SVENSKT NÄRINGSLIV

Vem tar ansvar för AI?

GÄLLANDE LAGSTIFTNING, FRAMLAGDA REFORMFÖRSLAG OCH ÖVERGRIPANDE ANALYS
MARS 2021

Författare: Daniel Westman

Innehåll

Förord

Grunden för all handel är kundrelationer. Bra kundrelationer byggs med tillit. Även när allt fler produkter och tjänster innehåller eller består av en AI-lösning ska kunder kunna vara trygga med att höga krav på säkerhet, etik och ansvar gäller. Det är viktigt att skapa incitament för tekniska lösningar som ökar säkerheten och därmed minskar riskerna för skador och kränkningar.

Dagens politiska inspel handlar dels om risk och rädslor, dels om vikten av att se det goda för medborgare, miljö och konkurrenskraft som kan komma med AI-användning. För att gå vidare med nya förslag behöver först existerande reglering som ger tillit identifieras, men också den som begränsar utvecklingen. Att regler är proportionerliga, förutsägbara och rimligt enkla att tillämpa är viktiga utgångspunkter, men lagstiftaren behöver också fokusera på hur vi kan skapa de bästa AI-tillämpningarna av högsta kvalitet.

Kanske leder nya branschövergripande AI-regleringar till både över- och underreglering? Kanske behöver AI regleras på många olika sätt beroende på typ av använd AI-teknik och bransch? Kanske är den redan reglerad?

Svenskt Näringsliv har bett juristen Daniel Westman, som skrivit om och arbetat praktiskt med it-rätt och dataskyddsfrågor i över 20 år, att beskriva gällande säkerhets- och ansvarslagstiftning, nyligen presenterade förslag samt föreslå åtgärder för att förbättra regelverket. Bedömningarna och slutsatserna är författarens egna.

Med denna rapport hoppas Svenskt Näringsliv kunna bidra till kunskap om gällande lagstiftning och hittills framlagda förslag, men framför allt presentera en analys av vad som kan tydliggöra ansvaret för AI i framtidens reglering.

Stockholm mars 2021

Carolina Brånby

1 Inledning

1.1 Rapportens innehåll och uppläggning

Artificiell intelligens (AI) spås förändra många delar av samhället i grunden. Redan idag används AI för att ställa medicinska diagnoser, minska energianvändningen, ta fram nya produkter och tjänster, effektivisera industriell produktion och förbättra kundservicen.

AI kan användas för att förbättra säkerheten och till att säkerställa skyddet för grundläggande rättigheter. Det kan handla om att förhindra olyckor i trafiken eller om att upptäcka och motverka it-baserade angrepp. Rätt använda kan AI-lösningar öka transparensen i samhället och förhindra kränkningar av privatlivet i onlinemiljöer.

Samtidigt är AI-tekniken i sig förenad med risker, där vissa är betydande. Om AI-system som används i självkörande fordon eller sjukvården fallerar kan det leda till sak- eller personskador. Om beslut fattas av bristfälliga AI-system kan individer utsättas för diskriminering eller felaktig hantering av personuppgifter.

Denna rapport handlar om den rättsliga reglering som ska möta denna typ av risker. Fokus riktas mot regler som syftar till att förhindra och ersätta sak- och personskador, men regler som syftar till att motverka kränkningar av mänskliga rättigheter (diskriminering, skydd för personuppgifter etc.) berörs också kortfattat.

Även om AI baserad på maskininlärning är en relativt ny företeelse saknas det inte tillämpliga rättsregler. Äldre regler har i huvudsak en teknikneutral utformning och är därför normalt tillämpliga också när AI används. Det gäller till exempel generella och sektorsspecifika produktsäkerhetsregler i EU-rätten, allmänna skadeståndsregler på nationell nivå och produktansvarsregler inom EU-rätten. Därutöver är regler som syftar till att skydda grundläggande rättigheter, till exempel EU-regler om diskriminering och behandling av personuppgifter, relevanta vid utveckling och användning av AI.

Egenskaper hos AI-tekniken gör dock att befintliga regler ibland blir svåra att tillämpa och att det uppstår oönskade luckor i skyddet. Det kan leda till att otillräckliga säkerhetskrav ställs på nya produkter och till att den som har orsakats en skada har praktiska svårigheter att få ersättning. I förlängningen kan detta leda till att medborgare, konsumenter och företag förlorar tilliten till produkter och tjänster som använder sig av AI. För företag som utvecklar eller använder AI-lösningar kan brister i lagstiftningen leda till onödiga rättsliga hinder, rättsosäkerhet och snedvriden konkurrens i förhållande till aktörer i andra länder.

Mot denna bakgrund har det inom EU inletts ett arbete med att utvärdera den befintliga regleringen och med att utarbeta uppdaterade och kompletterande regler. EU-kommissionen har presenterat en så kallad vitbok om AI,[1] med en tillhörande rapport om säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik.[2] Europaparlamentet har nyligen antagit två resolutioner som innehåller fullständiga lagstiftningsförslag. Det ena förslaget innehåller ett ramverk för etiska aspekter av artificiell intelligens, robotteknik och tillhörande teknik.[3] Det andra förslaget innehåller en ny skadeståndsordning för artificiell intelligens.[4] Kommissionen – som har exklusiv behörighet att inleda lagstiftningsärenden inom EU – har aviserat att den ska återkomma med sina förslag under första kvartalet 2021.

I denna rapport beskrivs hur gällande regelverk för framför allt produktsäkerhet och skadeståndsansvar kan tillämpas på AI-relaterade produkter eller tjänster. Därefter beskrivs de reformförslag som hittills presenterats på EU-nivå översiktligt. Avslutningsvis lämnas vissa rekommendationer om förhållningssättet till ny lagstiftning på området. Hur kan intresset av trygghet och tillit för AI balanseras mot intresset av att inte i onödan försvåra utvecklingen och användningen av nya innovativa lösningar?

Rapporten behandlar ett omfattande och komplext område, där utvecklingstakten dessutom är hög. Samtidigt är ambitionen att framställningen ska vara kortfattad och lättillgänglig. En naturlig konsekvens blir därmed att framställningen blir översiktlig.

1.2 Vad är AI? En policyanpassad definition

Begreppet artificiell intelligens (AI) har inte någon allmänt accepterad definition, utan används med olika innebörd i olika sammanhang. AI kan uppfattas som ett samlingsbegrepp för olika tekniker och forskningsområden som har det gemensamt att de är inriktade på att få datorprogram att utföra arbetsuppgifter som traditionellt har ansetts kräva mänsklig intelligens. Det kan till exempel handla om röstassistenter, bildanalysprogram, självkörande fordon och tal- och ansiktsigenkänningssystem.

Vid en policydiskussion är det viktigt att arbeta med en definition av AI som tar sikte på sådana aspekter som motiverar särskilda politiska och rättsliga överväganden. Inom EU:s policyarbete har det ansetts motiverat att arbeta med en definition som tar sikte på system som i någon mån är *autonoma* och innehåller ett moment av *självlärande*.

[1] Kommissionen, *Vitbok om artificiell intelligens – en EU-strategi för spetskompetens och förtroende* (COM(2020) 65 final), nedan ”EU-kommissionens vitbok”.

[2] Rapport från kommissionen till Europaparlamentet, rådet och Europeiska Ekonomiska och sociala kommittén, *Konsekvenser för säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik* (COM(2020) 64 final).

[3] Europaparlamentets resolution av den 20 oktober 2020 med rekommendationer till kommissionen om en ram för etiska aspekter av artificiell intelligens, robotteknik och tillhörande teknik (2020/2012(INL)), nedan ”Europaparlamentets etikresolution”.

[4] Europaparlamentets resolution av den 20 oktober 2020 med rekommendationer till kommissionen om en skadeståndsordning för artificiell intelligens (2020/2014(INL)), nedan ”Europaparlamentets skadeståndsresolution”.

”Artificiell intelligens avser system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder – med viss grad av självständighet – för att uppnå särskilda mål.

AI-baserade system kan vara helt programvarubaserade och fungera i den virtuella världen (till exempel röstassistenter, bildanalysprogram, sökmotorer, tal- och ansiktsigenkänningssystem), eller inbäddas i hårdvaruenheter (till exempel avancerade robotar, självkörande bilar, drönare eller tillämpningar för sakernas internet).”[5]

Genom denna definition riktas i praktiken stort fokus på lösningar som bygger på maskininlärning, det vill säga upplärning av *algoritmer* genom användning av *stora datamängder* (big data). Definitionen tydliggör också att AI har en nära koppling till *sakernas internet* (internet of things), det vill säga uppkopplad utrustning. Självkörande fordon är exempel på autonoma och självlärande system som bygger på hantering av stora datamängder, bland annat inhämtade genom uppkopplade sensorer av olika slag.

Begreppet AI för lätt tankarna till robotar eller system med människolik förmåga till lärande och analysförmåga (generell eller stark AI). Under en överskådlig framtid kommer AI dock vara synonymt med kvalificerad automatisering av vissa specifika arbetsuppgifter (specifik eller svag AI).

1.3 Stora möjligheter – men också risker

AI har potential att bidra med betydande nytta inom en mängd områden. Analyser pekar på att användningen av AI kan öka den ekonomiska tillväxten. AI används redan idag för att ställa medicinska diagnoser, minska energianvändningen, reducera trafikolyckor, skapa nya tjänster, effektivisera industriell produktion, utveckla nya läkemedel och förkorta handläggningstider.[6]

AI-teknologin är samtidigt förknippad med risker av olika slag och allvarlighetsgrad. I princip kan AI-system orsaka alla typer av skador som förekommer inom skadeståndsrätten, det vill säga personskador, sakskador, rena förmögenhetsskador[7] och kränkningar (till exempel av den personliga integriteten).

