Checklista

Juridik vid användning av AI

ES2022-08

Juni 2022

Innehåll

1. Inledning

Användning av artificiell intelligens (AI) i en myndighets verksamhet är ett område där flera frågor uppkommer om vilka förutsättningar som föreligger. Begreppet "användning" används i generell betydelse i denna checklista och inbegriper införskaffande, upphandling, utveckling, testning, drift och förvaltning m.m.

Checklistan fokuserar på rättsliga frågeställningar. För ytterligare beskrivningar om definitioner, risker och nyttor, organisatoriska förmågor, tillämpningar m.m. hänvisas till eSams rapport om *Samverkan kring tillämpad AI*.[1]

Många gånger är de rättsliga frågorna desamma som vid annan verksamhetsutveckling. I eSams *Checklista för jurister Introduktion i rättsliga förutsättningar i utvecklingsinsatser*,[2] ges ett stöd i vilka rättsliga frågor som kan behöva ställas i en utvecklingsinsats.

Syftet med checklistan *Juridik vid användning av AI* är att belysa vilka rättsliga frågor som särskilt behöver beaktas vid användning av AI. Denna checklista ska alltså ses som ett komplement till Checklista för jurister.

Checklistan *Juridik vid användning av AI* vänder sig till jurister inom statliga myndigheter och regioner som i sitt arbete kommer i kontakt med användning av AI. Därutöver kan den också vara ett hjälpmedel för att skapa förståelse, även för en vidare krets, för vilka rättsliga frågor som uppkommer i samband med AI-användning.

Checklistan syftar inte till att besvara de olika juridiska frågeställningar som kan vara aktuella. Den syftar heller inte till att redovisa en uttömmande förteckning över alla relevanta frågeställningar, författningar eller styrande dokument. Varje myndighet behöver vara uppmärksam på att det kan finnas specialförfattningar som kan påverka förutsättningarna i varje utvecklingsinsats.

Checklistan omfattar inte rättsliga frågor om pseudonymisering i samband med AI-användning, se istället i eSams vägledning om pseudonymisering av personuppgifter.[3] Checklistan avser inte heller att på djupet behandla frågor om etiska[4] och moraliska perspektiv, vilket är frågor som också måste beaktas vid en AI-användning. Frågor om mänskliga rättigheter berörs delvis i checklistan.

[1] ES2022-03 Rapport Samverkan kring tillämpad AI, juni 2022.
[2] Checklista för jurister Introduktion i rättsliga förutsättningar i utvecklingsinsatser, version 2.0, eSam juni 2019.
[3] ES2022-01 Vägledning Pseudonymisering av personuppgifter, juni 2022.
[4] Detta ämne behandlas bl.a. i Ethical guidelines for thrustworthy AI, High-Level Expert Group on Articifial Intelligence.

Arbetet med att ta fram denna checklista har genomförts av en särskild arbetsgrupp bestående av Charlotta Aggevall, Christina Wikström, Joakim Lundegård, Jonas Öhrnell, Katarina Lind, Tina Chavoshi, Sofie Wildiér, Ulrika Hedberg, Hans-Peter Erlingsson och Linda Lindström. Kvalitetssäkring har skett i eSams rättsliga expertgrupp, expertgruppen i säkerhet samt koordineringsgruppen för arkitektur. Beredning har skett via eSams samordningsgrupp.

2. Artificiell intelligens (AI)

Det saknas en vedertagen definition av begreppet AI. Europaparlamentet definierar t.ex. AI enligt följande: "AI är i korthet en maskins förmåga att visa människoliknande drag, såsom resonerande, inlärning, planering och kreativitet. AI möjliggör för tekniska system att uppfatta sin omgivning, hantera vad de uppfattar och lösa problem, med syfte att uppnå ett specifikt mål. Datorn mottar information (redan förberedd eller insamlad genom sina egna sensorer, t.ex. via en kamera), behandlar den och svarar. AI-system är kapabla till att anpassa sitt beteende, till en viss grad, genom att analysera effekterna av tidigare åtgärder, och att arbeta självständigt."[5] I den AI-förordning som håller på att förhandlas inom EU, se avsnitt 3.1 nedan, föreslås i artikel 3.1 en definition av AI-system där större vikt läggs vid tekniken och vilka metoder som används. För ytterligare beskrivning av olika definitioner, se t.ex. eSams rapport om *Samverkan kring tillämpad AI*.

AI-system kan ha olika komplexitetsgrad (smal eller generell AI[6]), vilket kan påverka den juridiska bedömningen av t.ex. transparens och förklarbarhet. Vid användning och utveckling av AI är det vanligen fråga om AI som bygger på maskininlärning (ML), dvs. system som tränas till att identifiera statistiska samband i data (snarare än regler programmerade av människor). Systemet korrigerar och anpassar beräkningsalgoritmerna utifrån utfall och behov.[7]

AI kan t.ex. användas i webbsökningar, automatiska översättningar och smarta hem. Inom offentlig verksamhet används AI bl.a. i form av chattbotar för att ge service till allmänheten genom fråga-svar-funktioner, som maskningsverktyg, vid urval eller för att ta fram beslutsunderlag, samt som beslutsstöd vid ärendehandläggning eller faktiskt handlande.[8]

[5] Europaparlamentets definition, https://www.europarl.europa.eu/news/sv/headlines/priorities/artificiell-intelligens-i-eu/20200827STO85804/vad-ar-artificiell-intelligens-och-hur-anvands-det
[6] Artficial narrow intelligence (ANI), artificial general intelligence (AGI) och artificial super intelligence (ASI).
[7] Exempel på tekniker och metoder är dataigenkänning, kognitiv databehandling, neurala nätverk, djupinlärning, naturlig språkinlärning, prediktiv analys och stordata, se eSams rapport Samverkan kring tillämpad AI.
[8] SOU 2018:25 s 150.

