppfatta eventuella avvikelser samt möjliga tendenser till återfall i tid.¹⁰³

Andra exempel där AI-liknande lösningar används i socialtjänsten är automatiserade beslut om bistånd. SKR beskriver i sin *Omvärldsrappport digitalisering maj 2020* hur transparens har hamnat alltmer i fokus vid användning av algoritmer, i syfte att värna rättssäkerhets- och insynsfrågor för medborgaren.¹⁰⁴ I rapporten hänvisas till en studie om ramverk för algoritmiskt ansvar och öppenhet, från Europeiska parlamentets forskningstjänst: *A governance framework for algorithmic accountability and transparency*, EPRS - European Parliamentary Research Service, 2019.¹⁰⁵

Automatiserade beslut diskuteras utförligare utifrån legala aspekter i avsnitt 9.3.

8.3 Hälsoappar

En annan del av AI-sektorn inom e-hälsa är de appar som i växande antal finns tillgängliga och som faller in under kategorin medicintekniska produkter. Dessa används ofta av användare utan vårdprofessionens direkta medverkan. Detta förutsätter att användaren har god kännedom om hur appen ska användas och förstår dess begränsningar.

Det förekommer också allt fler appar som samlar in individens vitalparametrar (blodtryck, syresättning, kroppstemperatur mm). Dessa data lagras av leverantörerna under olika lång tid och förvaras även utanför de egna landsgränserna. Nyligen visades i en svensk studie att flera populära tränings-appar samlade in mer persondata än vad användaren givit tillåtelse till, exempelvis kontaktlistor och telefonloggar.¹⁰⁶

Data är en förutsättning för AI, behöver då finnas i stora mängder och inte innehålla bias, det vill säga snedvridning. Det innebär i så fall att många människor måste vara beredda att dela med sig av sina data. Tillit till att denna data hanteras på ett säkert sätt för de ändamål som man har gett sitt tillstånd till, är av avgörande betydelse för människornas vilja att dela sina eventuellt känsliga personuppgifter. Därför är det viktigt att följa utvecklingen av hälsoapparna och hur apparnas leverantörer hanterar den information som samlas in.

Begreppet m-hälsa (*mHealth*) står för mobil hälsa och kan sägas utgöra en underavdelning till e-hälsa. Hälsoappar kan i sin tur ses som en del av m-hälsa. EU anger att det finns mer än 100 000 m-hälsoappar på marknaden, och följaktligen kan denna samlade – och växande – resurs vara betydelsefull för den digitala

¹⁰³ <https://skr.se/integrationsocialomsorg/socialomsorg/digitaliseringinomsocialtjansten/utvecklingsarbetendigitaliseringsocialtjansten/projektasikanartificiellintelligensanvandasforattbattreforstabedomningsinstrumetetasiformissbruksochberoendevarden.27486.html>

¹⁰⁴ SKR. Omvärldsrappport digitalisering maj 2020. <https://webbutik.skr.se/sv/artiklar/omvarldsrappport-digitalisering-maj-2020.html>

¹⁰⁵ European Parliament Think Tank. A governance framework for algorithmic accountability and transparency. (2019-04-04)

[https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU\(2019\)624262](https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)624262)

¹⁰⁶ M. Hatamian et al. Karlstad Universitet. A Multilateral Privacy Impact Analysis Method for Android Apps. (2019) <http://kau.diva-portal.org/smash/record.jsf?pid=diva2%3A1323331&dswid=-2860>

transformationen av hälso- och sjukvården.¹⁰⁷

8.4 Covid-19 och AI

Under arbetet med denna rapport drabbades världen av en ny pandemi, covid-19 (*coronavirus infectious disease -19*). Samtidigt som det är alldeles för tidigt att här kunna redogöra för hur AI har kunnat användas på ett effektivt sätt mot denna pandemi så vore det fel att helt bortse från covid-19 i denna rapport. Följaktligen ger vi här en kortfattad överflygning för att visa på exempel där AI på olika sätt kan användas i kampen mot covid-19-pandemin och dess effekter. För en utförligare genomlysning av covid-19 och e-hälsa, i relation till de så kallade megatrenderna, hänvisas till E-hälsomyndighetens Årsrapport 2020.¹⁰⁸

När det gäller det inledande utbrottet av covid-19 så anses det kanadensiska AI-företaget BlueDot ha varit först med att förutspå internationell spridning, och redan i januari 2020 var företaget först med att publicera en vetenskaplig artikel om detta.^{109,110}

Massachusetts Institute of Technology (MIT) i USA har utvecklat en AI-algoritm som kan skilja en vanlig hosta från en hosta orsakad av covid-19. AI-algoritm har sedan tränats på att lära sig skilja på hostningar från en person med covid-19 och en person som hostar av andra skäl. Algoritmen utvecklad vid MIT, har nästan alltid rätt. Träffsäkerheten ligger i skrivande stund på 98 procent med att identifiera covid-19 hos personer med positivt coronatest.¹¹¹

Även i Sverige ses satsningar på AI mot covid-19. Forskare vid KTH och Karolinska universitetssjukhuset undersöker om bildanalys kan öka möjligheterna att sätta rätt diagnos vid covid-19. Resultaten får betydelse både för kommande pandemier och för sjukhus utan specialiserad vård. Studier visar att AI kan skilja covid-19 från andra typer av lunginflammation. Målet är att förbättra träffsäkerheten i diagnostiken ytterligare.¹¹²

SOS Alarm undersöker tillsammans med det danska företaget Corti möjligheten att skapa lägesbilder (hotspots) av covid-19, detta med hjälp av en AI-funktion

¹⁰⁷ Europeiska Kommissionen. Shaping Europe's digital future. Policy – mHealth. (2020-10-29)

<https://ec.europa.eu/digital-single-market/en/mhealth>

¹⁰⁸ Årsrapport 2020 – Pandemi och e-hälsa. Återrapportering enligt E-hälsomyndighetens regleringsbrev 2020 (S2020/05629/FS)

<https://www.ehalsomyndigheten.se/globalassets/dokument/arsrapporter/arsrapport-2020---pandemi-och-e-halsa.pdf>

¹⁰⁹ Eric Niiler. Wired. An AI Epidemiologist Sent the First Warnings of the Wuhan Virus (2020-01-25)