Dessa typer av skador kan orsakas *direkt* av ett AI-system, till exempel genom att systemet manövrerar ett fordon eller fattar ett beslut om hur en viss person ska behandlas. Men AI-systemens inverkan kan också vara mer *indirekt*; till exempel

[5] Kommissionens meddelande till Europaparlamentet, rådet, Europeiska ekonomiska och sociala kommittén och Regionkommittén, *Artificiell intelligens för Europa* (COM(2018) 237 final). Definitionen diskuteras och utvecklas av Kommissionens expertgrupp på hög nivå för AI-frågor, *En definition av AI: Viktigaste förmågor och discipliner*, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines.

[6] Regeringskansliet, *Nationell inriktning för artificiell intelligens*, Stockholm 2018, https://www.regeringen.se/informationsmaterial/2018/05/nationell-inriktning-for-artificiell-intelligens, s. 4.

[7] Ekonomisk skada som uppkommer utan samband med att någon lider person- eller sakskada (1 kap. 2 § skadeståndslagen (1972:207)). Se vidare avsnitt 2.

kan en skada uppstå i samband med att en mänsklig beslutsfattare ges ett felaktigt eller vilseledande underlag.

I takt med att allt fler verksamheter i samhället digitaliseras och den direkta kontrollen lämnas över till AI-system ökar naturligtvis risken för att brister i dessa system leder till allvarliga person- och sakskador. I ett konkret fall kan orsaken till en inträffad skada vara svår att fastställa. Möjliga förklaringar kan till exempel vara bristande utformning av ett AI-systems grundläggande algoritmer, användning av felaktiga eller ofullständiga data för maskininlärningen, handhavandefel eller externa cyberangrepp mot systemet.

I policydiskussionen kring AI har även riskerna för *kränkningar av mänskliga rättigheter* uppmärksammats. Särskilt riskerna för diskriminering samt kränkning av rätten till privatliv och rätten till skydd för personuppgifter har diskuterats. Om felaktiga eller ofullständiga datamängder används för maskininlärning kan det leda till diskriminering eller annan typ av oönskad särbehandling, till exempel på grund av kön eller etnicitet. Samtidigt kan en alltför omfattande behandling av personuppgifter eller en behandling av känsliga personuppgifter vara problematisk i förhållande till skyddet av personuppgifter, särskilt om personuppgifterna läggs till grund för profilering och automatiserade beslut. AI-systems bristande transparens liksom avsaknaden av möjligheter till omprövning genom mänsklig inblandning kan ibland leda till försämrad rättssäkerhet.

På en samhällsnivå finns det till och med *risker för demokratin*. Vissa särskilda typer av AI-system, till exempel system för massövervakning och autonoma vapensystem, utgör effektiva verktyg för totalitära stater. Men även i andra stater finns betydande risker. Det kan till exempel handla om att AI-system i allt högre grad kontrollerar det offentliga samtalet som sker i sociala medier genom att moderera yttranden som användarna gör eller om att insyn, kontroll och ansvarsutkrävande blir svårare inom en mängd samhällsområden.

Ovanstående inventering av skadetyper och risker förknippade med AI är central för att det ska vara möjligt att analysera behovet av åtgärder, till exempel behovet av rättsliga reformer. Samtidigt kan den ge en alltför mörk bild av AI-utvecklingen. Det är viktigt att komma ihåg att människor inte heller är ofelbara och att skador regelbundet uppkommer i verksamheter där arbetet utförs eller styrs av människor. Det bör även noteras att ett viktigt användningsområde för AI är just att upptäcka säkerhetsbrister och att varna för onormala eller riskfyllda aktiviteter.

1.4 AI-policy – strategi, etik och juridik

I mitten av år 2020 hade över 60 länder antagit nationella strategier för AI och flera hade inlett ett sådant arbete.[8] Gemensamma inslag i dessa strategier är åtgärder för att främja forskning, tillgång till data (för maskininlärning) samt kompetensutveck-

[8] OECD (2020), *OECD Digital Economy Outlook 2020*, OECD Publishing, Paris, *https://doi.org/10.1787/bb167041-en*, s. 272.

ling. Många länder strävar efter att användningen av AI ska vinna folks förtroende samtidigt som de försöker hitta metoder för att kontrollera riskerna förknippade med AI. Den svenska regeringen presenterade en AI-strategi med denna inriktning under 2018.[9] Strategin innehåller inte någon djupare analys av de rättsliga frågeställningarna kring säkerhet och ansvar. Bilden är att den rättsliga regleringen kring AI i huvudsak anses vara en fråga som ska hanteras på EU-nivå.[10]

EU-kommissionens vitbok om AI som presenterades i februari 2020 utgör en strategi för spetskompetens och tillit.[11] Utgångspunkten för EU:s policy på AI-området är att människan ska sättas i centrum. AI uppfattas inte som ett självändamål utan som ett sätt att förbättra för mänskligheten. AI ska utvecklas och användas med respekt för rättsliga och etiska krav. Det är människor som ska hållas ansvariga när AI utvecklas och används. Även om AI-system kan agera mer eller mindre autonomt är det därmed inte aktuellt att införa rättigheter och skyldigheter för systemen som sådana.

EU-kommissionens ställningstagande i vitboken ligger i linje med de *etiska riktlinjer för tillförlitlig AI* som kommissionens oberoende expertgrupp på hög nivå för AI-frågor presenterade i april 2019.[12]

”Tillförlitlig AI har tre komponenter som bör finnas med under systemets hela livscykel: a) den bör vara laglig och följa alla gällande lagar och förordningar, b) den bör vara etisk och säkerställa att etiska principer och värden upprätthålls, och c) den bör vara robust ur både teknisk och samhällelig synvinkel, eftersom AI-system kan orsaka oavsiktliga skador, trots goda intentioner. Varje komponent är nödvändig, men inte tillräcklig i sig för att skapa tillförlitlig AI. Helst ska alla tre komponenter fungera harmoniskt tillsammans och överlappa varandra. Om det i praktiken uppstår spänningar mellan dessa komponenter bör samhället sträva efter att lösa dem.”[13]

Expertgruppen behandlar inte frågan om vad som utgör laglig AI närmare. Däremot presenteras ett antal *etiska principer* för att skapa tillförlitlig AI, till exempel principerna om respekt för människans autonomi, skadeförebyggande, rättvisa och förklarbarhet. Gruppen anger att utveckling, spridning och användning av AI-system bör uppfylla kraven för tillförlitlig AI: 1) mänskligt agentskap och mänsklig tillsyn, 2) teknisk robusthet och säkerhet, 3) integritet och dataförvaltning, 4) transparens, 5) mångfald, icke-diskriminering och rättvisa, 6) samhällets och miljöns välbefinn

[9] Regeringskansliet, *Nationell inriktning för artificiell intelligens*, Stockholm 2018, https://www.regeringen.se/informationsmaterial/2018/05/nationell-inriktning-for-artificiell-intelligens.

[10] I ett så kallat non-paper har Sverige tillsammans med 13 andra medlemsstater i oktober 2020 lämnat vissa synpunkter på den framtida regleringen av AI (*Non-paper – Innovative and trustworthy AI: two sides of the same coin*, Position paper on behalf of Denmark, Belgium, the Czech Republic, Finland, France, Estonia, Ireland, Latvia, Luxembourg, the Netherlands, Poland, Portugal, Spain and Sweden on innovative and trustworthy AI, tillgängligt på https://www.permanentrepresentations.nl/documents/publications/2020/10/8/non-paper---innovative-and-trustworthy-ai). Se vidare avsnitt 3.5 nedan.

[11] Kommissionen, *Vitbok om artificiell intelligens – en EU-strategi för spetskompetens och förtroende* (COM (2020) 65 final), nedan ”EU-kommissionens vitbok”.

[12] Kommissionens expertgrupp på hög nivå för AI-frågor, *Etiska riktlinjer för tillförlitlig AI*, Bryssel 2019.

[13] A.a. s. 2.

ande samt 7) ansvarsskyldighet.[14] De etiska riktlinjerna innehåller vägledning kring ett AI-systems hela livscykel. Expertgruppen har även tagit fram en bedömningslista som kan användas av företag som utvecklar eller använder AI.[15]

De etiska riktlinjerna för tillförlitlig AI innehåller åtgärder som kan användas för att motverka de risker som beskrivits i föregående avsnitt. Men som expertgruppen själv framhåller ersätter de icke-bindande etiska principerna inte rättslig reglering av AI. I vitboken framhåller EU-kommissionen att tydliga regler kring användningen av AI skulle skapa förtroende hos konsumenterna och företagen och därmed påskynda användandet av tekniken. Det framhålls särskilt att regelverket bör främja Europas innovationskapacitet och konkurrenskraft.[16]

När det mer konkret gäller *säkerhets- och ansvarsregleringen* konstaterar EU-kommissionen att bristen på tydliga säkerhetsbestämmelser för AI-system inte bara riskerar att utsätta personer för risker, utan även skapar rättslig osäkerhet för företag som marknadsför produkter som innefattar AI. Vidare framhålls att de myndigheter som ansvarar för marknadsövervakning och kontroll av reglernas efterlevnad kan hamna i en situation där de är osäkra på om de kan ingripa. Orsaken kan vara osäkerhet kring den egna befogenheten och brist på teknisk kapacitet att inspektera systemen. Rättslig osäkerhet anses därför minska säkerheten totalt sett och undergräva de europeiska företagens konkurrenskraft.[17]

Om skador skulle uppkomma gör bristen på tydliga krav och egenskaperna hos AI-tekniken enligt EU-kommissionen det svårt att spåra potentiellt problematiska beslut som fattas med stöd av AI-system. Detta kan i sin tur göra det svårt för personer som lidit skada att få ersättning enligt befintlig ansvarslagstiftning på nationell nivå och EU-nivå.[18]

I vitboken och i den samtidigt publicerade rapporten om säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik analyseras tillämpningen av den gällande regleringen översiktligt.[19] Eftersom vissa medlemsstater redan överväger lagstiftning för att möta riskerna med AI menar EU-kommissionen att det ur ett inremarknadsperspektiv är angeläget med en översyn av den relevanta EU-lagstiftningen.[20]

Utöver eventuella anpassningar av det befintliga regelverket menar EU-kommissionen att det kan behövas *en ny särskild lagstiftning om AI*. För att den nya lagstiftningen

[14] A.a. s. 2 f. Det handlar som synes knappast om några helt nya etiska principer, utan om en tillämpning av kända principer just på AI.