3. Rättsområden som kan aktualiseras vid användning av AI

Generellt sett behövs vid användning av AI beaktas samma rättsområden som vid annan verksamhetsutveckling. Nedan redovisas i korthet några av de rättsområden som vanligen aktualiseras vid AI-användning, utifrån den struktur som används i Checklista för jurister. För en mer fullständig genomgång av de rättsområden och den lagstiftning som aktualiseras hänvisas till Checklista för jurister. I avsnitt 3.1 redovisas också det förslag till AI-förordning som är under förhandling.

3.1 AI-förordningen (förslag)

Inom EU förhandlas[9] förslaget till Europaparlamentets och rådets förordning om harmoniserade regler för artificiell intelligens (rättsakt om artificiell intelligens) och om ändring av vissa unionslagstiftningsakter (AI-förordningen).[10] Syftet med AI-förordningen är att harmonisera reglerna för AI inom EU.[11] Förordningen kommer att få betydelse för myndigheters AI-användning och behöver därmed tas i beaktande inför kommande AI-användning även om förordningen, vid denna checklistas utgivande, är under förhandling. Förslaget innehåller i huvudsak följande:

- Riskbaserat angreppsätt i syfte att skapa en strukturerad uppdelning mellan olika typer av AI-system och dess användning där vissa är förbjudna, medan andra är tillåtna men med restriktioner och krav i form av bl.a. tillsyn och registrering hos ansvarig myndighet.
- Definition av AI-system.
- Förbud mot viss användning av AI-system.
- Granskningsprocess och CE-certifiering av högrisksystem. Krav på att upprätthålla adekvat nivå av transparens i hur AI-systemet fungerar, tillhandahålla relevant dokumentation av systemets funktion för tillsyn och säkerställa användning av data som är av hög kvalitet.
- Informations- och transparenskrav för hur AI-system som är avsedda att interagera med fysiska personer ska utformas och utvecklas så att fysiska personer blir informerade om att de interagerar med AI-system.
- Användning av regulatoriska sandlådor.

[9] Regeringskansliet Faktapromemoria 2020/21:FPM109 Förordning om artificiell intelligens.
[10] Förslag till Europarlamentets och rådets förordning om harmoniserade regler för artificiell intelligens (rättsakt om artificiell intelligens) och om ändring av vissa unionslagstiftningar,
Bryssel den 21.4.2021, COM (2021) 206 final.
[11] Enligt förslaget kommer AI-förordningen inte att gälla vid åtgärder som syftar till att skydda nationell säkerhet.

- Rapporteringsskyldigheter för tillhandahållare när det gäller utredning av AI-relaterade incidenter och tekniska problem.
- Sanktioner.

3.2 Kompetensområde

Vid en AI-utveckling aktualiseras regeringsformens bestämmelser. Utvecklingen och resultatet måste överensstämma med grundlagskrav om allas likhet inför lagen, saklighet och opartiskhet samt får inte medföra ett betydande intrång i den personliga integriteten.[12] Vidare får en myndighet enligt legalitetsprincipen[13] endast vidta åtgärder som har stöd i rättsordningen. En myndighet måste således innan den inleder en AI-utveckling ha stöd för utvecklingen i myndighetens uppdrag och myndighetsspecifik lagstiftning. Proportionalitetsprincipen[14] måste beaktas.

3.3 God offentlighetsstruktur

En god ordning i hanteringen av myndigheternas information (handlingar) är av betydelse för den enskildes möjlighet att ta del av allmänna handlingar och möjligheten att bygga upp och bevara allmänna handlingar som en del av det nationella kulturarvet. Bestämmelserna i tryckfrihetsförordningen (1949:105), offentlighets- och sekretesslagen (2009:400), arkivlagen (1990:782) m.fl. måste beaktas. Vid en AI-utveckling behöver alltså identifieras om det uppstår handlingar och om dessa är allmänna samt se till diarieföring och registrering görs när detta krävs. Vidare behöver krav på bevarande och gallring säkerställas, liksom hantering av uppgifter som omfattas av sekretess. Exempelvis behöver det bedömas om tränings- och valideringsdata blir nya allmänna handlingar eller om de utgör kopior av tidigare allmänna handlingar. Källkoden kan i sig vara en allmän handling.[15] Förändringar av källkod kan innebära att det genereras nya allmänna handlingar och att det anses ske en gallring av tidigare version. Med AI kan nya sammanställningar göras utifrån stora mängder information. Information som var för sig inte omfattas av sekretess kan sammanställt komma att omfattas av sekretess. Det kan också finnas en risk att de sekretessregler som normalt sett är tillämpliga i verksamheten inte kan tillämpas i förhållande till de handlingar som upprättas inom ramen för utvecklingen.

Upphovsrättsliga bestämmelser samt bestämmelser om vidareutnyttjande och öppna data kan aktualiseras, såsom lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk och lagen (2010:566) om vidareutnyttjande av handlingar från den offentliga

[12] 1 kap. 9 § och 2 kap. 6 § regeringsformen (1974:152).
[13] 1 kap. 1 § tredje stycket regeringsformen och 5 § förvaltningslagen (2017:900).
[14] 2 kap. 21 § regeringsformen och 5 § förvaltningslagen.
[15] Kammarrättens i Stockholm dom i mål nr 3692-19.

förvaltningen.[16] AI-utveckling i form av datorprogram omfattas normalt av upphovsrätt. En AI-utveckling resulterar också ofta i en produkt som omfattas av ett vidare immaterialrättsligt skydd. Det kan också finnas upphovsrättsliga hinder för att använda AI på informationsmängder av olika slag, t.ex. kan i vissa fall användningen begränsas utifrån anställdas upphovsrätt. Det behöver klargöras vem som har äganderätt till produkten och vilken nyttjanderätt som föreligger. Vidare behöver klargöras om produkten får vidareutvecklas eller överlåtas och vilka förutsättningar som föreligger för vidareutnyttjande av data.