<https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/>

¹¹⁰ J Travel. NCBI. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. (2020-01-14) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7107534/>

¹¹¹ Jennifer Chu. MIT. Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs. (2020-10-29) <https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029>

¹¹² Peter Ardell. KTH. Bildanalys och AI ska rädda fler coronapatienter. (2020-05-20)

<https://www.kth.se/aktuellt/nyheter/bildanalys-och-ai-ska-radda-fler-coronapatienter-1.987755>

för att detektera 112-samtal där covid-19 kan misstänkas utifrån patientens symtombild.¹¹³

Vinnova bidrar till svensk forskning mot covid-19, bland annat genom stöd till projekt för AI-baserad identifiering av plasmadonatorer för covid-19-antikroppar¹¹⁴, men också genom bilateralt samarbete med Indien inom hälsa och AI, inklusive gemensamma satsningar mot covid-19.¹¹⁵

Kampen mot covid-19 bygger till stor del på information och data. I syfte att underlätta informationssökningen har Google utvecklat ett AI-verktyg baserat på naturlig språkförståelse, *natural language understanding* (NLU).^{116,117}

Europarådets Ad Hoc-kommitté för AI har sammanställt en resumé över redovisade tillämpningar av AI mot covid-19, med samtidig hänvisning till Europeiska sociala stadgan (*European Social Charter*) samt Konvention 108 om skydd för enskilda vid automatisk databehandling av personuppgifter.^{118,119}

OECD menar att AI har en avgörande roll i alla aspekter av kampen mot covid-19 och understryker vikten av att beslutsfattare uppmuntrar och underlättar forskares delande av data, samt att de också ska ha tillfredsställande resurser för databehandling. Även OECD påminner om nödvändig hänsyn till riktlinjer rörande etik och säkerhet, *OECD AI Principles*.¹²⁰

Avslutningsvis kan nämnas att Europeiska kommissionen investerar i användandet av AI för att effektivisera diagnostik och behandling av covid-19, till exempel genom en spridd satsning via *Connecting Europe Facility* (CEF) på AI-baserad datortomografi för snabbare lungdiagnostik.¹²¹

Flera exempel på hur AI kan användas mot covid-19 har presenterats, bland annat i april 2020, av en internationell forskargrupp (*Mapping the landscape of Artificial Intelligence applications against COVID-19*)¹²² som anger följande möjligheter, och hinder:

¹¹³ SOS Alarm. SOS Alarm skapar lägesbild av covid-19 med hjälp av AI (2020-06-16) <https://www.sosalarm.se/pressrum/pressmeddelanden/2020/sos-alarm-skapar-lagesbild-av-covid-19-med-hjalp-av-ai/>

¹¹⁴ Vinnova. AI-baserad identifiering av plasmadonatorer för Covid-19-antikroppar (2020-05-27) <https://www.vinnova.se/p/ai-baserad-identifiering-av-plasmadonatorer-for-covid-19-antikroppar/>

¹¹⁵ Vinnova. Bilateralt samarbete med Indien inom hälsa och AI inför utmaningar som Coronaviruset. (augusti 2020) <https://www.vinnova.se/e/samverkan-mellan-vinnova-och-indien/bilateralt-samarbete-med-indien-inom-halsa-och-ai/>

¹¹⁶ Keith Hall. Google AI Blog. An NLU-Powered Tool to Explore COVID-19 Scientific Literature. (2020-05-04) [An NLU-Powered Tool to Explore COVID-19 Scientific Literature](https://ai.googleblog.com/2020/05/an-nlu-powered-tool-to-explore-covid-19-scientific-literature.html)

¹¹⁷ Google. Get answers to complex scientific questions related to COVID-19. (2020) <https://covid19-research-explorer.appspot.com/>

¹¹⁸ Europeiska Kommissionen. AI and control of COVID-19. (2020) <https://www.coe.int/en/web/artificial-intelligence/ai-covid19>

¹¹⁹ Europeiska Kommissionen. (1985-10-01) <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108>

¹²⁰ OECD. Using artificial intelligence to help combat COVID-19. (2020-04-23) <https://www.oecd.org/coronavirus/policy-responses/using-artificial-intelligence-to-help-combat-covid-19-ae4c5c21/>

¹²¹ Europeiska Kommissionen. Using AI to fast and effectively diagnose COVID-19 in hospitals. (2020-05-19) <https://ec.europa.eu/digital-single-market/en/news/using-ai-fast-and-effectively-diagnose-covid-19-hospitals>

¹²² Joseph Bullock et al. United Nations Global Pulse. Mapping the landscape of Artificial Intelligence applications against COVID-19. (2020-04-23) <https://arxiv.org/pdf/2003.11336v2.pdf>

- AI erbjuder en mängd olika potentiella applikationer för att tillgodose medicinska och samhällsliga behov som uppkommit på grund av covid-19, men få av dessa har nått en tillräcklig mognadsgrad för att idag ha en tillämpbar effekt.
- Utifrån ett molekylärt perspektiv kan AI användas för att uppskatta virusproteiners struktur, identifiera existerande läkemedel som kan vara användbara, föreslå nya substanser, identifiera tänkbara vaccinkandidater, förbättra diagnostik samt öka förståelsen för virusets infektionsförmåga och patogenicitet¹²³.
- Utifrån ett kliniskt perspektiv kan AI stödja diagnostik av covid-19 genom medicinsk bildanalys, erbjuda alternativ att spåra sjukdomsutveckling genom icke-invasiv metodik, samt bidra till medicinsk prognos baserat på stora mängder vårddata.
- Utifrån ett samhällsperspektiv har AI redan använts i ett flertal olika epidemiologiska modeller genom empirisk data, för att förutspå antal infekterade fall utifrån olika strategier för allmänheten, men också för att identifiera likheter och skillnader i pandemins utveckling i olika regioner. AI kan också bidra till att undersöka vidden och spridningen av felaktig information och medveten desinformation.
- Att dela och tillhandahålla data och modeller är avgörande för att AI på snabbaste och effektivaste sätt ska kunna bidra till att motverka covid-19-pandemin och dess effekter.
- Samtidigt är det nödvändigt att ta hänsyn till regelverk och kvalitetssäkring för att säkerställa säkerhet och minimera risk för skada som för övrigt gäller för all AI-utveckling.
- Internationell samverkan baserad på multidisciplinär och öppen forskning är nödvändig för att påskynda tillämpning av forskningsrön och ny kunskap i globala lösningar som kan anpassas till lokal kontext.