[15] En uppdaterad version av bedömningslistan samt vissa kompletterande verktyg för praktiskt etiskt arbete med AI publicerades under 2020, se *The assessment list for trustworthy artificial intellegens* (ALTAI) for self assessment.

[16] EU-kommissionens vitbok, s. 11.

[17] EU-kommissionens vitbok, s. 13.

[18] EU-kommissionens vitbok, s. 14.

[19] Rapport från kommissionen till Europaparlamentet, rådet och Europeiska Ekonomiska och sociala kommittén, *Konsekvenser för säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik* (COM (2020) 64 final). Se analysen i avsnitt 2 nedan.

[20] EU-kommissionens vitbok, s. 17.

ska vara proportionerlig och inte i onödan begränsa användningen av AI bör den inriktas mot användningar som innebär hög risk. I vitboken skisserar kommissionen på kriterier för att identifiera högriskverksamhet samt möjliga krav som skulle kunna ställas på sådan verksamhet.[21]

1.5 AI är inte oreglerat

Det saknas idag i princip helt regler som är särskilt inriktade på AI. Befintliga regler inom en mängd olika rättsområden är emellertid tillämpliga även på AI. Såväl sektorsspecifika regler (till exempel regler rörande fordon och medicinska produkter), som generella regler (till exempel regler om skadeståndsansvar och behandling av personuppgifter) är styrande för utformningen och användningen av AI. Vissa av dessa regler är resultatet av EU-rättslig harmonisering, medan andra har ett rent nationellt ursprung.

I avsnitt 2 beskrivs och analyseras säkerhets- och ansvarsregleringens tillämpning på AI något mer utförligt. Dessutom behandlas kortfattat vissa kompletterande regler som syftar till att ge ett skydd mot kränkningar av mänskliga rättigheter (till exempel diskrimineringslagstiftningen och dataskyddslagstiftningen).

Vid sidan av de regler som behandlas i denna rapport kan en mängd andra regler bli relevanta för utformningen av användningen av AI. Tillgången till och möjligheterna att använda data för maskininlärning bestäms till exempel av ett antal olika regelverk (regler om tillgång till och rätten att vidareutnyttja myndighetsinformation, skyddet för företagshemligheter, immaterialrätt och dataskyddslagstiftningen). När en färdig AI-applikation används i en viss typ av verksamhet blir eventuell särskild reglering som finns för denna verksamhet tillämplig, till exempel regler för den finansiella sektorn eller hälso- och sjukvården.

I praktiken kan reglering som är relevant för AI innebära krav på systemens utformning eller på utvecklingsprocessen. Vanligast torde dock vara att de generellt tillämpliga regler som finns idag är inriktade på *effekterna* av att ett AI-system används.

Utvecklingen och användningen av AI befinner sig med andra ord inte i ett rättsligt vakuum – tvärtom. Det faktum att det finns regler som är tillämpliga också på AI är emellertid inte detsamma som att reglerna är adekvata. Särskilda egenskaper hos AI-tekniken kan till exempel göra reglerna svåra att tillämpa och leda till en bristande uppfyllelse av reglernas ändamål, till exempel att skydda medborgarna mot skador och kränkningar. Men gällande regler kan också innebära omotiverade begränsningar av möjligheterna att utveckla och använda AI. Det kan till exempel handla om begränsningar av möjligheterna att få tillgång till och använda relevanta data för maskininlärning. Ytterligare ett problem kan vara att bristande harmonisering motverkar en fungerande inre marknad för AI-tillämpningar.

[21] EU-kommissionens vitbok, s. 18 ff. Kommissionen regleringsskiss beskrivs i avsnitt 3.2 nedan.

1.6 Egenskaper som leder till rättsliga utmaningar

AI-systemens *autonoma och självlärande funktion* kan göra det svårt att avgöra vem som ska hållas ansvarig för ett systems åtgärder. Situationen kompliceras av att *många aktörer* kan vara inblandade under ett AI-systems livscykel. Exempelvis kan ett företag utveckla systemets grundläggande funktionssätt, ett annat företag tillhandahålla data för maskininlärningen och ett tredje företag ta systemet i bruk i sin verksamhet.

AI-system kopplas ofta samman med andra uppkopplade enheter och tar emot impulser från dessa. Systemen som sådana är inte heller några statiska produkter, utan tvärtom föremål för förändring, till exempel i form av programvaruuppdateringar. Denna *öppenhet och föränderlighet* innebär särskilda utmaningar vid tillämpning och utformning av säkerhets- och ansvarsregler.

AI-system som är baserade på maskininlärning är i huvudsak uppbyggda kring prediktiv analys, det vill säga användning av stora datamängder för att hitta mönster som sedan läggs till grund för algoritmernas utformning och funktionssätt. Detta arbetssätt, i kombination med systemens tekniska komplexitet, leder inte sällan till en *black-box-problematik*, det vill säga att det blir svårt att få insyn i och förklaring till ett systems agerande.

2 Gällande säkerhets- och ansvarsreglering

2.1 Inledande översikt

Produktsäkerhetsregleringen innehåller säkerhetskrav som måste vara uppfyllda för att produkter och tjänster ska få göras tillgängliga på marknaden. Den generella produktsäkerhetsregleringen kompletterar och fyller ut den mer detaljerade och krävande reglering som finns för vissa typer av produkter, till exempel medicintekniska produkter och bilar. Eftersom säkerhetskrav på varor och tjänster får stor betydelse för den inre marknaden har det skett en omfattande EU-harmonisering av produktsäkerhetsregleringen.

Även regleringen rörande ansvar för skador består av flera delar. I nationell skadeståndsrätt finns ett långtgående skadeståndsansvar för den som genom oaktsamhet orsakar en skada. För vissa aktörer och vissa verksamheter är skadeståndsansvaret strikt. På EU-nivå finns en harmoniserad lagstiftning om så kallat produktansvar som innebär att tillverkaren åläggs ett strikt ansvar för personskador som en produkt har orsakat på grund av en säkerhetsbrist.

Produktsäkerhetsregleringen och ansvarsregleringen kompletterar varandra. Produktsäkerhetsregleringen anger under vilka förutsättningar en produkt får släppas ut på marknaden. Ansvarsregleringen bestämmer vem som får ta konsekvensen av att en skada faktiskt inträffar.

När det gäller kränkningar av grundläggande rättigheter som kan uppstå vid användningen av AI kompletteras produktsäkerhets- och ansvarsreglerna av vissa andra regelverk. Till denna grupp räknas till exempel regler om diskriminering samt regler om behandling av personuppgifter.

2.2 Produktsäkerhetsregleringen

Det EU-rättsliga regelverket på produktsäkerhetsområdet består dels av direktivet om allmän produktsäkerhet[22], dels av olika sektorsspecifika förordningar, till exempel för fordon och medicinteknisk utrustning.[23]

Produktsäkerhetslagen (2004:451) syftar till att säkerställa att varor och tjänster som tillhandahålls konsumenter inte orsakar personskador (1 §).[24] Lagen kompletterar och fyller ut den sektorsspecifika lagstiftningen på produktsäkerhetsområdet. Därmed garanteras att det finns grundläggande krav på produktsäkerhet för alla varor och tjänster som är avsedda för konsumenter eller kan antas komma att användas av konsumenter (2 §). Lagens krav riktas primärt mot tillverkare, men även distributörer åläggs skyldigheter enligt lagen.

Produktsäkerhetslagens grundkrav är att varor och tjänster som tillhandahålls av näringsidkare till konsumenter ska vara säkra (7 §). En vara eller tjänst anses säker om den inte för med sig någon risk för människors hälsa och säkerhet eller bara en låg risk. Tillhandahållande av varor eller tjänster med låg risk är dock bara tillåtet om risken anses vara godtagbar med beaktande av hur varan eller tjänsten används. Risken ska dessutom vara förenlig med en hög skyddsnivå när det gäller människors hälsa och säkerhet (8 §). I lagen anges vissa faktorer som ska beaktas vid prövningen av om dessa förutsättningar är uppfyllda (9–10 §§).

Ett centralt moment i den EU-baserade produktsäkerhetsregleringen är presumtionen för att en vara eller tjänst är säker om den uppfyller en tillämplig nationell säkerhetsstandard som genomför en EU-gemensam säkerhetsstandard (11 §). Eftersom en presumtion innebär stora praktiska fördelar för en näringsidkare skapas starka incitament att tillämpa dessa standarder. Det gäller inte minst för företag som är verksamma i flera medlemsstater och som därmed kan förhålla sig till liknande krav i alla dessa stater.

Dessa säkerhetsstandarder är dock inte i sig bindande. En näringsidkare kan välja andra metoder för att visa att en produkt eller tjänst är säker, till exempel genom att förhålla sig till ”god sed för produktsäkerhet i den berörda branschen” eller till ”den aktuella vetenskapliga och tekniska kunskapsnivån” (12 §).

Produktsäkerhetslagen innehåller vidare krav på att en näringsidkare ska lämna viss säkerhetsinformation (13 §). Om näringsidkaren får kännedom om att en vara eller tjänst som tillhandahållits är farlig kan hen bland annat vara skyldig att sända ut ett varningsmeddelande, återkalla varan eller tjänsten och underrätta tillsynsmyndigheten (14–19 §§ samt 23 §). Tillverkare har dessutom vissa skyldigheter att märka varan, utföra stickprovskontroller av tillhandahållna varor, granska och föra

[22] Europaparlamentets och rådets direktiv 2001/95/EG av den 3 december 2001 om allmän produktsäkerhet (senast ä+ndrat genom Europaparlamentets och rådets förordning (EG) nr 596/2009).