3.4 Personuppgifter

Om AI-utvecklingen innebär att personuppgifter behandlas, aktualiseras bestämmelserna i dataskyddsförordningen,[17] brottsdatalagen (2018:1177), dataskyddslagen (2018:218) med tillhörande förordning samt myndighetsspecifika registerförfattningar. Det måste finnas en rättslig grund för behandlingen och uppgifterna får endast behandlas för särskilda, uttryckligt angivna och berättigade ändamål. Detta gäller även om personuppgifter behandlas vid testverksamhet och det kan då bli fråga om en tolkning av utrymmet enligt finalitetsprincipen.[18] Testning är en typ av behandling, men anses i vissa fall även vara ett eget ändamål.[19] Det senare torde främst bli aktuellt då testning inte sker inom ursprungsverksamheten.[20] Vissa registerförfattningar har testverksamhet angivet som ett särskilt ändamål. eSams uppfattning är att testning som huvudregel är en typ av behandling och inte ett eget ändamål, se eSams promemoria *En modern registerförfattning*.[21]

Om en behandling av personuppgifter sannolikt leder till en hög risk för fysiska personers rättigheter och friheter ska det före behandlingen göras en bedömning av den planerade behandlingens konsekvenser för skyddet av personuppgifter.[22] Myndigheten behöver också säkerställa de registrerades rättigheter, såsom information till den registrerade, rättelse och radering. Det finns krav på dokumentation för att möjliggöra transparens. Det ska även vidtas lämpliga tekniska och organisatoriska åtgärder.[23] Sannolikt kommer det många gånger vid AI-användning finnas sådana risker som föranleder att det behöver genomföras en konsekvensbedömning.

---

[16] Regeringen föreslog den 7 april 2022 en ny lag om den offentliga sektorns tillgängliggörande av data (prop. 2021/22:225). Lagen ska genomföra Europarlamentets och rådets direktiv öppna data och vidareutnyttjande av information från den offentliga sektorn. Lagändringarna kommer träda i kraft den 1 augusti 2022. Genom ikraftträdandet av den nya lagen upphävs den äldre lagen.
[17] 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).
[18] Artikel 5.1.b och artikel 6.4 dataskyddsförordningen.
[19] Datainspektionen (nu Integritetsskyddsmyndigheten) har gjort en skillnad på testning för att säkerställa befintliga uppgifter mot testning för utveckling av nya system (jfr Datainspektionens beslut den 15 september 2014, diarienr. 1275-2013).
[20] I prop. 2019/20:113 s 19-20 anges: ”att testverksamhet inte behöver regleras som ett särskilt ändamål måste således anses vara en slags huvudprincip”.
[21] ES2022-06 Promemoria En modern registerförfattning, juni 2022.
[22] Artikel 35.1 dataskyddsförordningen.
[23] Artikel 32 dataskyddsförordningen.

Dataskyddsförordningens principer för behandling av personuppgifter måste beaktas, bl.a. principerna om uppgiftsminimering och lagringsminimering, liksom kravet på korrekthet.[24] Eftersom AI-system ofta behöver en större mängd personuppgifter för att kunna finna tillförlitliga samband, måste dessa principer även ställas mot krav som objektivitet och likabehandling. Ibland kan det uppstå en konflikt mellan att å ena sidan minimera uppgiftsmängder och å andra sidan behöva en större uppgiftsmängd för ett korrekt utfall.

Vad gäller profilering och automatiserade beslut, se avsnitt 4.2.

3.5 Service, tillgänglighet och ärendehandläggning

Bestämmelserna om service och tillgänglighet i regeringsformen, förvaltningslagen och myndighetsförordningen m.m. gäller såväl vid ärendehandläggning som vid annan förvaltningsverksamhet. Om AI används inom ärendehandläggning gäller samma krav på kommunicering, partsinsyn och dokumentation av beslutsunderlag samt beslutsmotivering som vid ordinarie handläggning. Det kan finnas särskild anledning att se över frågor om transparens och automatiserade beslut, se avsnitten 4.1 och 4.2.

Krav på tillgänglighet enligt lagen (2018:1937) om tillgänglighet i digital offentlig service, språklagen (2009:600) m.m. behöver beaktas vid AI-användning, t.ex. om det avser en rösttjänst eller en chattbot. Frågor kring diskriminering beskrivs särskilt i avsnitt 4.3.

3.6 Informationssäkerhet

Informationssäkerhet handlar om ansvar, riskmedvetenhet och helhetssyn. Inför en AI-användning behöver det göras en informationsklassning, en hot- och riskanalys samt övervägas om det behövs en konsekvensbedömning enligt dataskyddsregleringen. Bestämmelser om krisberedskap behöver beaktas. Det behöver finnas en hantering för personuppgiftsincidenter och upptäckt av dessa. Det behöver även bedömas om säkerhetsskyddslagens bestämmelser kan bli tillämpliga och utifrån det avgöra om det är lämpligt med en AI-tillämpning vid sådana uppgifter.

3.7 Upphandling och konkurrensfrågor

Många gånger föreligger liknande förutsättningar vid införskaffande av AI som vid anskaffning av annan programvara. eSam har tagit fram olika it-villkor som kan ge stöd i avtalsutformning vid införskaffande av AI.

[24] Artikel 5 dataskyddsförordningen.