¹²³ Patogenicitet är den potentiella kapaciteten hos specifika mikroorganismer (exempelvis virus, bakterier) att orsaka sjukdom. Virulens används ofta synonymt men betecknar mer specifikt graden av patogenicitet.

9. Juridiken

9.1 Regelverket i korthet

- Det finns ingen särskild författning som reglerar ansvarsförhållanden, beslut, transparens och risker med AI och maskininlärning. Den reglering som finns är till stor del teknikneutral. De författningar som är tillämplbara för AI-användning inom hälso- och sjukvård och socialtjänst reglerar t.ex. olika aspekter av ansvar, personuppgiftsbehandling och informationssäkerhet.
- EU:s dataskyddsförordning, *General Data Protection Regulation (GDPR)*¹²⁴ reglerar hur personuppgifter får användas inom hela unionen. I Sverige finns flera författningar som kompletterar förordningen.
- Patientdatalagen (2008:355) kompletterar dataskyddsförordningen och reglerar vårdgivares behandling av personuppgifter inom hälso- och sjukvården. I Socialstyrelsens föreskrifter och allmänna råd (HSLF-FS 2016:40) om journalföring och behandling av personuppgifter inom hälso- och sjukvården finns detaljerade bestämmelser om hur vårdgivaren ska ta ansvar för informationssäkerheten i verksamheten. Det finns också en lag (2001:454) om behandling av personuppgifter inom socialtjänsten, som även den kompletterar dataskyddsförordningen.
- I lagen (2018:1174) om informationssäkerhet för samhällsviktiga och digitala tjänster finns bestämmelser som syftar till att uppnå en hög nivå på säkerheten i nätverk och informationssystem för samhällsviktiga tjänster, som t.ex. hälso- och sjukvård.
- Hälso- och sjukvårdsverksamhet ska enligt hälso- och sjukvårdslagen (2017:30) bedrivas så att kraven på en god vård uppfylls. Det innebär till exempel att vården ska tillgodose patientens behov av trygghet, kontinuitet och säkerhet (5 kap. 1 §). Där verksamheten bedrivs ska det bland annat finnas den utrustning som behövs för att god vård ska kunna ges (5 kap. 2 §).
- Vårdgivaren, den som bedriver verksamheten, ansvarar enligt 3 kap. 1 § patientsäkerhetslagen (2010:659) bland annat för att planera, leda och kontrollera verksamheten så att den lever upp till kravet på god vård. För att uppfylla detta ansvar ska vårdgivaren enligt Socialstyrelsens föreskrifter och allmänna råd (SOSFS 2011:9)¹²⁵ bedriva ett systematiskt kvalitetsarbete i form av att fastställa rutiner för arbetet och göra riskanalyser. Riskanalyser ska t.ex. göras före förändringar av en verksamhet och innan nya arbetssätt eller

¹²⁴ Europaparlamentets och rådets förordning (EU) 2016/679 av den 27 april 2016 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter och om upphävande av direktiv 95/46/EG (allmän dataskyddsförordning).

¹²⁵ Socialstyrelsens föreskrifter och allmänna råd (SOSFS 2011:9) om ledningssystem för systematiskt kvalitetsarbete.

metoder börjar tillämpas. Föreskrifterna om systematiskt kvalitetsarbete ska tillämpas även av den som bedriver socialtjänst.

- Legitimerad hälso- och sjukvårdspersonal har enligt patientsäkerhetslagen ett personligt yrkesansvar som bl.a. innebär att de ska utföra sitt arbete i överensstämmelse med vetenskap och beprövad erfarenhet. Patienter ska ges sakkunnig och omsorgsfull hälso- och sjukvård som uppfyller dessa krav (6 kap. 1 §).
- Socialstyrelsen har meddelat föreskrifter (SOSFS 2008:1) om användning av medicintekniska produkter i hälso- och sjukvården.¹²⁶
- Kommunallagen (2017:725), som huvudsakligen reglerar förtroendevalda fullmäktiges och nämndernas verksamhet, tillåter inte automatiserade beslut.
- Förvaltningslagen (2017:900), som reglerar offentlig förvaltning, tillåter automatiserade beslut och möjliggör ett undantag från dataskyddsförordningens principiella förbud mot automatiserade beslut.
- Medicintekniska regelverket i form av *Medical Devices Regulation* (MDR) samt *In-vitro Diagnostic Medical Devices Regulation* (IVDR) anger att en medicinsk mjukvara måste vara fullt validerad innan den kan CE-märkas och släppas ut på marknaden.¹²⁷

9.2 AI och medicintekniska regelverket

När det gäller AI inom e-hälsa så är den grundläggande frågan om lagstiftningen tillåter användning av självlärande AI-system inom hälso- och sjukvården?

Frågan måste bland annat kopplas till EU:s förordning för medicintekniska produkter (*Medical Devices Regulation*, MDR 2017/745) och dess krav på produktens avsedda användning samt att det ska råda en rimlig balans mellan acceptabel risk och prestanda.

Läkemedelsverket¹²⁸ menar att det är svårt att ge ett enkelt svar på frågan eftersom AI är ett mycket brett område. AI finns inte specifikt omnämnt i det nuvarande regelverket eller den nya förordningen och än så länge finns inte heller någon tydlig vägledning från EU-kommissionen om hur AI ska hanteras juridiskt. Läkemedelsverkets tolkning blir därför att AI ska betraktas och hanteras som mjukvara. Vilka implikationer detta får beskrivs nedan.