[23] Europaparlamentets och rådets förordning (EU) 2019/2144 om typgodkännande av fordon och Europaparlamentets och rådets förordning (EU) 2017/745 om medicintekniska produkter.

[24] Lagen genomför direktivet om allmän produktsäkerhet men går längre än direktivet på vissa punkter, bland annat genom att den också omfattar (vissa typer av) tjänster.

register över inkomna klagomål samt informera distributörerna om det förebyggande säkerhetsarbetet (20 §). Distributörer ska medverka till att tillverkaren fullgör sina skyldigheter, bland annat genom att bevara sådan dokumentation som behövs för att varornas ursprung ska kunna spåras (21 §).

En tillsynsmyndighet (oftast Konsumentverket) kan ingripa mot farliga varor och tjänster med vitesföreläggande eller vitesförbud (27–31 §§). Den näringsidkare som uppsåtligt eller oaktsamt har åsidosatt sina skyldigheter kan åläggas att betala en sanktionsavgift på mellan 5 000 och 5 miljoner kronor (37–41 §§).

Beträffande vissa produktkategorier krävs utöver efterlevnad av de allmänna och sektorsspecifika säkerhetskraven att produkten CE-märks. Proceduren för CE-märkning skiftar, men typiskt sett förkommer två varianter. Den ena varianten innebär att tillverkaren själv bedömer och intygar att de krav som ställs på produkten är uppfyllda. Den andra varianten innebär att granskningen utförs av ett oberoende organ som också utfärdar intyget.

Produktsäkerhetsregleringen är i grunden tillämplig även när AI-lösningar ingår i de varor och tjänster som tillhandahålls. Vissa av de sektorsspecifika produktsäkerhetsreglerna har anpassats till en verklighet där allt fler produkter innehåller digitala komponenter. Det gäller till exempel reglerna om medicintekniska produkter. Några specifika anpassningar till just AI har dock inte gjorts så här långt.[25]

Efter en genomgång har EU-kommissionen gjort bedömningen att produktsäkerhetsregleringen i huvudsak hanterar AI-teknikens utmaningar på ett tillfredsställande sätt. Vissa reformbehov har dock identifierats.[26]

Det autonoma beteendet hos vissa AI-system kan medföra viktiga produktförändringar som påverkar säkerheten under systemets livscykel. Det kan därför vara aktuellt att kräva att en ny riskbedömning görs. Mänsklig tillsyn, från AI-produkternas och AI-systemens konstruktionsfas och under hela livscykeln, kan dessutom behövas som en skyddsåtgärd. En fråga som enligt kommissionen kan övervägas är om regelverket uttryckligen bör adressera risken för att data av bristande kvalitet påverkar systemets säkerhet.[27]

Vidare konstateras att produktsäkerhetsregleringen skulle behöva kompletteras med krav för att motverka black-box-effekten, det vill säga systemens komplexitet och bristen på insyn. Kommissionen skissar bland annat på ett krav på att göra designparametrar för algoritmer och metadata rörande använda datamängder tillgängliga om en skada inträffar.[28]

[25] Rapport från kommissionen till Europaparlamentet, rådet och Europeiska Ekonomiska och sociala kommittén, *Konsekvenser för säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik* (COM (2020) 64 final), s. 4. Se även Europaparlamentets och rådets förordning (EU) 2017/745 av den 5 april 2017 om medicintekniska produkter.

[26] EU-kommissionens vitbok, s. 14.

[27] Rapport från kommissionen till Europaparlamentet, rådet och Europeiska Ekonomiska och sociala kommittén, *Konsekvenser för säkerhet och ansvar när det gäller artificiell intelligens, sakernas internet och robotteknik* (COM (2020) 64 final), s. 7-9.

[28] A.a. s. 9 f.

Slutligen bör det enligt kommissionen övervägas om det behövs särskilda regelförändringar för att motverka att programvaruuppdateringar under en produkts livscykel leder till skador.[29]

2.3 Ansvarsregleringen

Skadeståndsansvaret för AI-relaterade skador bestäms i huvudsak av nationella regler. På EU-nivå finns emellertid regler om produktansvar för skadevållande produkter,[30] om ansvar för överträdelse av dataskyddsregler och om konkurrensrelaterade skador, som kan bli tillämpliga.

Den generella svenska regleringen av rätten till skadestånd finns i *skadeståndslagen* (1972:207). Lagen har karaktär av en ramlag och innefattar således inte någon uttömmande reglering av rätten till skadestånd. Den kompletteras av allmänna rättsgrundsatser som har utvecklats i praxis.

En grundläggande regel i skadeståndslagen, liksom i de flesta andra länders lagstiftning, är att den som uppsåtligen eller av oaktsamhet orsakar en sak- eller personskada är skyldig att ersätta denna skada (2 kap. 1 §). Oaktsamheten kan till exempel bestå i att avvika från de aktsamhetsnormer som har utvecklats inom ett visst område. Den som kräver skadestånd måste kunna visa att han eller hon har lidit en ersättningsgill skada och att denna skada är en konsekvens av motpartens uppsåtliga eller oaktsamma agerande.

I skadeståndsrätten finns särskilda regler som ålägger vissa subjekt ett strikt ansvar eller åtminstone omvänd bevisbörda. Vissa länder har ett flertal regler av detta slag, medan andra, till exempel Sverige, har få (till exempel för hundar och katter samt för atomskador).

Ansvaret för så kallad *ren förmögenhetsskada*, det vill säga ekonomisk skada som uppkommer utan samband med att någon lider person- eller sakskada, är mer begränsat i svensk rätt. Sådan skada ersätts som huvudregel endast om den vållats genom brott, om den uppkommer i ett avtalsförhållande eller har särskilt lagstöd (2 kap. 2 §). Exempel på sådan lagreglering är rätten till ersättning enligt immaterialrättslagstiftningen. Av rättspraxis framgår att ren förmögenhetsskada även ska ersättas i vissa speciella fall, till exempel när någon har lämnat vilseledande information som leder till skada för den part som litade på informationen, under förutsättning att den skadelidande hade en befogad anledning att lita på informationen i det aktuella fallet (jämför till exempel NJA 1987 s. 692).

Skadeståndsansvaret i *avtalsrelationer* är reglerat i särskild lagstiftning, till exempel köplagen (1990:931) och konsumentköplagen (1990:932). Ansvaret enligt dessa lagar är mer långtgående, till exempel är ansvaret inte begränsat till skador som orsakas

[29] A.a. s. 10 f.

[30] Rådets direktiv 85/374/EEG av den 25 juli 1985 om tillnärmning av medlemsstaternas lagar och andra författningar om skadeståndsansvar för produkter med säkerhetsbrister, nedan ”produktansvarsdirektivet”.

genom oaktsamhet. I andra avtal än konsumentavtal kan parterna samtidigt komma överens om begränsningar av skadeståndsansvaret.

Enligt *produktansvarslagen* (1992:18) – som genomför EU:s produktansvarsdirektiv – ska skadestånd betalas för personskada som en produkt har orsakat på grund av en säkerhetsbrist. Motsvarande skadeståndsansvar gäller för sakskador, men bara beträffande egendom som till sin typ vanligen är avsedd för enskilt ändamål och om den skadelidande vid tiden för skadan använde egendomen huvudsakligen för sådant ändamål (1 §).[31]

Med produkt avses i lagen lösa saker (2 §), det vill säga fysiska manifestationer. En produkt anses ha en säkerhetsbrist om produkten inte är så säker som skäligen kan förväntas. Säkerheten ska bedömas med hänsyn till hur produkten kunnat förutses bli använd och hur den har marknadsförts samt med hänsyn till bruksanvisningar, tidpunkt då produkten satts i omlopp och övriga omständigheter (3 §).

Inte bara tillverkaren utan en mängd andra aktörer (till exempel den som importerat, marknadsfört eller tillhandahållit produkten) kan hållas ansvariga för en produkts säkerhetsbrist (6–7 §§). Därigenom förenklas de praktiska möjligheterna för den skadelidande att få ersättning.

Föreligger det en säkerhetsbrist är ansvaret som utgångspunkt strikt (1 §). Detta underlättar för den skadelidande som inte behöver visa att den ansvarige varit oaktsam och att denna oaktsamhet lett till skadan. Den ansvarige kan dock undgå ansvar i vissa uppräknade situationer, till exempel om denne kan göra sannolikt att säkerhetsbristen inte fanns när produkten sattes i omlopp eller visar att det på grundval av det vetenskapliga och tekniska vetandet vid den tidpunkt då produkten sattes i omlopp inte var möjligt att upptäcka säkerhetsbristen (8 §).

Produktansvaret är tvingande och kan följaktligen inte begränsas genom avtal.

De beskrivna skadeståndsreglerna är tillämpliga även när skador är kopplade till *användning av AI*. Den skadelidande kan ofta välja att rikta sina krav mot olika aktörer (tillverkare av AI-system, företag som tagit systemet i bruk etc.) och basera sin talan på olika rättsliga grunder (allmänt oaktsamhetsansvar, produktansvar, avtalsrättslig grund etc.). AI-teknikens egenskaper kan dock i praktiken göra det svårt för en skadelidande att få ersättning. Samtidigt skapas många gånger en oklar rättslig situation för företag som utvecklar eller använder ny teknik.

Vid tillämpningen av den allmänna oaktsamhetsregeln kan det på grund av AI-systemens komplexitet och föränderlighet samt mångfalden av inblandade aktörer vara svårt att bevisa vem som har orsakat en skada och att utesluta alternativa eller samverkande orsaker till en skada. Eftersom det handlar om en ny teknik kan det också vara svårt att visa vad som utgör en avvikelse från aktsamhetsstandarden.

[31] Produktansvarslagen reglerar inte skador på själva produkten. Denna typ av skador kan omfattas av köprättsliga regler (se till exempel 67 § köplagen).