Vid införskaffande av teknik, system eller lösningar med AI kan det bli fråga om en mer rörlig leverans än vid traditionell införskaffning. Det kan därmed behöva övervägas hur lämpliga krav kan ställas på garantier att AI-modellen är utvecklad med hänsyn till etik samt med objektivitet och saklighet som grund vid träning, test och verifiering. Dessutom kan det vara bra att i avtalet reglera ansvaret för AI-specifika fel och lämpliga påföljder vid avtalsbrott. Det är också viktigt att reglera granskningskriterier för att kunna värdera leverantörens AI-teknik och bedöma denna utifrån de egna kraven och kontraktsuppfyllnad.

Vid anskaffning av AI-programvara behöver den upphandlande myndigheten även se till att den förädlade AI-modellen ingår i användarlicensen alternativt tillfaller myndigheten med äganderätt. Ingår träningsdata i leveransen bör även denna materia regleras i avtalet, även om data i sig inte är en immaterialrättsligt skyddad tillgång. Dessutom kan data utgöra s.k. nödvändiga nyttigheter (essential facilities) som enligt konkurrensrätten kan ge en monopolliknande ställning för innehavaren.

Service-level agreement (SLA) kan behöva utformas på annat sätt vid AI-utveckling. Till skillnad från traditionella Service-level agreement (SLA:er) som främst riktar in sig på att upprätthålla en förbestämd nivå på tillgänglighet och pålitlighet, kan AI-specifika SLA:er fokusera på effekterna, exempelvis i form av effektivitets- eller kvalitetsmål.

Det är inte ovanligt att stora resurser och tid har investerats i träning av AI-modellen, som riskerar att gå förlorade vid avtalets upphörande. Det behöver därmed beaktas vid avtalets ingång vad som är en lämplig reglering vid exit.

3.8 Drift och förvaltning

Användning av AI kan innebära att ansvarsskyldigheter uppstår inom nya områden. Ansvaret kan avse hantering av risk för kränkningar av grundläggande fri- och rättigheter, skydd av personuppgifter, hantering av risk för diskriminering, konsumentskydd, produktsäkerhet och produktansvar, ansvar för person- eller sakskador, se avsnitt 4.4.

4. Några särskilt intressanta rättsfrågor

4.1 Transparens och förklarbarhet

Användningen av AI med dess specifika egenskaper (t.ex. bristande insyn, komplexitet, beroende av data och autonomt beteende) kan komma att inverka negativt på ett antal grundläggande rättigheter som bl.a. fastställs i Europeiska unionens stadgan om de grundläggande rättigheterna och Europakonventionen.[25] Dagens tekniska möjligheter att analysera stora datamängder med hjälp av artificiell intelligens och maskininlärning har gjort det enklare att göra profileringar och fatta automatiserade beslut. Det kan potentiellt få stora konsekvenser för enskildas rättigheter och friheter. För att säkerställa möjligheten till insyn och för att motverka kränkningar av mänskliga rättigheter eller andra negativa utfall är det viktigt med krav på transparens och öppenhet när AI-system används. Insyn och transparens bygger ett förtroende för AI-användning.

Vid AI-användning behöver det säkerställas att myndigheten klart och tydligt kan redogöra på vilka grunder ett beslut har fattats enligt förvaltningslagen.[26] Detta innebär att den enskilde som beslutet avser ska ges möjlighet att förstå hur myndigheten har resonerat i det enskilda fallet,[27] vilka omständigheter som myndigheten har tillmätt betydelse och hur myndigheten bedömt eventuella invändningar från den enskilde. Det måste också, när lagstiftningen kräver, ges tillfälle för den enskilde att yttra sig innan beslutet fattas.[28] Vid automatiserade beslut saknas av naturliga skäl delar av den information som ska dokumenteras enligt förvaltningslagen och myndighetsförordningen (2007:515).[29] Av förarbetena till förvaltningslagen framgår att de i lagtexten angivna uppgifterna inte alltid måste anges utan endast i förekommande fall.[30] Observera att avvikande bestämmelser kan finnas i speciallagstiftning.

Det behöver också säkerställas att myndigheten behandlar personuppgifter på ett lagligt, korrekt och öppet sätt och ger tydlig information om den behandling av personuppgifer som sker enligt dataskyddsförordningen.[31] Utöver den grundläggande information som alltid ska lämnas till den registrerade när personuppgifter behandlas uppställs särskilda krav på information till den registrerade när en myndighet tillämpar automatiserat

[25] Europeiska konventionen den 4 november 1950 angående skydd för de mänskliga rättigheterna och de grundläggande friheterna.
[26] Se regler om motiveringsskyldighet i 32 § i förvaltningslagen.
[27] Prop. 2016/17:180 s. 320.
[28] 25 § förvaltningslagen.
[29] 31 § förvaltningslagen och 21 § myndighetsförordningen.
[30] Prop. 2016/17:180 s. 185.
[31] Artiklarna 5.1a och 13 i datasskyddsförordningen.

beslutsfattande eller profilering.[32] Artikel 29-gruppen[33] har tagit fram riktlinjer, som ratificerats av Europeiska dataskyddsstyrelsen, om automatiserat individuellt beslutsfattande och profilering där ytterligare vägledning om hur bl.a. öppenhet bör tillämpas i fråga om profilering ges.[34] Enligt riktlinjerna rör det sig inte nödvändigtvis om en komplex förklaring av de algoritmer som används eller att lämna ut den fullständiga algoritmen. Den information som tillhandahålls bör emellertid vara tillräckligt heltäckande för att den registrerade ska förstå skälen till beslutet.