Det som är speciellt med AI är dess förmåga att lära sig med tiden, vilket kan jämföras med en mjukvara som förbättras genom regelbundna uppdateringar. Skillnaden är att ett AI-system uppdaterar sig själv, utan inblandning och kontroll av någon tillverkare. Förändringen är också i princip oförutsägbar då en tillverkare inte kan förutsäga vilka

¹²⁶ <https://lakemedelsverket.se/malgrupp/Halso---sjukvard/Medicinteknisk-sakerhetsinformation/>

¹²⁷ Läkemedelsverket; kommunikation med enhetschef på Medicinteknik (2020-11-18)

¹²⁸ Läkemedelsverket; kommunikation med utredare på Medicinteknik (2019-09-05)

stimuli AI-systemet utsätts för och lär sig av. I fallet då en mjukvara uppdateras av tillverkaren, så är tillverkaren också ansvarig för och kan säkerställa att den uppdaterade mjukvaran är validerad.

Frågeställningen med AI blir då hur en tillverkare kan garantera att en AI-produkt är validerad genom hela sin livscykel, eftersom den förändras kontinuerligt. Tillverkarens möjlighet att ansvara för att AI-systemet är validerat kan endast säkerställas om systemet tränas med hjälp av stora mängder relevant data under kontrollerade former – sannolikt inte i skarp användning – samt under översyn av tillverkaren. När träningen är klar stängs inlärningsfunktionen av för ny verifiering och validering av systemet innan det tas i bruk.

Det medicintekniska regelverket är tydligt på den punkten att en medicinsk mjukvara måste vara fullt validerad innan den kan CE-märkas¹²⁹ och sedan släppas ut på marknaden.

Läkemedelsverkets bedömning i frågan blir följaktligen att ett AI-system som medicinteknisk produkt inte kan ha sin ”lärandefunktion” aktiverad när den används kliniskt, eftersom kravet på validering svårligen kan uppfyllas av tillverkaren.¹³⁰ En tänkbar nackdel med denna tolkning är att AI inte kan utnyttjas fullt ut på ett kontinuerligt sätt, och därmed förloras åtminstone delvis vinsterna med AI som ett ständigt lärande system. Här bör dock beaktas att regleringen tillkommit i syfte bland annat att skydda patienternas säkerhet.

Denna fråga beskrivs även i Socialstyrelsens rapport *Digitala vårdtjänster och artificiell intelligens inom hälso- och sjukvården* (sid 55 och sid 76).¹³¹ Där hänvisas till en rapport sammanställd av *British Standards Institution* (BSI), och *Medicines & Healthcare products Regulatory Agency* (MHRA) i Storbritannien, samt *Association for the Advancement of Medical Instrumentation* (AAMI), USA, där expertis, i syfte att minska patientsäkerhetsrisker när det gäller AI med kontinuerligt lärandesystem, rekommenderar att efter validering läsa algoritmerna i olika kontrollerade och definierade mjukvaruversioner.¹³² I rapporten påtalas dock värdet av att utveckla användarfall för att utvärdera denna typ av rutiner.

Intressant i detta sammanhang är att amerikanska *U.S. Food and Drug Administration* (FDA) utreder möjligheterna för ett livscykelbaserat regelverk som skulle kunna tillåta modifieringar av vissa kontinuerligt lärande AI-system, samtidigt som säkerheten i

¹²⁹ CE-märkning är en produktmärkning inom främst EU men även inom EES. CE står för Conformité Européenne ('i överensstämmelse med EG-direktiven'). En produkt med CE-märkning får säljas i EES-området utan ytterligare krav. Förutsättningarna för att få CE-märka en produkt är att: Produkten överensstämmer med grundläggande krav på exempelvis hälsa, säkerhet, miljö, samt att föreskriven kontrollprocedur har följts.

¹³⁰ Läkemedelsverkets skriftliga yttranden är inte ett beslut i rättslig mening utan ska ses som verkets bedömning vid aktuell tidpunkt (2019-09-05 respektive 2020-11-18) grundat på det material som skickats till Läkemedelsverket i samband med förfrågan. Om det, efter att Läkemedelsverket meddelat sin bedömning, framkommer ny information kan Läkemedelsverkets ställningstagande ändras.

¹³¹ Socialstyrelsen. *Digitala vårdtjänster och artificiell intelligens i hälso- och sjukvården*. (oktober 2019) <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2019-10-6431.pdf>

¹³² The emergence of artificial intelligence and machine learning algorithms in healthcare: Recommendations to support governance and regulation - Position paper. London, UK: Medicines & Healthcare products Regulatory Agency and BSI (2019) <https://www.bsigroup.com/globalassets/localfiles/en-gb/about-bsi/nsb/innovation/mhra-ai-paper-2019.pdf>

egenskap av medicinteknisk utrustning bibehålls.¹³³ I en förslagsskrift, tillika diskussionsunderlag, beskriver FDA hur AI-/ML-baserad mjukvara som utgör medicinteknisk utrustning (*Software as Medical Device*, SaMD) befinner sig på ett spektrum utifrån graden av risk för patienten, samtidigt som samma AI-system befinner sig på ett spektrum från låst eller fixerat till kontinuerligt lärande. Utifrån en matris bestående av tre olika medicinska risknivåer (Kritisk, Allvarlig, Icke allvarlig) och tre olika nivåer av vilken betydelse informationen från AI-systemet har (Behandling eller diagnos, Styra klinisk ledning, Informera klinisk ledning), erhålls en samlad risk-klassificering (ett slags ”risk-poäng”) som kan avgöra hur strikt valideringsproceduren måste vara. Som exempel nämns AI-baserade bildanalyssystem för bedömning av sår och sårsläkning (lägre risknivå) jämfört med system för bedömning av hudförändringar såsom melanom (högre risknivå), där det förstnämnda systemet, på lägre risknivå, kan antas ha större möjlighet att få behålla sin lärandefunktion under validering och kanske även i klinisk användning.¹³⁴

En observation är att den klassificering i fyra olika risknivåer för mjukvarubaserad medicinteknisk utrustning som beskrivs från FDA ovan har likheter med den klassificering som EU:s MDR innehåller.