När skadelidande på detta sätt har svårt att fullgöra sin bevisbörda kan en lösning vara att införa regler om strikt ansvar. Samtidigt är det inte helt självklart vem som i sådana fall skulle åläggas det strikta ansvaret: tillverkaren av ett system, den som förser ett system med data, den som använder system eller någon annan. Sett från ett rättspolitiskt perspektiv är det dessutom viktigt att inte skapa en ordning som skapar incitament mot att utveckla och använda nya tekniska lösningar, som på en samhällsövergripande nivå är säkrare än lösningar som bygger på mänskliga insatser. Utmaningen är alltså att skapa en balans mellan å ena sidan skadelidandes intresse av att enkelt kunna få ersättning, och därmed i förlängningen möjligheten att skapa tillit till användningen av AI, och å andra sidan incitamenten att våga satsa på ny teknik med stor samhällspotential. Viktiga inslag i en sådan balans kommer troligen att vara olika former av (obligatoriska) försäkringslösningar.

I studier och rapporter från senare år dras ofta slutsatsen att EU:s *produktansvarsreglering* i grunden är relevant och effektiv, men att vissa egenskaper hos den nya tekniken i praktiken gör det svårare för skadelidande att få ersättning för skador. I andra sammanhang har det dock framhållits att produktansvarsregleringen lider av strukturella brister som resulterar i att skadelidande på grund av risken för höga rättegångskostnader avstår från att göra sina rättigheter gällande. Det har också framhållits att produktansvarsdirektivet har resulterat i bristande harmonisering.[32]

Produktansvarets tillämpning på tillhandahållande av programvara är dessutom osäker. När programvara är inbyggd i en vara är produktansvarslagen utan tvekan tillämplig på personskador som orsakas av programvaran. Men om samma programvara tillhandahålls på en fristående bärare eller i form av en onlinetjänst är det mer tveksamt om produktansvarslagen är tillämplig. EU-kommissionen har i sin genomgång framhållit att det särskilt finns en osäkerhet kring hur en uppdatering av programvara ska betraktas. En lösning skulle möjligen vara att betrakta denna som tillhandahållande av en ny produkt.

En annan svaghet som har påtalats är att den skadelidande måste kunna styrka såväl skadan som säkerhetsbristen och orsakssambandet mellan dessa två. Den skadelidande måste även kunna visa att produkten inte är så säker som rimligen kan förväntas. Sett ur den skadelidandes perspektiv devalverar detta värdet av lagens i grunden strikta ansvar. För tillverkare med flera är det samtidigt angeläget att inte kunna hållas ansvarig för förhållanden som ligger utanför deras kontroll.

2.4 Diskrimineringslagstiftningen, dataskyddslagstiftningen med mera

Flera av de risker som är förknippade med AI är inte kopplade till risken för person- eller sakskador, utan till risken för kränkningar av grundläggande fri- och rättigheter. Det handlar till exempel om diskriminerande AI-system eller system som inte tillgodo-

[32] European Parliament, *Artificial Intelligence and Civil Liability,* Study Requested by the JURI committee, Author: Andrea Bertolini, Bryssel 2020, s. 54 ff.

ser enskildas anspråk på rättssäkerhet eller rätt till insyn. I policydiskussionen kring AI är det ofta risker av detta slag som får störst uppmärksamhet.

Av diskussionen är det ibland lätt att få intrycket att tillämplig rättslig reglering för att motverka denna typ av kränkningar helt saknas. Så är emellertid inte fallet. Två omfattande och relativt krävande regelverk som påverkar utveckling och användning av AI är diskrimineringslagstiftningen och dataskyddslagstiftningen. Båda dessa regleringar är direkt tillämpliga även i den privata sektorn. Därutöver måste alla grundläggande rättigheter i EU:s rättighetsstadga beaktas vid genomförandet och tillämpningen av andra EU-rättsliga bestämmelser, till exempel sektorsspecifik reglering inom den finansiella sektorn. En förklaring till att skyddet mot AI-relaterade kränkningar trots detta ibland uppfattas som bristande kan vara att de relevanta reglerna är relativt komplexa samtidigt som den exakta tillämpningen på AI-relaterade fenomen ännu inte är fastställd genom rättspraxis eller vägledning från tillsynsmyndigheter.

Inte heller på detta område är det möjligt att göra en fullständig genomgång av alla relevanta regler och deras betydelse för utvecklingen och användningen av AI. Framställningen begränsas till en kort redogörelse för diskriminerings- och dataskyddslagstiftningens relevans i sammanhanget.

Ett *förbud mot diskriminering* finns uttryckt i grundläggande instrument som skyddar mänskliga rättigheter, till exempel Europeiska konventionen om skydd för de mänskliga rättigheterna och Europeiska unionens stadga om de grundläggande rättigheterna. Den svenska regleringen finns i huvudsak i diskrimineringslagen (2008:567), som bland annat genomför ett antal EU-direktiv.

I diskrimineringsrätten görs en åtskillnad mellan direkt och indirekt diskriminering, men båda formerna är ofta otillåtna. *Direkt diskriminering* innebär att någon ”missgynnas genom att behandlas sämre än någon annan behandlas, har behandlats eller skulle ha behandlats i en jämförbar situation, om missgynnandet har samband med kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning eller ålder”. *Indirekt diskriminering* innebär att någon ”missgynnas genom tillämpning av en bestämmelse, ett kriterium eller ett förfaringssätt som framstår som neutralt men som kan komma att särskilt missgynna personer med visst kön, viss könsöverskridande identitet eller uttryck, viss etnisk tillhörighet, viss religion eller annan trosuppfattning, viss funktionsnedsättning, viss sexuell läggning eller viss ålder, såvida inte bestämmelsen, kriteriet eller förfaringssättet har ett berättigat syfte och de medel som används är lämpliga och nödvändiga för att uppnå syftet” (1 kap. 4 § diskrimineringslagen).

Även om syftet med ett AI-system inte varit att missgynna, till exempel på grund av kön, kan detta bli resultatet, till exempel på grund av programmeringsmisstag eller på grund av bristande datakvalitet i maskininlärningsfasen. Det kan då vara fråga om indirekt diskriminering. Den som använder sig av ett AI-system inom sådana områden som enligt 2 kap. diskrimineringslagen omfattas av förbud mot diskriminering (till exempel arbetslivet, tillhandahållande av varor och tjänster till allmänheten samt hälso- och sjukvården) måste därför försäkra sig om systemets

kvalitet och sina egna arbetsprocesser. Den som bryter mot diskrimineringslagen kan åläggas att betala en relativt kännbar diskrimineringsersättning till den som kränkts (5 kap.).

Diskrimineringslagstiftningens effektivitet i relation till AI-tillämpningar kan samtidigt ifrågasättas i vissa hänseenden. För det första utgör förbudet mot indirekt diskriminering inte någon helt tydlig norm att förhålla sig till i den praktiska tillämpningen. Om det finns objektiva skäl till att ett visst bedömningskriterium etc. har valts framgår det av rättspraxis att tillåtligheten ska avgöras genom en proportionalitetsbedömning i det enskilda fallet. För det andra innebär maskininlärningstekniken ofta att AI-systemet på egen hand får identifiera kriterier och tillvägagångssätt genom analys av tillgängliga datamängder. Dessa kriterier och tillvägagångssätt tillämpas sedan när AI-systemet utför sina uppgifter, till exempel fattar beslut i ett ärende. För det tredje är det ofta svårt för berörda personer eller tillsynsmyndigheter att upptäcka och bevisa otillåten diskriminering när många AI-system fungerar som en black-box. I praktiken krävs tillgång till en översikt över systemets åtgärder samt underliggande data för att lyckas i detta hänseende.

Diskrimineringslagen ställer visserligen också krav på aktiva åtgärder för att motverka diskriminering. I dagsläget gäller dessa krav emellertid bara vissa områden (arbetsliv och utbildning). Kraven är dessutom inte särskilt utformade för att motverka sådana risker som uppkommer till följd av användningen av AI. Men det är givetvis fullt möjligt att utveckla denna del av diskrimineringslagstiftningen på ett sätt som skapar en rimlig balans mellan olika aktörers legitima intressen. Det bör i detta sammanhang uppmärksammas att AI-utvecklingen inte alltid skapar nya problem, utan ofta bidrar till att redan existerande problem uppmärksammas. Det gäller såväl förekomsten av diskriminering i samband med mänskligt beslutsfattande som svagheter i lagstiftningen. Lösningen är därför inte alltid att specifikt reglera just AI.

Dataskyddslagstiftningen är ett rättsligt verktyg för att skydda enskildas grundläggande rättigheter, särskilt skyddet för personuppgifter och rätten till skydd för den personliga integriteten. Det centrala regelverket på området är EU:s dataskyddsförordning (GDPR).[33]

Dataskyddslagstiftningen innehåller en omfattande och relativt komplex reglering av i princip all behandling av personuppgifter. Grunden i dataskyddslagstiftningen utgörs av vissa principer (ändamålsbegränsning, uppgiftsminimering med mera) samt kravet på rättsligt stöd för varje behandling (samtycke, intresseavvägning med mera). Behandlingen av så kallade känsliga personuppgifter (uppgifter om hälsa, biometriska uppgifter som används för identifiering med mera) och överföringar till ett land utanför EU/EES måste uppfylla särskilda förutsättningar för att vara tillåtna. Reglerna om när behandlingen av personuppgifter är tillåten kompletteras av särskilda rättigheter för de registrerade (rätt till information, rätt till radering

[33] 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). Se även lagen (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning (”dataskyddslagen”). Inom vissa områden finns också sektorsspecifika dataskyddsregler.

med mera) och av krav på aktiva åtgärder från den personuppgiftsansvarige, det vill säga den organisation som behandlar personuppgifterna (säkerhetskrav, krav på att utse ett så kallat dataskyddsombud med mera). Överträdelser av dataskyddsreglering kan straffas genom kännbara sanktionsavgifter utfärdade av tillsynsmyndigheten (Integritetsskyddsmyndigheten). Personer som lidit materiell eller immateriell skada har även rätt till ersättning.