Vid AI-utveckling framhålls ofta en ökad risk för att myndigheten inte kan tillhandahålla information om beslutsmotiveringar och den personuppgiftsbehandling som sker avseende beslut som fattas av ett AI-system. Detta är ett problem som framförallt uppstår i samband med användning av så kallad djupinlärning. Vid djupinlärning använder man stora neurala nätverk som har flera lager och dessa är för människor svåra att studera och begripa när de har tränats klart. Det leder till brister i förklarbarhet, då det inte går att veta med säkerhet vilka mönster modellen använt för att nå sitt resultat. Det kan alltså vara svårt att ge en tillfredsställande motivering till ett beslut som tagits med hjälp av en sådan modell. Sådan karaktäristik kallas ofta för svarta lådor (black box).

Inför en AI-utveckling eller AI-användning i offentlig verksamhet bör särskilt övervägas vilken form och nivå av AI som bör tillämpas så att myndigheten kan säkerställa transparens[35] och insyn och upprätthålla medborgarnas förtroende. Ju mer avancerad AI, desto mer komplex blir frågan om transparens och förklarbarhet.

4.2 Automatiserade beslut

Förvaltningsmyndigheter får enligt 28 § förvaltningslagen använda automatiserat beslutsfattande.[36] Med automatiserat beslutsfattande avses beslut som fattas maskinellt utan att någon enskild befattningshavare på myndigheten tar aktiv del i själva beslutsfattandet i det enskilda fallet. AI och automatiserade beslut är ofta en möjlighet att effektivisera, förbättra eller förenkla handläggningen av ärenden, men det behöver beaktas att långtgående effektivitetsåtgärder riskerar att ske på bekostnad av både rättssäkerheten och legitimiteten. Om den data som AI-systemet tränas och programmeras med antingen är bristfällig eller innehåller skevheter (bias) finns det risk för diskriminerande, partiska eller i övrigt felaktiga beslut. En helt automatiserad process

[32] Artikel 22 och skäl 60 i dataskyddförordningen. Oavsett om behandlingen omfattas av bestämmelserna i artikel 22 måste det förtydligas för den registrerade att behandlingen avser både a) profilering och b) beslutsfattande grundat på den profil som skapats enligt artiklarna 13.1 c och 14.1 c.
[33] Artikel 29-gruppen är den oberoende europeiska arbetsgruppen som behandlade frågor om integritetsskydd och skydd av personuppgifter fram till den 25 maj 2018 (införande av den allmänna dataskyddsförordningen).
[34] Riktlinjer om automatiserat beslutsfattande och profilering enligt förordning (EU) 2016/679).
[35] "Decision tree" är ett exempel på en transparant väg till resultatet där man kan svara på ex. frågan varför fick jag inget lån från banken.
[36] Automatiserat beslutsfattande kan ske både med och utan AI. Bestämmelsen i 28 § förvaltningslagen gäller endast själva beslutsfattandet. I övrigt saknas bestämmelser i förvaltningslagen som tar sikte på automatiserade förfaranden.

förutsätter att samtliga omständigheter som kan bli aktuella programmeras i en algoritm, i annat fall finns det risk för att beslutet blir felaktigt.

I dataskyddsförordningen finns regler som begränsar möjligheten att använda AI för automatiserat beslutsfattande. Det finns flera bestämmelser enligt vilka den registrerade ska få information om förekomsten av automatiserat beslutsfattande, se avsnitt 4.1. Enligt artikel 22.1 i dataskyddsförordningen ska den registrerade även ha rätt att inte bli föremål för ett beslut som enbart grundas på automatiserad behandling, inbegripet profilering, vilket har rättsliga följder för honom eller henne eller på liknande sätt i betydande grad påverkar honom eller henne. Profilering definieras i art. 4.4 dataskyddsförordningen som automatisk behandling av personuppgifter för att bedöma vissa personliga egenskaper hos en person i synnerhet för att förutsäga bland annat pålitlighet. Automatiserat beslutsfattande hos myndigheter som innebär behandling av personuppgifter för att analysera eller förutse till exempel pålitlighet, beteende, arbetsprestation, hälsa eller dylikt kan därför vara otillåten. Det ska tilläggas att automatiserade beslut kan fattas med eller utan profilering. Vidare kan profilering ske utan att det fattas ett automatiserat beslut.

Det är oklart hur artikel 22.1 dataskyddsförordningen ska tolkas och därmed hur långt begränsningen att använda AI sträcker sig. Artikel 29-gruppen har i riktlinjer, som ratificerats av Europeiska dataskyddsstyrelsen, tolkat bestämmelsen som att det rör sig om ett *generellt förbud* mot automatiserat beslutsfattande och att det inte behövs någon invändning från en enskild.[37] Andra menar att en sådan tolkning inte stämmer överens med ordalydelsen i artikeln och systematiken i dataskyddsförordningen där det i andra artiklar i samma kapitel talas om *en rätt* som den registrerade har, men där det förutsätts en invändning från den registrerade. Vidare kan artikeln antingen tolkas som att förbudet träffar alla former av automatiserat beslutsfattande eller som att den endast gäller automatiserat beslutsfattande som inbegriper profilering. Det kan också nämnas att artikeln gäller beslut som *enbart* grundar sig på automatiserad behandling. Enligt artikel 29-gruppen är artikeln inte tillämplig om en människa med befogenhet att ändra beslutet granskar det på ett meningsfullt sätt. Beslutet grundas då inte uteslutande på automatiserad behandling.