Ett aktuellt exempel på AI-baserad diagnostik som berörts tidigare i denna fokusrapport rör den AI-lösning som utvecklats på Danderyds sjukhus, för bedömning av röntgenbilder. Eftersom denna AI-algoritm nu har integrerats mot bildarkiv och är användbar kliniskt så beslutade IVO i februari 2020 att öppna ett tillsynsärende. Motiveringen var att AI-algoritmen måste uppfylla kraven enligt MDR, det vill säga antingen vara CE-märkt (enligt de rutiner rörande validering som det kräver i sin tur) eller ses som egentillverkad, där vårdgivaren själv tar ansvar för patientsäkerheten vid användning av AI-lösningen.¹³⁵ I skrivande stund har inget utfall rapporterats i detta tillsynsärende, men i sin årliga rapport, *Vad har IVO sett 2019?*, anger IVO tydligt att man kommer följa hur vården inför och använder nya tekniska lösningar såsom AI.

Som framgår ovan så ingår även IVDR 2017/746 i det medicintekniska regelverket. I en nyligen publicerad kommentar till EU:s vitbok om AI menar författaren att IVDR är mer skraddarsytt för AI än MDR eftersom det förstnämnda är mer detaljerat och mer fokuserat på kvalitet och relevans hos data. Av den anledningen anføres att IVDR kan utgöra en bra startpunkt för att tydliggöra och implementera ett juridiskt ramverk för AI-applikationer av hög-risk-karaktär inom hälso- och sjukvård.¹³⁶

När det gäller digitala chatbotar i hälso- och sjukvården så utgör dessa inte ”allmän rådgivning” utan vård och behandling. Följaktligen måste vårdgivare finnas och journal upprättas. Detta medför att även chatbotar i hälso- och sjukvården bör ses i ljuset av

¹³³ U.S. FDA. Artificial Intelligence and Machine Learning in Software as a Medical Device (2020-10-05) <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>

¹³⁴ Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback (2019-04-02) <https://www.fda.gov/media/122535/download>

¹³⁵ Linda Berglund. Dagens Medicin. Ivo granskar Danderyds AI-projekt. (2020-02-03) <https://www.dagensmedicin.se/artiklar/2020/02/03/ivo-granskar-danderyds-ai-projekt/>

¹³⁶ A. Kiseleva. AI as a Medical Device: Is it Enough to Ensure Performance Transparency and Accountability? (2020-06-25) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3627741

MDR. De bestämmelser som gäller telefonrådgivning bör även kunna tillämpas på liknande sätt för chatbotar, men ifall det är tillåtet för den enskilde (läs: patienten) att vara anonym bortfaller många skyldigheter för vårdgivaren.

9.3 Automatiserade beslut

AI (eller AI-liknande system baserade på fasta algoritmer) kan användas för automatiserade beslut av olika slag. Här redogörs för lagstiftningen när det gäller användning av automatiserade beslut, i första hand inom socialtjänsten.

Dataskyddsförordningen (GDPR)

Enligt artikel 22.1 i dataskyddsförordningen (*General Data Protection Regulation*) ska den registrerade ha rätt att inte bli föremål för ett beslut som enbart grundas på automatiserad behandling, inbegripen profilering, vilket har rättsliga följder för honom eller henne eller på liknande sätt i betydande grad påverkar honom eller henne.

Undantag från rättigheten att inte bli föremål för automatiserat individuellt beslutsfattande gäller enligt dataskyddsförordningen för beslut som tillåts enligt nationell rätt som den personuppgiftsansvarige omfattas av och som fastställer lämpliga åtgärder till skydd för den registrerades rättigheter, friheter och berättigade intressen.

Förvaltningslagen

Enligt 28 § första stycket förvaltningslagen (2017:900) kan beslut fattas automatiserat. Bestämmelsen trädde i kraft den 1 juli 2018. Den förvaltningsrättsliga osäkerhet som tidigare funnits rörande myndigheternas rätt att, utan särskild tillåtande författningsreglering, fatta beslut automatiserat har således undanröjts med denna reglering.

Förvaltningslagen innehåller även bestämmelser om kommunikation och rätt att lämna uppgifter muntligen, rätt till omprövning och rätt till överklagande. Därutöver finns även generella krav på legalitet, objektivitet och proportionalitet.

Mot den bakgrunden har eSam¹³⁷ i ett rättsligt uttalande gjort bedömningen att de generellt tillämpliga bestämmelserna i förvaltningslagen får anses innefatta sådana lämpliga skyddsåtgärder för den registrerade att automatiserade beslut får fattas när de faller inom ramen för förvaltningslagens tillämpningsområde och är överklagbara.¹³⁸

Enligt eSams uppfattning utgör dataskyddsförordningen således inget hinder mot automatiserat beslutsfattande enligt förvaltningslagen.

¹³⁷ eSam (eSamverkansprogrammet) är ett medlemsdrivet program för samverkan mellan 23 myndigheter och SKL.

¹³⁸ eSams rättsliga uttalande, dnr/ref: VER 2017:57. (2018-03-19)
<http://www.esamverka.se/download/18.1d63d1d4162d1fa77f4264b/1523951705890/Ra%CC%88ttsligt%20uttalande%20art%2022%20autobeslut.pdf>

En kommun som utnyttjat denna möjlighet är Trelleborg, där RPA (*robotic process automation*) använts för automatiserade beslut inom socialtjänsten för biståndsansökningar.¹³⁹ Detta har samtidigt varit omdebatterat.¹⁴⁰

I ett projekt med stöd från Vinnova använder sig Norrtälje kommun av AI/ML som beslutsstöd vid orosanmälningar rörande barn och unga.¹⁴¹

Kommunallagen

När det emellertid gäller beslut som fattas enligt kommunallagen (2017:725) av kommunala myndigheter, har Sveriges Kommuner och Regioner (SKR) gjort bedömningen att det i dagsläget inte finns något utrymme för helt automatiserade beslut där ingen människa deltar i beslutet.