Eftersom begreppet personuppgift har givits en vid innebörd (varje upplysning som avser en identifierad eller identifierbar fysisk person) blir dataskyddslagstiftningen av stor betydelse för de flesta former av AI. En analytisk uppdelning kan göras mellan sådan behandling av personuppgifter som sker i samband med maskininlärningen, det vill säga för att utveckla systemets bedömningskriterier med mera, och sådan behandling som sker i samband med att systemet används, till exempel för att fatta beslut som rör enskilda personer. I det första fallet handlar det om en behandling som kan liknas vid den behandling som sker vid framställning av statistik, där det i och för sig kan handla om en hantering av känsliga personuppgifter eller av stora mängder personuppgifter, men där intresset inte är riktat mot den enskilda individen utan mot principiella (statistiska) samband. I det senare fallet kan behandlingen däremot gå ut på att AI-systemet ska fatta beslut eller vidta åtgärder som rör en enskild individ baserat på tillgängliga personuppgifter om personen i fråga. Vid maskininlärningen finns det visserligen risk för läckage av uppgifter eller annan obehörig användning, men de ovan behandlade riskerna med AI är annars hänförliga till den behandling som sker i samband med beslutsfattandet etc.

Det är inte möjligt att i detta sammanhang fullständigt redogöra för hur dataskyddslagstiftningen motverkar riskerna för kränkning förknippade med AI. På ett övergripande plan kan det dock konstateras att det finns flera olika regler i dataskyddslagstiftningen som är av särskilt intresse i sammanhanget.

- När AI-system ges rätt att fatta beslut baserat på personuppgifter kan den särskilda skyddsregleringen om automatiserat beslutsfattande i artikel 22 i GDPR bli tillämplig. Därmed ges den registrerade i praktiken ett extra starkt inflytande över om behandlingen av uppgifterna ska få ske.
- Dataskyddslagstiftningen innehåller vidare flera transparenskrav som kan motverka AI-systemens black-box-problematik. I samband med automatiserat beslutsfattande har den personuppgiftsansvarige en särskild skyldighet att aktivt lämna ”meningsfull information om logiken bakom samt betydelsen och de förutsedda följderna av sådan behandling för den registrerade” (artikel 13.2 f och 14.2 g).
- I praktiken kommer den behandling av personuppgifter som sker vid användningen av AI ofta att utgöra en sådan högriskbehandling som kräver att det görs en så kallad konsekvensbedömning (en slags riskanalys) innan behandlingen inleds (artikel 35). Den personuppgiftsansvarige är även skyldig att aktivt vidta adekvata säkerhetsåtgärder (artikel 32) och att skapa system som uppfyller principerna om inbyggt dataskydd och dataskydd som standard (artikel 25).
- Biometriska uppgifter som används för identifiering räknas som känsliga personuppgifter, vilket innebär att möjligheterna att hantera denna typ av uppgifter i

AI-system är begränsade (artikel 9), särskilt i kombination med automatiserat beslutsfattande (artikel 22).

- Integritetsskyddsmyndigheten är behörig att utöva tillsyn över i princip all behandling av personuppgifter och att utdöma sanktionsavgifter vid överträdelser.

Den samlade bilden är att dataskyddslagstiftningen innehåller flera av de verktyg som behövs för att hantera de risker för kränkningar som är förknippade med AI-system. Därmed inte sagt att det inte finns utrymme för förbättringar.

Samtidigt kan det konstateras att dataskyddsregleringen i vissa hänseenden framstår som alltför restriktiv och riskerar att motverka möjligheterna till en legitim AI-utveckling. Det handlar framför allt om behandlingsreglernas inverkan på möjligheterna att använda stora mängder personuppgifter, inklusive känsliga personuppgifter, för maskininlärning. Den begränsade möjligheten att behandla känsliga personuppgifter kan också försvåra arbetet med att upptäcka och motverka indirekt diskriminering i AI-system. Att inte behandla några känsliga personuppgifter skyddar nämligen inte mot att AI-systemet väljer att använda kriterier som starkt samvarierar med en diskrimineringsgrund.

3 Förslag till ny rättslig reglering av AI

3.1 Inledande anmärkningar

I tidigare avsnitt har behoven av att anpassa existerande regelverk till AI-utvecklingen uppmärksammats. I detta avsnitt beskrivs vissa större rättsliga reformförslag som har presenterats under det senaste året. Framställningen är kortfattad och tar endast sikte på de mest centrala elementen i förslagen. Syftet är framför allt att lägga grunden för den principiella diskussionen om regleringen av AI i rapportens avslutande del.

3.2 EU-kommissionens vitbok

EU-kommissionen har som uppmärksammats i det föregående gjort en översiktlig genomgång av delar av säkerhets- och ansvarsregleringen och därvid pekat på vissa brister.[34] Möjliga lagstiftningsförändringar har skisserats, dock utan att några konkreta förslag har lagts fram. För att möta de tekniska och kommersiella utmaningarna förknippade med AI menar kommissionen att det utöver sådana eventuella ändringar krävs en lagstiftning som särskilt är inriktad på AI.[35]

För att tillgodose behovet av säkerhet och tillit utan att begränsa innovation och utveckling i onödan förordar kommissionen att den nya regleringen bygger på en *riskbaserad ansats*. Med det avses att den nya kompletterande regleringen bara ska gälla AI-tillämpningar som bedöms vara förknippade med hög risk.[36]

Vad som utgör hög risk ska enligt kommissionens skiss bestämmas utifrån den sektor där AI-tillämpningen används och den aktuella *användningens karaktär*. Exempel på sektorer med hög risk som nämns är hälso- och sjukvård, transport och energi. Enligt kommissionen bör reglerade sektorer förtecknas i en bilaga till regelverket som regelbundet uppdateras. Eftersom varje användning inom en sektor inte medför särskilda risker ska ett kompletterande kriterium användas för att identifiera högriskverksamhet. Centralt i detta sammanhang är enligt kommissionen vilken påverkan systemet har på berörda parter. Som hög risk i detta hänseende kan till exempel räknas att beslut med

[34] Se avsnitt 1.4 och 2. Någon närmare analys av hur lagstiftningen rörande mänskliga rättigheter, till exempel dataskydds- och diskrimineringslagstiftningen, skyddar medborgarna vid användningen av AI görs inte i vitboken eller i den anknytande rapporten.

[35] EU-kommissionens vitbok, s. 18.

[36] EU-kommissionens vitbok, s. 19 f.

rättslig effekt fattas eller att åtgärder som riskerar att skada en person fysiskt vidtas av systemet. Vissa applikationer, till exempel fjärridentifiering med biometri (såsom ansiktsigenkänning), ska emellertid alltid anses utgöra en högrisktillämpning.[37]

Kommissionen skissar i vitboken på möjliga *krav* som kan ställas på AI-applikationer med hög risk. I huvudsak handlar det om förebyggande krav som syftar till att förhindra kränkningar (till exempel diskriminering) och uppkomsten av säkerhetsbrister. Kraven rör bland annat hanteringen av träningsdata, transparenskrav, robusthet och korrekthet samt tillgången till mänsklig insyn och medverkan. Dessutom diskuteras särskilt strikta villkor för vissa utvalda tillämpningar, till exempel fjärridentifiering med biometri.[38]

Kommissionen anger att den överväger om de förebyggande kraven som ställs på högrisktillämpningar ska bli föremål för *förhandskontroll*, till exempel genom testning, inspektion eller certifiering. Sådana eventuella krav måste i sådana fall samordnas med redan befintliga kontrollordningar kopplade till sektorsspecifik produktsäkerhetsreglering. Vidare framhåller kommissionen behovet av behöriga och kompetenta tillsynsmyndigheter.[39]

För AI-tillämpningar som inte utgör högrisk-AI överväger kommissionen att föreslå att det inrättas ett *frivilligt märkningssystem*.[40]

3.3 Europaparlamentets skadeståndsresolution

Europaparlamentet har i oktober 2020, i form av en resolution riktad till EU-kommissionen, presenterat ett förslag till en ny EU-förordning innehållande en skadeståndsordning för artificiell intelligens.[41] I korthet innebär förslaget att en operatör av ett autonomt AI-system som innebär hög risk ska åläggas ett strikt ansvar för de skador som systemet orsakar. Förslaget omfattar personskador, sakskador och rena förmögenhetsskador.

Förslaget utgår från att produktansvaret för tillverkare av AI-system behålls och att det görs vissa förtydliganden av produktansvarsdirektivet i linje med vad som diskuterats i avsnitt 2.3 ovan. Den nya regleringen föreslås komplettera *tillverkarnas* produktansvar med en reglering som tar sikte på operatörer av AI.

Förslaget tar sikte både på så kallade frontend- och backend-operatörer. Med *frontend-operatör* avses i förslaget ”varje fysisk eller juridisk person som utövar en viss kontroll över en risk som är förknippad med AI-systemets drift och sätt att fungera och som drar nytta av dess drift”. Med *backend-operatör* avses ”varje

[37] EU-kommissionens vitbok, s. 20.

[38] EU-kommissionens vitbok, s. 20 ff.

[39] EU-kommissionens vitbok, s. 25 f.

[40] EU-kommissionens vitbok, s. 26 f.

[41] Europaparlamentets resolution av den 20 oktober 2020 med rekommendationer till kommissionen om en skadeståndsordning för artificiell intelligens (2020/2014(INL)), nedan ”Europaparlamentets skadeståndsresolution”.

fysisk eller juridisk person som fortlöpande definierar teknikens egenskaper och tillhandahåller data och ett väsentligt backend-stöd och därför även utövar en viss kontroll över den risk som är förknippad med AI-systemets drift och sätt att fungera" (artikel 3).

Operatörer av AI-system med hög risk åläggs ett *strikt ansvar* för de skador som systemet orsakar. Operatören ska inte kunna friskriva sig eller demonstrera frånvaro av oaktsamhet för att kunna undgå ansvar. Däremot begränsas ansvaret vid force majeure (artikel 4).