I art. 22.1.b dataskyddsförordningen finns ett undantag som ger medlemsstaterna möjlighet att i nationell lag godkänna automatiserat beslutsfattande om den nationella lagen fastställer lämpliga åtgärder till skydd för den registrerades rättigheter, friheter och berättigade intressen. Regeringen anser att förvaltningslagen innehåller sådana

[37] Artikel 29-gruppen för uppgiftsskydd, Riktlinjer om automatiserat individuellt beslutsfattande och profilering enligt förordning (EU) 2016/679, godkända som fortsatt gällande av Europeiska dataskyddstyrelsen den 25 maj 2018.

skyddsåtgärder som avses i dataskyddsförordningen.[38] Även eSams bedömning är att de generellt tillämpliga bestämmelserna i förvaltningslagen innefattar sådana lämpliga skyddsåtgärder för den registrerade att automatiserat beslutsfattande med stöd av 28 § förvaltningslagen är tillåtet.[39]

4.3 Diskriminering

Diskrimineringsombudsmannen (DO) och flera europeiska likabehandlingsorgan har uppmärksammat förekomsten av AI och automatiserat beslutsfattande samt de risker för diskriminering som kan uppstå när algoritmer används vid databearbetning och beslutsfattande.[40]

Risk för diskriminering vid AI-användning kan uppstå t.ex. på grund av bristande kvalitet i data, bristande träning av AI samt bristande transparens och svarta lådor. Brister i data kan utgöra en risk för diskriminering genom att nödvändiga data saknas eller att data som används inte återspeglar faktiska förhållanden utan har påverkats av skevheter (bias) redan vid insamling. Data kan återspegla historisk diskriminering och därmed reproducera den. Data kan också tolkas på ett sätt som leder till risk för diskriminering. Brister i träning kan föreligga när inmatade data innehåller skevheter (bias) och inlärningsalgoritmen därmed tränar en modell som diskriminerar. Brist på transparens innebär en risk för att diskriminering inte upptäcks.[41]

Vid en AI-användning behöver de olika formerna[42] för diskriminering beaktas och användningen av AI bedömas utifrån samtliga diskrimineringsgrunder i diskrimineringslagen (2008:567).[43] Det behöver också genom uppföljning säkerställas att beslut är korrekta utifrån ett diskrimineringsperspektiv.

4.4 Ansvar

Användningen av AI är förknippad med både möjligheter samt risker och det finns anledning att överväga vilka ansvarsskyldigheter som föreligger. Ansvaret kan handla om hantering av risk för kränkningar av grundläggande fri- och rättigheter (inklusive skydd av personuppgifter, integritet och icke-diskriminering), men även reglering av

[38] Se bl.a. prop. 2017/18:95 s. 100, prop. 2017/18:112 s. 64-65, prop. 2017/18:115 s. 31, prop. 2017/18:254 s. 55 och prop. 2018/19:33 s. 164.
[39] eSam Rättsligt uttalande om automatiserade beslut 2018-03-19.
[40] Diskrimineringsombudsmannens rapport 2022:1, Transparens, träning och data, Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering.
[41] Diskrimineringsombudsmannens rapport 2022:1, Transparens, träning och data, Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering.
[42] 4 § diskrimineringslagen; Direkt diskriminering, indirekt diskriminering, bristande tillgänglighet, trakasserier, sexuella trakasserier, instruktioner att diskriminera.
[43] 1 § och 5 § diskrimineringslagen; Kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning, ålder.

konsumentskydd, produktsäkerhet och produktansvar (inklusive reglering av person- eller sakskador).

Användningen av AI kan riskera att kränka, eller åtminstone påverka, grundläggande fri- och rättigheter så som yttrande- och mötesfrihet eller rätten till mänsklig värdighet. Utökade möjligheter att analysera människors vanor skulle till exempel kunna användas för spårning och massövervakning av bland annat anställda eller medborgare.

Beroende på hur en AI-modell programmeras, tränas, utvecklas och används kan diskriminering uppkomma, se avsnitt 4.3. En överträdelse av diskrimineringsrätten kan medföra en ersättningsskyldighet gentemot den enskilde.

I dataskyddsförordningen har den personuppgiftsansvarige ett ansvar för och ska kunna visa att principerna för behandling av personuppgifter efterlevs. Dataskyddsförordningen är omfattande men särskilt bör regler kring automatiserat beslutsfattande, information som ska tillhandahållas och konsekvensbedömning beaktas. Överträdelser av dataskyddsförordningen kan medföra sanktioner, inklusive administrativa sanktionsavgifter och/eller att skadestånd behöver betalas till personer som lidit skada.[44]

AI-teknik kan innebära nya säkerhetsrisker för användare när den är inbäddad i produkter och tjänster. Om användandet av AI i produkter och tjänster leder till skador kan det bli fråga om ansvar enligt produktsäkerhetsreglering eller ansvarsreglering, två områden som kompletterar varandra men tar sikte på olika delar.

Det EU-rättsliga regelverket på produktsäkerhetsområdet består dels av direktivet om allmän produktsäkerhet,[45] dels av olika sektorsspecifika förordningar. I nationell lagstiftning syftar produktsäkerhetslagen till att säkerställa att varor och tjänster som tillhandahålls konsumenter inte orsakar skador på personer. Lagen tillämpas i fråga om varor och tjänster som tillhandahålls i näringsverksamhet och varor som tillhandahålls i offentlig verksamhet. En förutsättning för att lagen ska tillämpas är att varan eller tjänsten är avsedd för konsumenter eller kan antas komma att användas av konsumenter.

Skadeståndsansvaret för AI-relaterade skador bestäms i huvudsak av nationella regler. Den generella svenska regleringen av rätten till skadestånd finns i skadeståndslagen (1972:207), men även produktansvarslagen (1992:18) kan komma att aktualiseras vid en person- eller sakskada som en produkt har orsakat på grund av en säkerhetsbrist. Det finns även bestämmelser på EU-nivå att beakta, t.ex. regler om produktansvar för

[44] Art. 82 dataskyddsförordningen.
[45] Europaparlamentets och rådets direktiv 2001/95/EG av den 3 december 2001 om allmän produktsäkerhet. Kommissionen har lagt fram ett förslag till EU-förordning om allmän produktsäkerhet vars syfte är att ersätta direktivet.

skadevållande produkter, regler om ansvar för överträdelse av dataskyddsregler och regler om konkurrensrelaterade skador.[46]

I förslag till AI-förordning föreslås liknande sanktionsmöjligheter som i dataskyddsförordningen.