I enlighet med detta har SKR under våren 2019 riktat en hemställan till regeringen att ändra kommunallagen så att den inte hindrar automatiserat beslutsfattande.¹⁴² Det har resulterat i utredningen ”*En väl fungerande ordning för val och beslutsfattande i kommuner och regioner*” Dir. 2020:10. Utredaren ska bland annat (i) ta ställning till hur en möjlighet till automatiserat beslutsfattande kan införas i kommuner och regioner, (ii) särskilt överväga hur ett automatiserat beslutsfattande bör förenas med principerna för det kommunala beslutsfattandet, stadgan och dataskyddsförordningen, och (iii) föreslå de författningsändringar som bedöms lämpliga. Uppdraget ska redovisas till regeringen senast den 24 mars 2021.¹⁴³

Information om algoritmer

Vid automatiserade beslut inom offentlig förvaltning finns idag ingen skyldighet för myndigheter att redovisa algoritmen (eller metodiken) bakom ett förvaltningsbeslut. Emellertid har Digitaliseringsrättsutredningen i betänkandet *Juridik som stöd för förvaltningens digitalisering* (SOU 2018:25) föreslagit att en myndighet ska se till att information kan lämnas om hur myndigheten vid handläggning av mål eller ärenden använder algoritmer eller datorprogram som helt eller delvis påverkar utfallet eller beslutet vid automatiserade urval eller beslut.¹⁴⁴

9.4 Ansvarsfrågan

I tidigare avsnitt har beskrivits risker och olika fel som kan uppkomma vid användning av AI inom olika branscher och samhällssektorer. En viktig fråga i samtliga fall är – vem

¹³⁹ <https://moten.trelleborg.se/welcome-sv/namnder-styrelser/arbetsmarknadsnamnden/arbetsmarknadsnamnden-2018-09-17/agenda/automatiserat-beslutsfattande-i-kommunal-forvaltningpdf-1?downloadMode=open>

¹⁴⁰ <https://www.svd.se/obehorig-algoritm-tar-beslut-i-socialtjansten>

¹⁴¹ <https://www.norrtalje.se/info/stod-och-omsorg/familj-barn-och-unga/projekt-for-ai-och-robotisering-av-orosnmalan/>

¹⁴² SKR. Hemställan om ändring i kommunallagen med anledning av bestämmelse i 28 § förvaltningslagen om automatiserat beslutsfattande. (2019-05-23) <https://skr.se/download/18.3654812616b816209bf6889a/1561463149520/190625%20Skrivelse%20till%20Regeringen%20om%20automatiskt%20beslutsfattande.docx.pdf>

¹⁴³ <https://www.regeringen.se/4a4a88/contentassets/ff4ff714e67248aa8cafc2a49b4e8d54/en-val-fungerande-ordning-for-val-och-beslutsfattande-i-kommuner-och-regioner-dir.-202010.pdf>

¹⁴⁴ Digitaliseringsrättsutredningen SOU 2018:25 <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2018/03/sou-201825/>

bär ansvaret? Användaren av AI-tjänsten? Leverantören av tjänsten? AI-programmeraren? Eller andra aktörer som varit involverade i tjänsten?

Ansvarsfrågan kommer att bli allt viktigare, till exempel för försäkringsbolagen som har anledning att förvänta sig ett tydligt regelverk som grund för ersättningsprinciper. När avancerade robotar, autonoma maskiner och AI-system blir allt vanligare i samhället kommer de också oundvikligen att orsaka olyckor och även dödsfall.¹⁴⁵

Dessa frågor angående ansvar och vad tekniken får användas till, kan förväntas bli fler allt eftersom tekniken utvecklas och dess användningsområden breddas. AI-system kan antingen vara vad vi här kallar ”äkta AI”, det vill säga ett autonomt system med adaptiva algoritmer för ständigt lärande och ständig optimering, eller AI i form av fasta algoritmer som ligger till grund för en statisk regelstyrning. Det finns juridiska implikationer kopplade till skillnaden mellan dessa båda typer av AI, inte minst inom e-hälsa. Idag ses AI som mjukvara och regleras därmed av den medicintekniska förordningen, MDR. Enligt tolkning av regelverket, som beskrivits ovan, är ”äkta AI” inte tillåtet i klinisk användning inom hälso- och sjukvård, eftersom den är att se som en adaptiv, kontinuerligt och autonomt uppdaterande mjukvara. Enligt MDR måste en AI-lösning valideras och CE-märkas i statiskt format med lärandefunktionen avstängd inför klinisk användning.

Denna lagtolkning, som syftar till att värna patientsäkerheten, innebär också att fördelar med AI som ett ständigt lärande och optimerande verktyg kan komma att hämmas och därmed kan dess fulla potential inte nyttjas.

Avslutningsvis ska nämnas att ansvar sannolikt kommer utkrävas på flera nivåer. När det gäller hälso- och sjukvård och socialtjänst är vårdgivare och den som bedriver socialtjänst alltid ytterst ansvarig för den verksamhet som bedrivs och för den utrustning som man väljer att använda. Om det inträffar en skada vid användning av AI kan det också röra sig om ett handhavandefel som faller inom det personliga yrkesansvaret. Men om det är fel på utrustningen lär ansvar också kunna utkrävas av tillverkare eller leverantör. Sannolikt måste ansvarsfrågan utredas i varje enskilt fall.

9.5 Ny lagstiftning för AI-forskning?

Som nämnts ett flertal gånger ovan så behövs vid AI-utveckling stora mängder relevant data. I Finland stiftades 2019 en lag om sekundär användning av personuppgifter från social- och hälsovård. Denna lag syftar till att skapa en bättre grund för prediktiv analys, prognoser baserade på matematiska modeller, och för användning av AI. ”Dessa teknologier gör det möjligt att utnyttja stora datamassor bättre än tidigare till exempel i det professionella vårdarbetet och i beslutsfattandet.”¹⁴⁶ Lagens syfte är också att trygga skyddet för individens tillitsskydd samt rättigheter och friheter vid behandlingen av personuppgifter.

¹⁴⁵ Maria Gunther. Dagens Nyheter. Vem får skulden när en robot dödar? (2019-04-06) <https://www.dn.se/nyheter/vetenskap/vem-far-skulden-nar-en-robot-dodar/>

¹⁴⁶ Lag om sekundär användning av personuppgifter inom social- och hälsovården (552/2019). (2019-04-26) <https://finlex.fi/sv/laki/alkup/2019/20190552>

Den svenska Kommittén för teknologisk innovation och etik (Komet) har identifierat regelhinder som hämmar ansvarsfull utveckling och användning av nya tekniska lösningar inom hälsodataområdet. Därför har Komet nyligen föreslagit i en skrivelse till regeringen att en utredning skyndsamt bör tillsättas för att se över lagstiftningen som styr insamling och delning av hälsodata. Utredningen bör ha särskilt fokus på organisation och tillämpningsområde för sekundäranvändning av hälsodata. Sekundäranvändning av hälsodata innebär att använda personuppgifter som samlats in inom hälso- och sjukvård och socialtjänst för andra ändamål.