Hög risk definieras i detta sammanhang som "en betydande risk för att ett autonomt fungerande AI-system kan orsaka skada för en eller flera personer på ett sätt som är slumpmässigt och som går utöver det som rimligen kan förväntas; hur betydande denna risk är beror på samspelet mellan hur allvarlig den eventuella skadan är, graden av självständigt beslutsfattande, sannolikheten för att risken förverkligas samt hur och i vilken miljö AI-systemet används" (artikel 3).

Alla AI-system med hög risk och alla kritiska sektorer där de används ska förtecknas i en bilaga till förordningen. EU-kommissionen ska bemyndigas att besluta om uppdateringar av denna bilaga. Frontend- och backend-operatörer av AI-system med hög risk åläggs att teckna relevanta försäkringar (artikel 4).

Förordningsförslaget innehåller en reglering av högsta ersättningsbelopp. Beträffande personskador är det högsta beloppet 2 miljoner euro och för övriga skador 1 miljon euro. Takbeloppen gäller för samtliga som lidit skada "genom en och samma process som utförts av ett och samma AI-system med hög risk" (artikel 5).

Beträffande sådana AI-system som inte kategoriseras som system med hög risk har operatören ett oaktsamhetsansvar. Vissa uttryckliga grunder som kan åberopas för att undgå ansvar räknas upp i förslaget (artikel 8).

Som utgångspunkt är flera inblandade operatörer solidariskt ansvariga och har regressrätt mot varandra. Om en aktör räknas både som en operatör enligt den föreslagna förordningen och som tillverkare enligt produktansvarsdirektivet innehåller förslaget vissa konfliktregler (artikel 11).

Regleringen föreslås bli tvingande, det vill säga den ska inte kunna åsidosättas genom avtal. Rätten till ersättning enligt andra författningar eller på grund av avtal kvarstår (artikel 1).

3.4 Europaparlamentets etikresolution

Europaparlamentet har i oktober 2020, i form av en resolution riktad till EU-kommissionen, presenterat ett förslag till en helt ny EU-förordning om etiska principer för utveckling, spridning och användning av artificiell intelligens, robotteknik och

tillhörande teknik.[42] Förslaget kan ses som ett försök att omvandla kommissionens expertgrupp för AI-frågors förslag till etiska riktlinjer till ett bindande rättsligt regelverk. Det bygger vidare på kommissionens skisser i vitboken (jämför avsnitt 3.2 ovan).

Den föreslagna förordningen är i huvudsak inriktad på AI-tillämpningar med hög risk, men innehåller vissa regler som gäller all användning av AI.

Tillämpningsområdet är brett. Regleringen gäller inte bara utvecklare av AI, utan även den som distribuerar eller använder AI. All användning av AI i EU, oavsett var tillhandahållaren är etablerad eller var utvecklingen har skett, omfattas. Tekniker som berörs är AI-system, robottekniker samt tillhörande teknik (till exempel system för att identifiera biometriska eller genetiska data).

I artikel 5 anges grundkraven för all utveckling, spridning och användning av AI, nämligen att unionslagstiftningen (till exempel GDPR) ska efterlevas samt att människans värdighet, självbestämmanderätt och säkerhet och andra grundläggande rättigheter enligt stadgan ska respekteras.

Beträffande AI med *hög risk* gäller enligt artikel 6–12 mer strikta krav. Hög risk definieras i förordningen som ”en betydande risk för att utveckling, spridning och användning av artificiell intelligens, robotteknik och tillhörande teknik kan bli till skada eller förfång för enskilda personer eller för samhället, genom att det därvid bryts mot i unionslagstiftningen förankrade grundläggande rättigheter och säkerhetsbestämmelser, varvid hänsyn ska tas till användningens eller ändamålets särdrag, samt till sektorn inom vilken tekniken utvecklas, sprids eller används och till svårhetsgraden av den skada eller det förfång som kan förväntas uppkomma” (artikel 4). I en bilaga till förordningen anges ett antal ändamål och sektorer som ska göra att AI-verksamheten alltid anses vara förknippad med hög risk.

Högrisk-AI får bara utvecklas, spridas och användas om den

- garanterar full mänsklig tillsyn vid varje tillfälle och på ett sätt som gör det möjligt att återfå full mänsklig kontroll när så behövs (artikel 7)
- uppfyller vissa uppräknade krav på cybersäkerhet, transparens och förklarbarhet (artikel 8)
- inte fungerar snedvridande eller diskriminerande (artikel 9)
- uppfyller etiska principer om socialt ansvarstagande, jämställdhet och miljöansvar (artikel 10–11)
- respekterar strikta integritets- och dataskyddsprinciper för fjärridentifiering, till exempel ansiktsigenkänning, på allmän plats (artikel 12).

Varje fysisk eller juridisk person ska ha rätt att få ersättning för skada som orsakats av utveckling, spridning och användning av artificiell intelligens i strid med unionsrätten och de skyldigheter som fastställs i förordningen (artikel 13).

[42] Europaparlamentets resolution av den 20 oktober 2020 med rekommendationer till kommissionen om en ram för etiska aspekter av artificiell intelligens, robotteknik och tillhörande teknik (2020/2012(INL)), nedan ”Europaparlamentets etikresolution”.

Efterlevnaden av kraven på högrisk-AI ska enligt förslaget övervakas av nationella tillsynsmyndigheter. För övrig AI införs ett frivilligt europeiskt intyg om etisk efterlevnad (artikel 15–16).

Den sammantagna bilden av förslaget är att det är mycket långtgående och krävande för den som vill utveckla, sprida och använda AI-lösningar. Ersättningsskyldigheten i förhållande till de vaga materiella kraven skulle innebära en stor riskexponering för alla inblandade aktörer.

3.5 Vissa reaktioner på förslagen

Många aktörer har lämnat synpunkter på den inriktning för AI-relaterad lagstiftning som EU-kommissionen har stakat ut i vitboken och som sedermera har följts upp genom Europaparlamentets båda resolutioner.

Ett politiskt tungt uttalande är det som Sverige och 13 andra medlemsstater gjorde i oktober 2020. I ett så kallat *non-paper* framhålls att innovation och tillförlitlig AI utgör två sidor av samma mynt.[43] Gruppen uttalar sitt stöd för en riskbaserad ansats vid regleringen av AI, men tycks vilja betona att den av kommissionen skisserade regleringsmodellen riskerar att leda till att balansen mellan de relevanta intressena inte blir korrekt.

I uttalandet framhålls behovet av en noggrann analys av det gällande regelverket och dess eventuella brister i förhållande till AI-tekniken. Vidare understryks behovet av ett helt harmoniserat regelverk för den inre marknaden.

Sektor och användningsområde kan enligt gruppen inte vara de enda kriterierna för att avgöra vad som utgör en sådan högrisktillämpning som ska vara föremål för särskilda rättsliga krav i en eventuell ny lagstiftning. Istället förordas en lösning som innebär att det etableras ett proportionerligt regelverk och en metodologi för att göra riskbedömningar i det enskilda fallet. Vid riskbedömningen bör den potentiella effekten av en skada och risken för att skadan inträffar beaktas. Vidare framhålls att reglerna bör utformas så att kategorin högrisktillämpning blir undantaget, inte huvudregeln.

Slutligen menar de 14 staterna att den europeiska AI-policyn bör vara att etablera tillitsfull AI som en konkurrensfördel. Istället för att införa tvingande regler som kan fungera som hinder för innovationen bör det skapas incitament för utvecklare att proaktivt och systematiskt stimulera användningen av tillitsfull AI. Soft law-lösningar i form av självreglering och frivilliga lösningar, liksom en robust standardiseringsprocess, förordas. Särskilt framhålls värdet av en frivillig europeisk märkning av sådan AI som bedöms vara säker, robust och etisk.

[43] Trustworthy and innovative AI, Position paper on behalf of Denmark, Belgium, the Czech Republic, Finland, France, Estonia, Ireland, Latvia, Luxembourg, the Netherlands, Poland, Portugal, Spain and Sweden on innovative and trustworthy AI.

4 Framtidens AI-reglering – några kommentarer

4.1 Inledning

Denna rapport har handlat om hur den rättsliga regleringen hanterar vissa typer av risker förknippade med AI. Gällande reglers tillämpning har analyserats och rättsliga reformförslag har beskrivits. Mot denna bakgrund lämnas här vissa synpunkter på framtidens AI-reglering.

4.2 Vilka krav kan ställas på regleringen av AI?

Ett centralt syfte med regleringen av AI är att skapa tillit till tekniken och därmed främja dess användning. Rätt utformad kan regleringen minska riskerna för skador och kränkningar och samtidigt se till att de som ändå drabbas får ersättning. För att detta ska uppnås måste det finnas tydliga regler och mekanismer för att se till att dessa regler efterlevs på ett effektivt sätt.

Med tanke på den stora potential som är förknippad med AI är det samtidigt viktigt att regleringen är *proportionerlig* och inte i onödan hindrar användningen av AI eller försämrar teknikens kvalitet genom alltför långtgående begränsningar. Det är också viktigt att regleringen är *förutsebar* och rimligt enkel att tillämpa. Dessa krav gäller både befintlig och eventuell ny reglering.

I det omfattande policy- och lagstiftningsarbete som dragits igång inom EU saknas ofta en analys av hur *befintlig lagstiftning* riskerar att begränsa AI-utvecklingen. Vi har till exempel i denna rapport sett att gällande dataskyddsregler riskerar att motverka maskininlärning av hög kvalitet i vissa sammanhang. Sådana analyser skulle naturligtvis sällan kunna leda till att de aktuella reglerna helt togs bort. Däremot skulle det ibland vara möjligt att identifiera möjliggörande särlösningar som samtidigt innehåller kompletterande skyddsmekanismer. Analyserna skulle också kunna leda till tydligare vägledning kring hur befintliga regler ska kunna tolkas och tillämpas på ett sätt som också beaktar teknikutvecklingens fördelar.

