



Juridisk Publikation

STOCKHOLM - UPPSALA - LUND - GÖTEBORG - UMEÅ

Norea Normelli

Artificiell intelligens och rätten till en rättvis rättegång

Särtryck ur häfte 1/2021

Nummer 1/2021

ARTIFICIELL INTELLIGENS OCH RÄTTEN TILL EN RÄTTVIS RÄTTEGÅNG

Av Norea Normelli¹

Än så länge används artificiell intelligens (AI) inte i svenska domstolar, men det är möjligt att AI-system kommer att få en större roll där framöver. I denna artikel görs en ansats till att belysa hur framtida användning av AI i domstolarna skulle kunna både stärka och urholka rätten till en rättvis rättegång. Vidare diskuteras och analyseras hur befintliga regler kan begränsa när och hur AI får användas i domstolarna. I artikeln lyfts också ett behov av ytterligare regler på området. De bör utformas för att säkerställa att AI används för att stärka individens rätt till en rättvis rättegång, utan att kränka densamma.

I. INLEDNING

Redan i dag finns en lagstadgad möjlighet att använda AI i den svenska förvaltningen. Det framgår av 28 § förvaltningslagen (2017:900) att förvaltningsmyndigheter får använda automatiserat beslutsfattande. Möjligheten utnyttjas främst av myndigheter med ärendetäta verksamheter, såsom Försäkringskassan, för att effektivisera handläggningen.² AI har inte introducerats i de svenska domstolarna, men det finns exempel på länder som använder AI i rättsväsendet. I USA används till exempel algoritmer för att bedöma risken för återfall i brottslighet.³ Därutöver finns det AI-system som, även om tekniken inte används än, kan förutse hur Europadomstolen (ED) kommer att döma i enskilda fall.⁴ Med tanke på den ökande användningen av AI i allmänhet och det faktum att några länder redan introducerat AI i rättsväsendet, är det möjligt att en liknande utveckling kommer att ske i de svenska domstolarna framöver.

¹ Jur. kand., Uppsala universitet. Artikeln bygger huvudsakligen på en uppsats som skrevs inom ramen för en fördjupningskurs i konstitutionell rätt. Ett stort tack till expertgranskaren och redaktionen för att ni har läst och lämnat värdefulla synpunkter på tidigare utkast. Tack också tack till Henrietta Cahn för dina ständigt goda råd, Hedvig Keussen för kloka inspel och Agnes Liminga för konstruktiv kritik.

² Prop. 2016/17:180, *En modern och rättssäker förvaltning – ny förvaltningslag*, s. 179 f.

³ *Loomis mot Wisconsin* 881 N.W.2d 749, 2016, p. 13–15.

⁴ Se t.ex. Aletras, N., Dimitrios, T., Preotiuc-Pietro, D. & Lampos, V., *Predicting Judicial Decisions of the European Court of Human Rights: a Natural Language Processing Perspective*, PeerJ Computer Science, vol. 2:e93, 2016. Jfr även avsnitt 3. För ytterligare exempel på länder som använder AI i domstolarna, se <https://www.coe.int/en/web/cepej/practical-examples-of-ai-implemented-in-other-countries> (besökt 10 april 2021).

Det faktum att AI kan användas i domstolarna betyder dock inte att AI bör användas.⁵ Ett sätt att angripa frågan om lämplighet är att diskutera om och i så fall när AI kan komma att stärka eller riskera att kränka rätten till en rättvis rättegång. Om AI kan användas för att stärka den rättigheten kan det finnas skäl att använda AI för det ändamålet. Samtidigt får AI inte användas på ett sätt som kränker rätten till en rättvis rättegång. För att kunna utnyttja AI och samtidigt säkerställa ett effektivt skydd för rättigheten är det viktigt att redan nu identifiera riskområden och upprätta regler om hur och när AI får användas i de svenska domstolarna.

Syftet med denna artikel är att illustrera hur en framtida användning av AI i domstolarna kan påverka rätten till en rättvis rättegång, såsom den uttrycks i 2 kap. 11 § st. 2 regeringsformen (RF) och i art. 6.1⁶ i den europeiska konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna (EKMR), i positiv och negativ bemärkelse. Vidare ska det i artikeln diskuteras vilka regler som kan behövas för att säkerställa att användningen av AI inte riskerar att kränka denna rättighet. Möjligheterna och utmaningarna med att använda AI i domstolarna skiljer sig något åt beroende på om det rör sig om brottmål, tvistemål eller förvaltningsprocess, eftersom exempelvis partsställningen och bevisningen skiljer sig åt mellan de olika processtyperna. Artikeln belyser vissa aspekter som är relevanta för alla de tre processtyperna, men också aspekter som främst är framträdande inom ett särskilt processområde. På så sätt kan artikeln ligga till grund för fortsatta och fördjupade diskussioner inom ramen för de olika processtyperna.