Angelägna frågor i utredningsarbetet rör:

- Ansvarsfördelning mellan staten, huvudmännen och andra berörda aktörer avseende hälsodata.
- Nationell reglering av sekundäranvändning av hälsodata.
- Ändamål för sekundäranvändning av hälsodata, till exempel utveckling och innovation, styrning och tillsyn av myndigheter, planering och utredning inom myndigheter samt undervisning.
- Organisering och tillämpning av sekundäranvändning av hälsodata, till exempel om någon myndighet bör få i uppdrag att på begäran om information samla in, samköra och lämna ut uppgifter som bygger på sekundäranvändning av hälsodata. Med detta avses att på begäran samla in personuppgifter från olika personuppgiftsansvariga och genom att samköra dem producera oidentifierad, anonymiserad eller pseudonymiserad information åt den som framställt begäran.
- Förutsättningar för en federerad lösning där hälsodata hanteras i flera distribuerade datakällor, i förhållande till att data i stället lagras centralt inom en myndighet eller motsvarande.
- Eventuella hinder, oklarheter eller motsättningar i gällande lagstiftning och, där det är relevant och möjligt, lämna förslag på eventuella författningsändringar.

Komet lyfter tydliga etiska aspekter, framför allt målkonflikten mellan den enskilda patientens integritet och möjligheten att i framtiden erbjuda bättre vård till följd av ny kunskap från forskning och innovation.¹⁴⁷

I syfte att förbereda för gränsöverskridande sekundäranvändning av hälsodata har EU-kommissionen utlyst en *Joint Action – Towards a European Health Data Space (TEHDAS)*. Finland har genom den statliga innovationsmyndigheten Sitra utsetts till koordinator för detta konsortium om drygt 20 länder. När detta skrivs befinner sig TEHDAS i ansökningsfasen. Sverige ingår i konsortiet, bland annat representerat av E-hälsomyndigheten i egenskap av *competent authority* (av regeringen utsedd expertmyndighet). E-hälsomyndigheten medverkar aktivt som ledare av ett work package inriktat på de olika ländernas tolkning av GDPR, tillståndsmyndigheter samt handläggning av ansökningar om att använda data för forskning, innovation eller framtagande av regelverk (*polymaking*). TEHDAS-programmet kan ses som en del av EU-kommissionens fortsatta arbete i syfte att ta fram ett regelverk för säkerhet, ansvar, grundläggande rättigheter och data-aspekter inom AI, och en dataförordning som

¹⁴⁷ https://www.kometinfo.se/wp-content/uploads/2020/10/Komet_pressmeddelande_20201015.pdf

fastställer villkor för kontroll och förutsättningar för att dela data, för medborgare och näringsliv.¹⁴⁸

När det gäller sekundäranvändning av potentiellt integritetskänslig hälsodata så kan etablerandet av regulatoriska sandlådor (*regulatory sandboxes*) bidra till att säkerställa en tillfällig skyddad miljö för detta. I sin nationella AI-strategi lyfter Finland fram ett generellt behov av regulatoriska sandlådor¹⁴⁹. Storbritannien (genom NHSx) och Singapore har beskrivit satsningar på regulatoriska sandlådor för AI med särskilt fokus på hälso- och sjukvård samt omsorg.^{150, 151, 152}

10. Kompetensförsörjning

Allt fler svenska företag såväl som myndigheter satsar på AI.¹⁵³ I en artikel om detta redogjordes för de hetaste jobbtitlarna, här i försök att översätta dessa till svenska: Maskininlärningsingenjör (Machine Learning Engineer), Djupinlärningsingenjör (Deep Learning Engineer), Senior datavetenskapare (Senior Data Scientist), Visuell dataingenjör (Computer Vision Engineer), Datavetenskapare (Data Scientist), Algoritmutvecklare (Algorithm Developer), Junior datavetenskapare (Junior Data Scientist), Utvecklarkonsult (Developer Consultant), AI-chef (Director of Data Science), AI-ledare och dataanalytiker (Lead Data Scientist).¹⁵⁴

Denna lista stärker bilden av att AI inte enbart kommer ta över ett flertal jobb i samhället, utan även kommer att bidra till att skapa nya.

Det finns tydliga tecken på att bristen på dessa kompetenser kommer att öka både i Sverige och internationellt framöver. Arbetsförmedlingen bedömer att efterfrågan på mjukvaru- och systemutvecklare kommer vara fortsatt stor de kommande åren. Även om många utbildas inom området väntas det inte vara tillräckligt för att möta efterfrågan på fem års sikt.¹⁵⁵ En kompetensbrist av detta slag riskerar att hämma utveckling och implementering av AI.

¹⁴⁸ Commission Work Programme 2021 A Union of vitality in a world of fragility. (2020-10-19) https://ec.europa.eu/info/sites/info/files/2021_commission_work_programme_en.pdf

¹⁴⁹ Publications of the Ministry of Economic Affairs and Employment Ministry 47/2017 Finland's Age of Artificial Intelligence, Turning Finland into a leading country in the application of artificial intelligence Objective and recommendations for measures, s. 44:

https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf

¹⁵⁰ Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care. NHSx, (oktober 2019) https://www.nhs.uk/media/documents/NHSX_AI_report.pdf

¹⁵¹ Regulatory sandbox for AI needed to test and build systems, NHSX says. Digitalhealth. (2020-02-12) <https://www.digitalhealth.net/2020/02/regulatory-sandbox-for-ai-needed-to-test-and-build-systems-nhsx-says/>

¹⁵² Licensing Experimentation and Adaptation Programme (LEAP) - A MOH Regulatory Sandbox, Ministry of Health, Singapore. (september 2018)

<https://www.moh.gov.sg/home/our-healthcare-system/licensing-experimentation-and-adaptation-programme-leap---a-moh-regulatory-sandbox>

¹⁵³ Martin Wallström. SvD. Nästan alla företag satsar på AI-projekt. (2019-07-21)

<https://www.svd.se/kraftig-okning-av-ai-projekt-nastan-alla-foretag-satsar>

¹⁵⁴ M. Wallström. SvD. Sug efter AI-expert – här är nya heta jobben. (2019-07-21)

<https://www.svd.se/sug-efter-ai-expert--har-ar-hetaste-jobbtitlarna>

¹⁵⁵ Arbetsförmedlingen. Mjukvaru- och systemutvecklare. <https://arbetsformedlingen.se/for-arbetssookande/valj-yrke/framtidsutsikter/yrkeskompassen/#/yrkesprognos/2512>

Som framgår av denna rapport är det också av stor betydelse att de programmerare, datavetenskapare och AI-chefer som ska forma en långsiktigt hållbar AI inom exempelvis e-hälsa, har adekvat utbildning och kunskap även inom etik, juridik och mänskliga rättigheter.