Inom EU är ett vanligt sätt att visa politisk handlingskraft att anta *nya regler*. Nya regler kan många gånger vara motiverade, men det är viktigt att reglerna inte görs mer

långtgående än vad som krävs och därmed hämmar värdefull användning. Här bör det särskilt beaktas att nya regler inriktade på AI ofta föreslås gälla utöver redan befintliga regler. Även konsekvenserna av en eventuell ny reglering inriktad mot vad som beskrivs som AI med hög risk måste därför utvärderas noggrant. Är tillämpningsområdet tillräckligt tydligt preciserat och avgränsat? Är de materiella krav som ställs nödvändiga i all verksamhet som kommer att omfattas?

För att investering och användning av AI ska främjas är det också viktigt med *enhetliga spelregler*. För det första krävs en harmoniserad reglering på den inre marknaden, men även gemensam vägledning kring hur gällande regler ska tolkas och tillämpas på AI-relaterade fenomen. För det andra måste det beaktas att AI är ett i högsta grad internationellt fenomen. Utveckling och maskininlärning kan till exempel förläggas till länder utanför unionen som saknar ett adekvat rättsligt skydd, till exempel skydd för personuppgifter. Samtidigt kan den färdiga AI-lösningen många gånger, helt eller delvis, tillhandahållas på distans från länder utanför unionen. Det är viktigt att den europeiska regleringen inte missgynnar företag som väljer att utveckla och tillhandahålla tjänster inom unionen.

4.3 En svårreglerad företeelse

AI utgör – vilket bland annat denna rapport visat – inte en företeelse som är enkel att reglera. AI utgör i praktiken en bred samling tekniker som dessutom är i snabb utveckling och som hela tiden får ett bredare användningsområde. Bristen på en tydlig definition av vad som avses med AI är i detta sammanhang ett problem. Det finns en tydlig risk att ny lagstiftning utformas med vissa specifika tillämpningar för ögonen, till exempel självkörande fordon eller automatiserade system för rekrytering, samtidigt som den färdiga regleringen får ett betydligt bredare tillämpningsområde.

Det måste därför ställas höga krav på kunskap och öppen debatt när nya regler tas fram. Det behövs forskning, men även möten mellan företrädare för olika professioner för att kartlägga regleringsförutsättningarna. Exempel på andra åtgärder för att få mer djupgående kunskap om förutsättningarna för reglering är *praktiska testbäddar såsom reglerade sandlådor*, för att undersöka tillämpningen av befintlig lagstiftning på olika typer av AI-lösningar, och simuleringar av hur ny lagstiftning skulle påverka utvecklingen och användningen av AI inom EU.

Denna typ av arbete tar tid och därmed väcks frågan om vad som är rätt tidpunkt för ny reglering. Den nya EU-kommissionen har presenterat en ambitiös agenda med snäva tidsramar för att ”reglera AI”. Mycket talar för att det är bättre att inte stressa igenom en ny heltäckande reglering, utan istället utföra arbetet mer stegvis, till exempel genom en noggrann översyn av olika befintliga regelverk, inte minst de sektorsspecifika produktsäkerhetsreglerna samt dataskydds- och diskrimineringslagstiftningen, innan en helt ny bredare regleringsansats eventuellt tar vid.

På ett principiellt plan kan det diskuteras om en sådan bredare generell AI-reglering alls är lämplig eller önskvärd. Det handlar som sagt om många olika tekniker med en mängd olika användningsområden, vilket gör att en mängd olika reglerings-

modaliteter aktualiseras. Det finns en betydande risk att en one-size-fits-all-reglering, även om den inriktas på högrisk-AI, kommer att brista i tydlighet och leda till såväl över- som underreglering. Som framgått i rapporten råder det knappast någon brist på regler som är tillämpliga på AI, och de risker med AI-tillämpningar som har identifierats är av en sådan karaktär som normalt hanteras inom ramen för de befintliga regelverken.

4.4 Etik och standarder utgör en viktig grund

Som framhållits i rapporten ersätter etisk vägledning, självregleringsinitiativ och standarder inte en adekvat rättslig reglering av AI-relaterade företeelser. Men även det omvända gäller; all styrning bör inte göras till rättslig styrning. Europaparlamentets förslag till etikresolution utgör till exempel ett tveksamt försök i denna riktning.

Etiska regler med mera kan många gånger, särskilt i ett tidigt utvecklingsskede, vara ett viktigt komplement till befintlig lagstiftning. De dokument som har tagits fram av kommissionens expertgrupp på hög nivå för AI-frågor ger till exempel värdefull vägledning kring hur det praktiska arbetet kan bedrivas samtidigt som de erbjuder en nödvändig grad av flexibilitet med tanke på typ av tillämpning och användningsområde.[44]

Etiska regler och standarder kan också med tiden få direkt rättslig betydelse. Etablerade och allmänt accepterade principer för utveckling och användning av AI kan till exempel påverka skadeståndsrättsliga aktsamhetsbedömningar och bedömningen av vad som utgör en korrekt (fair) behandling av personuppgifter enligt artikel 5 i GDPR. Standarder kan användas för att visa att en produkt uppfyller produktsäkerhetsregleringens säkerhetskrav.

Om etiska regler och standarder kombineras med en *frivillig märkning* kan marknadskrafter och opinionsbildning användas för att främja utvecklingen av en tillitsfull AI. Det finns därför anledning att gå vidare med de förslag till en gemensam frivillig europeisk märkning som kommissionen och Europaparlamentet tagit fram.

4.5 Befintliga regelverk bör vara utgångspunkten

Det finns ingen genväg; vägen till en adekvat rättslig reglering av AI går via en noggrann analys av den befintliga rättsliga regleringens eventuella tillkortakommanden. En sådan analys har bara delvis genomförts innan förslag på helt nya regelverk läggs fram. EU-kommissionens vitbok innehåller ingen redovisning av hur befintliga regler om diskriminering och dataskydd kan tillämpas på AI och vari eventuella brister består. Samtidigt tar det skisserade förslaget till helt ny reglering sikte på just aspekter

[44] Etiska riktlinjer för tillförlitlig AI, 2019 samt Checklista för tillförlitlig AI: Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment, 2020.

som borde kunna hanteras inom ramen för dessa båda regelverk (med eller utan ändringar). Som påpekats ovan saknas även en tydlig analys av vilka omotiverade rättsliga hinder som finns för utvecklingen av AI.

Även om analysen har varit översiktlig indikerar denna rapport att de flesta utmaningarna förknippade med AI kan hanteras genom ändringar i befintliga regelverk. Därmed inte sagt att detta arbete är enkelt, tvärtom leder AI:s speciella egenskaper många gånger till fundamentala regleringsutmaningar. Dessa utmaningar kan emellertid inte rundas genom att ett helt nytt regleringslager inriktat på just AI läggs över det redan befintliga, utan utmaningarna måste ändå hanteras inom respektive reglering.

Möjliga reformer av befintliga regelverk har berörts kortfattat i avsnitt 2, men en djupare analys krävs naturligtvis på flera punkter.

4.6 Välj i första hand sektorsspecifika regler

Det är sällan rimligt att göra samma avvägning mellan nytta och risk på alla områden där AI kan användas. Vissa tillämpningar bör till exempel förbjudas helt i vissa sektorer, medan andra tillämpningar inte behöver någon ytterligare reglering än den som redan finns.

Eftersom en generell reglering av AI riskerar att leda till såväl över- som underreglering bör risker förknippade med AI i första hand motverkas genom sektorsspecifika regler. Ett sådant förhållningssätt har exempelvis präglat utformningen av produktsäkerhetsregleringen. Diskrimineringslagstiftning bygger på ett liknande förhållningssätt och innehåller en uppräkning av verksamheter där diskrimineringsförbud av visst slag gäller och anger vilka aktiva åtgärder som ska vidtas mot diskriminering på olika områden.

Den riskbaserade ansatsen, som bör prägla förhållandet till reglering av AI, bör alltså tillämpas också av lagstiftaren. Regler bör bara införas för sektorer och användningar där de verkligen behövs. Ett sådant angreppssätt möjliggör samtidigt att rättsliga lösningar enklare kan utvärderas och att slutsatser kan dras till exempel inför regleringen av andra sektorer eller inför införandet av en mer generell reglering.

4.7 En helt ny generell AI-reglering?

Det finns, som framgått i detta avsnitt, anledning att ställa sig tveksam till en helt ny generell reglering av AI av det slag som till exempel har föreslagits i Europaparlamentets etikresolution.

Om en sådan reglering ändå införs är det viktigt att *tillämpningsområdet* görs tydligt. Som framgått är själva fenomenet AI svårt att ringa in med precision och att avgränsa på ett tydligt sätt. Det är viktigt att den nya regleringen inte i praktiken blir tillämplig på all verksamhet som bedrivs med ett modernt it-stöd.

Som framhållits av flera aktörer bör en eventuell generell reglering – i enlighet med den riskbaserade ansatsen – inriktas på *högrisktillämpningar*. Som framhållits i de 14 medlemsstaternas non-paper är det i detta fall viktigt att identifieringen av vad som utgör hög risk görs på ett tydligt sätt och att regleringen inte i praktiken innebär att de flesta verksamheter ändå omfattas av de utökade materiella kraven. När syftet med regleringen i första hand är att motverka kränkningar av mänskliga rättigheter är risken stor att det blir konsekvensen.

Det är vidare viktigt att den materiella reglering som ska gälla för högrisktillämpningar får en *proportionerlig* utformning. Vid proportionalitetsbedömningen är det viktigt att analysera vad som är rimligt i förhållande till flera olika tillämpningar, aktörer med mera. De nya kraven får naturligtvis inte hamna i konflikt med befintliga krav i andra regleringar, men nya regler ska helst inte heller överlappa befintliga regler, eftersom detta riskerar att försvåra tillämpningen och minska reglernas effektivitet. Som ett exempel kan nämnas att nya krav på lagring av träningsdata kan skapa en spänning med dataskyddsregleringen och dess bakomliggande intressen.

www.svensktnaringsliv.se
Storgatan 19, 114 82 Stockholm
Telefon 08-553 430 00

Tryck: Arkitektkopia AB, Bromma, 2021