Det är viktigt att ansvarsförhållandena säkerställs och att det är tydligt vem som bär ansvaret vid en AI-användning.

4.5 Testverksamhet

Vid utveckling av AI finns det ett behov av att kunna utveckla algoritmer, träna modeller, optimera parametrar, validera kvalitet på data och modeller m.m. inom en utforskande aktivitet, inför ett beslut om eventuellt genomförande i verksamheten. Det vill säga det behöver utföras aktiviteter som sker i en s.k. testmiljö och inte i produktionsmiljö. Det kan också finnas ett behov av att testa en AI-produkt innan införskaffande.

För den juridiska bedömningen saknar det ofta betydelse om det är fråga om en behandling i testmiljö eller produktionsmiljö och samma rättsområden aktualiseras. Det som istället är avgörande är vilken informationsmängd som kommer att användas vid AI-utvecklingen, eftersom det påverkar förutsättningarna för den juridiska bedömningen och hur ingående denna behöver vara.

Är det fråga om behandling av personuppgifter måste det finnas en rättslig grund för behandlingen. Vidare behöver det bedömas om behandlingen är förenlig med de ursprungliga ändamål för vilka uppgifterna samlats in, se avsnitt 3.4. Det behöver också finnas en transparens i förhållande till den registrerade och beroende på risk kan även en konsekvensbedömning behöva genomföras. Detta gäller oaktat om det är fråga om ett tidigare stadium av utveckling av en algoritm eller i ett senare stadium vid träning av en modell. Det spelar inte heller någon roll om uppgifterna är pseudonymiserade.[47] Förekommer sekretessbelagda uppgifter behöver de hanteras på lämpligt sätt. Om det är möjligt att använda syntetisk data (fabricerad/påhittad data) eller anonymiserade uppgifter[48] föreligger inte samma rättsliga krav.

[46] Westman Daniel, Svenskt näringslivs rapport Vem tar ansvar för AI?, Mars 2021.
[47] Pseudonymiserade personuppgifter innebär fortfarande att dataskyddsregelverket måste tillämpas, se ES2022-01 Vägledning Pseudonymisering av personuppgifter.
[48] Anonymisering av personuppgifter inom en verksamhet kan vara svår att uppnå, se ES2022-01 Vägledning Pseudonymisering av personuppgifter.

I förslag till kommande AI-förordning finns bestämmelser för regulatoriska sandlådor[49] där utveckling kan ske och personuppgifter behandlas för andra syften än insamlingsändamålet, under vissa kontrollerade former. En sådan bestämmelse skulle ge vissa lättnader i förhållande till vad som annars behöver uppfyllas enligt dataskyddsförordningen. Bestämmelsen föreslås inte påverka nationella bestämmelser som utesluter behandling för andra syften.

[49] Regulatoriska sandlådor kan beskrivas som försöksverksamhet, begränsad i tid och rum, med syfte att utveckla teknik och regelverk under trygga former. Se t.ex. Rådets slutsatser om regulatoriska sandlådor och experimentklausuler som verktyg för ett innovationsvänligt, framtidssäkrat och motståndskraftigt regelverk som hanterar omvälvande utmaningar i en digital tidsålder, Bryssel den 16 november 2020, ST 13026/20.

5. Kompletterande checklistefrågor

Vid en rättslig bedömning av AI-användning bör utgångspunkt tas i frågorna i eSams Checklista för jurister. Nedan anges kompletterande frågor att ställa vid en AI-användning utifrån tidigare redovisade rättsområden i avsnitt 3 och 4. Beroende på användningsområde och informationsmängd aktualiseras frågorna i olika omfattning.

Behov av att använda AI
• Vilket behov ska användning av AI fylla? Finns rättsliga förutsättningar att behandla data för detta behov?
• Har datamängd och innehåll anpassats för den tänkta AI-användningen?
• Kan datat användas för AI-lösningens ändamål? Har data samlats in för andra syften. Finns det därmed risk att den inte innehåller alla faktorer som behöver beaktas i AI-lösningen?

Fri- och rättigheter, diskriminering m.m.
• Kommer AI:n ha en negativ inverkan på den enskildes grundläggande fri- och rättigheter?
• Kommer AI:n leda till diskriminerande resultat? Har diskrimieringsgrunderna beaktats?
• Används AI:n utan legitimt konkret syfte? Kommer AI:n kunna användas för spårning eller massövervakning?

Ansvar
• Är det fastställt vem som ansvarar för utveckling av AI:n?
• Vem har ansvaret för förvaltning av AI:n?
• Hur är ansvaret reglerat om AI:n bidrar till att felaktiga beslut fattas?
• Vem ansvarar för kvalitet i utfallet vid AI-användning?
• Vem är ansvarig vid skadeståndsanspråk?

Transparens
• Är det möjligt att ge information om hur träning och utvärdering av algoritmen gått till?
• Går det att förklara vad ett utfall eller resultat baseras på (utifrån s.k.black box-problematik)?
• Hur har risken för att maskininlärda algoritmer ”smittas” med felkällor eller felaktiga utgångspunkter (bias), med risk för såväl rättsosäkra som diskriminerande förfaranden, hanterats?
• Är det möjligt att ge information om hur maskininlärda algoritmer, som helt eller delvis påverkar utfallet vid automatiserade urval eller beslut, har kommit fram till sitt resultat?
• Hur kan algoritmen bevakas och följas upp? Hur ser tillsyn och granskningsprocess ut? Vilken systemdokumentation finns?
• Finns möjligheter till sökning, kontroll och sammanställning av uppgifter i allmänna handlingar? Uppstår nya handlingar?
• Finns det spårbarhet i tillämpningen av AI:n?