Enligt rapporten *Discriminating systems, Gender, Race, and Power in AI*¹⁵⁶ är det 15 procent av AI-forskarna på Facebook och 10 procent på Google som är kvinnor.¹⁵⁷ Enligt rapporten är situationen liknande inom akademien. Nyligen genomförda studier visar att 18 procent av författarna på ledande AI-konferenser är kvinnor¹⁵⁸, samtidigt som mer än 80 procent av AI-professorerna är män¹⁵⁹.

Detta understryker vikten av att mångfalden stärks på området, för att minska risken för bias/diskriminering i framtidens AI-lösningar, till exempel inom e-hälsa.

För framgångsrika AI-satsningar inom e-hälsa så är det naturligtvis av betydelse att även personal inom hälso- och sjukvård samt omsorg får utbildning inom detta område. Därutöver finns sannolikt ett allmänt behov av ökad kunskap om AI och dess betydelse för vårt framtida samhälle. En omfattande satsning på att tillgodose detta behov är den online-kurs som Finland erbjuder sin befolkning – och andra. Under sin tid som ordförandeland i EU uttalade Finland ambitionen att utbilda 1 procent av EU:s befolkning, och hittills har drygt en halv miljon människor över hela världen klickat sig igenom kursen.¹⁶⁰

11. Frågor för fortsatt diskussion

E-hälsomyndighetens fokusrapporter, varav föreliggande är en, ska beskriva och diskutera ett aktuellt område som har betydelse för e-hälsoutvecklingen i Sverige, och därigenom bidra till uppfyllandet av Vision e-hälsa 2025. Denna rapport är ett diskussionsunderlag som belyser i första hand övergripande etiska och juridiska frågor som kan ses som avgörande för en god AI-utveckling, i synnerhet de som uppstår inom hälso- och sjukvård och omsorg. Även andra viktiga aspekter har berörts, såsom kompetensförsörjning inom området.

Rapporten har visat på AI:s potential men även på de risker som föreligger. AI kommer med stor sannolikhet att förändra vårt samhälle och bidra till nytta och stora vinster inom vitt skilda områden, bland annat inom hälso- och sjukvård samt omsorg. Några frågor är emellertid viktiga att fortsätta belysa och diskutera:

¹⁵⁶ Sarah Myers West et al., DISCRIMINATING SYSTEMS - Gender, Race, and Power in AI. (april 2019) <https://ainowinstitute.org/discriminating-systems.pdf>

¹⁵⁷ T. Simonite. WIRED. AI is the future - but where are the women?. (2018) <https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/>

¹⁵⁸ Element AI. Global AI Talent Report. (2019) <https://ifgagne.ai/talent-2019/>

¹⁵⁹ Artificial Intelligence Index 2018. <http://cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf>

¹⁶⁰ Welcome to the Elements of AI free online course! <https://www.elementsofai.com/>

- AI förutsätter stora mängder data som bör vara strukturerad¹⁶¹, adekvat och av god kvalitet. Data ska vara representativt och utan bias samtidigt som hänsyn tas till den personliga integriteten. Hur skapar vi rätt avvägning mellan samhällets och individens bästa – och behöver dessa stå i motsats till varandra?
- Idag ses AI som mjukvara och regleras därmed av den medicintekniska förordningen, MDR. ”Äkta AI” är enligt regelverket inte tillåtet i klinisk användning inom hälso- och sjukvård, eftersom den är att se som en adaptiv, kontinuerligt och autonomt uppdaterande mjukvara. Medför regelverket att AI:s fulla potential inom hälso- och sjukvården som ett ständigt lärande och optimerande verktyg hämmas?
- AI-liknande lösningar kan användas för automatiserade beslut, t ex inom socialtjänsten. Dataskyddsförordningen GDPR innehåller ett förbud mot detta, vilket emellertid kan upphävas genom nationell lag vilket anses vara fallet genom förvaltningslagen. Samtidigt förbjuder dock kommunallagen automatiserade beslut. Utgör detta ett hinder för en önskad AI-utveckling? Och om så är fallet, hur kan hindret i så fall undanröjas på ett säkert sätt? (frågan utreds enligt Dir 2020:10¹⁶²)
- Arbetsmarknadsprognoserna förutspår en brist på kompetens för framtida utveckling av AI och hantering av Big Data. Hur säkerställs den framtida kompetensförsörjningen?
- Hur säkerställer vi att alla människor får ta del av den optimering av välfärden som AI kan möjliggöra? Det är grundläggande utifrån hälso- och sjukvårdslagens portalparagraf rörande vård på lika villkor för hela befolkningen. Grundläggande för en långsiktigt framgångsrik och hållbar utveckling och tillämpning av AI inom e-hälsa är också att medborgare såväl som vård- och omsorgspersonal har förtroende för, och tillit till dessa hjälpmedel. En avslutande viktig fråga är då: Hur informerar och utbildar vi alla berörda i samhället – medborgare, patienter och brukare, professionerna, beslutsfattare? Detta för att nödvändig tillit ska kunna etableras och bibehållas.

¹⁶¹ Strukturerad data exempelvis att grunddata dels på organisations nivå men även kliniska grunddata kopplat till standard som exempelvis SNOMED CT

¹⁶² En väl fungerande ordning för val och beslutsfattande i kommuner och regioner (2020 feb)
<https://www.regeringen.se/rattsliga-dokument/kommittedirektiv/2020/02/dir.-202010/>