Ärendehandläggning och beslutsfattande
• Ska AI:n användas i någon form av ärendehandläggning som utmynnar i ett myndighetsbeslut?
• Kommer AI:n användas i någon form av ärendehandläggning som mynnar ut i ett beslutsstöd?
• Används AI:n vid ett automatiserat beslutsfattande?
• Inbegriper det automatiserade beslutsfattandet profilering?
• Har information lämnats till de registrerade om förekomsten av automatiserat beslutsfattande och/eller profilering?

- Kan myndigheten i beslutsunderlaget redovisa vilka omständigheter som den har tillmätt betydelse och hur den har värderat dessa?

Avtalsfrågor

- Har det ställts krav på att leverantören lämnar garantier att algoritmen eller AI-modellen utvecklats med hänsyn till etik samt med objektivitet och saklighet som grund vid träning, testning och verifiering?
- Är det i avtalet reglerat ansvar för AI-specifika fel (bias)?
- För att kunna granska leverantörens AI-teknik, finns granskningskriterier reglerade om hur själva granskningen ska utföras samt i vilka fall det kan vara motiverat att ta hjälp av utomstående specialister och hur dessa ska bekostas?
- Omfattas produkten av ett immaterialrättsligt skydd? Regleras detta i så fall i avtalet?
- Har det tydliggjorts vem som har äganderätt till produkten och vilken nyttjanderätt som föreligger?
- Har det klargjorts om produkten får vidareutvecklas eller överlåtas och vilka förutsättningar som föreligger för vidareutnyttjande av data?
- Har det vid avtalets ingång beaktats vad som är en lämplig reglering vid avtalets upphörande (exit), utifrån tillgång m.m.?

Behandling och uppföljning

- Går det att visa att AI-lösningen tränats med en tillräcklig mängd kvalitativ data?
- Är utfallet ändamålsenligt och rimligt utifrån i förväg definierade kvalitetsnormer?
- Finns möjlighet att bevaka och upptäcka om en angripare har funnit ett sätt att förmå AI-lösningen att generera fel resultat?
- Finns det en kontroll av generationer av träningsdata som ligger till grund för AI-lösningars funktionalitet?
- Hur hanteras data över tid? Finns stöd för återanvändning av datat?

- Finns det rutiner på plats för hur bristande kvalitet eller rent felaktiga utfall ska hanteras (t.ex. så att enskilda drabbas i så liten utsträckning som möjligt?)

Frågor vid ikraftträdande av AI-förordningen
• Är det säkerställt att det inte är fråga om en förbjuden användning?
• Vilken risknivå tillhör AI-systemet?
• Uppfylls kraven på granskningsprocess?
• Är det fråga om användning av regulatoriska sandlådor? Är kraven för sådan användning uppfyllda?

6. Vägledningar med mera

I detta avsnitt ges några exempel på förordningar, rapporter och annat material som kan vara relevant att ta del av vid användning av AI. Listan ska inte ses som en uttömmande uppräkning.

Europeiska kommissionens förslag till Europaparlamentets och Rådet förordning om harmoniserade regler för artificiell intelligens (rättsakt om artificiell intelligens) och om ändring av vissa unionslagstiftningsakter, COM (2021) 206, 21 april 2021.

Europeiska kommissionens vitbok om artificiell intelligens – en EU-strategi för spetskompetens och förtroende, COM (2020) 65, 19 februari 2020.

Europeiska kommissionens oberoende expertgrupp på hög nivå för AI-frågor publikation Etiska riktlinjer för tillförlitlig AI, 8 april 2019.

Regeringskansliets faktapromemoria Förordning om artificiell intelligens, 2020/21:FPM109, 26 maj 2021.

Svenskt näringslivs rapport Vem tar ansvar för AI? – gällande lagstiftning, framlagda reformförslag och övergripande analys, mars 2021.

Diskrimineringsombudsmannens rapport Transparens, träning och data – Myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering, 2022:1, 17 februari 2022.

Patent- och registreringsverkets tankepapper Artificiell intelligens och immaterialrätt, 27 maj 2021.

Riksrevisionens rapport om automatiserat beslutsfattande i statsförvaltningen, Skr 2020/21:88, 25 februari 2021.

eSam är ett medlemsdrivet program för samverkan mellan myndigheter för att underlätta och påskynda digitaliseringen inom det offentliga. eSam bildades 2015 som en frivillig fortsättning på E-delegationen. En viktig uppgift för eSam är att ta fram stöd och vägledningar som ger förutsättningar för att öka den digitala samverkan inom offentlig förvaltning.

Alla stöddokument finns på esamverka.se

I eSam ingår Arbetsförmedlingen, Bolagsverket, Boverket, Centrala Studiestödsnämnden, Domstolsverket, eHälso-myndigheten, Ekonomistyrningsverket, Folkhälsomyndigheten, Försäkringskassan, Havs- och vattenmyndigheten, Inspektionen för vård och omsorg, Jordbruksverket, Kriminalvården, Kronofogdemyndigheten, Lantmäteriet, Länsstyrelserna, Migrationsverket, Naturvårdsverket, Patent- och Registreringsverket, Pensionsmyndigheten, Polisen, Riksarkivet, Rättsmedicinalverket, Sida, Skatteverket, Skolverket, Statens institutionsstyrelse, Statens servicecenter, Statens tjänstepensionsverk, Statistiska centralbyrån, Tillväxtverket, Trafikverket, Transportstyrelsen, Tullverket och Universitets- och högskolerådet.