Efter detta inledande avsnitt definieras begreppet AI i avsnitt 2. Därefter presenteras i avsnitt 3 exempel på hur AI kan användas i domstolarna. En diskussion om AI och rätten till en rättvis rättegång förs i avsnitt 4 och följs av ett avsnitt om reglering av AI i avsnitt 5. Det sistnämnda avsnittet diskuterar nuvarande regleringar som kan påverka när samt hur AI får användas i domstolarna och lyfter behovet av nya regler för att säkerställa att AI inte kränker rätten till en rättvis rättegång. Artikeln avslutas i avsnitt 6 med några kommentarer om behovet av att fortsätta utreda möjligheterna och utmaningarna med AI i rättsväsendet.

5 Jfr Bylander, E., *Den datoriserade domaren*, Svensk Juristtidning, 2019, s. 7. Bylander diskuterar när AI bör användas i domstolarna och lyfter vikten av att skilja på dömandet såsom konst och såsom teknik.

6 Anledningen till att artikeln är avgränsad till art. 6.1 EKMR är att art. 6.2–3 EKMR främst preciserar kraven i art. 6.1 EKMR.

2. BEGREPPET ARTIFICIELL INTELLIGENS

AI är ett samlingsbegrepp som kan definieras på olika sätt. I det följande används en definition från EU-kommissionen, då den kan sägas ge uttryck för en samsyn inom Europa när det gäller vad AI är:

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 (t.ex. röstassistenter, bildanalysprogram, sökmotorer, tal- och ansiktsgenkänningsystem), eller inbäddas i hårdvaruenheter (t.ex. avancerade robotar, självkörande bilar, drönare eller applikationer för sakernas internet).⁷

I princip all AI instrueras av algoritmer. En algoritm är en uppsättning regler eller instruktioner för att lösa ett problem.⁸ Med hjälp av maskininlärning kan AI lösa problem som inte kan definieras på förhand. AI-systemet delges vissa data (indata) som analyseras och tolkas. Efter att informationen bearbetats drar systemet en slutsats om vilket som är det bästa handlingsalternativet i den givna situationen och agerar i enlighet med det beslutet (utdata).⁹

3. EXEMPEL PÅ HUR AI KAN ANVÄNDAS I DOMSTOLARNA

AI kan både fatta egna beslut och assistera i beslutsprocesser. I teorin kan domstolarna därför använda AI för att stödja eller på sikt ersätta mänskligt beslutsfattande.¹⁰ Det finns AI-system som kan förutse domslut, även om de inte används i praktiken än. Två olika forskningsstudier visar till exempel att AI kunnat förutse domslut från ED i 79 % respektive 75 % av fallen.¹¹ En studie från USA visar att AI, förutom att kunna förutse domslut med 70,2 % säkerhet, också kunnat förutse de enskilda domarnas ställningstagande

7 Meddelande från kommissionen till Europaparlamentet, Europeiska rådet, Rådet, Europeiska ekonomiska och sociala kommittén och Regionkommittén: *Artificiell intelligens för Europa*, COM (2018) 237 final av den 25 april 2018, s. 1.

8 Gonfalonieri, A., *What is an AI Algorithm*, Medium, 22 april 2019, tillgänglig på <https://medium.com/predict/what-is-an-ai-algorithm-aceeab80e7e3> (besökt 20 januari 2021).

9 Europeiska kommissionens expertgrupp på hög nivå för artificiell intelligens, *A definition of Artificial Intelligence: main capabilities and disciplines*, av den 8 april 2019, s. 3 f.

10 Committee of Experts on Internet Intermediaries (MSI-NET), *Algorithms and human rights study on the human rights dimension of automated data processing techniques and possible regulatory implications*, 2017, s. 3.

11 Se Medvedeva, M., Vols, M. & Wieling, M., *Using machine learning to predict decisions of the European Court of Human Rights*, *Artificial Intelligence and Law*, 28(2), 2020, s. 263; Aletras m.fl., 2016, s. 2 f.

med 71,9 % säkerhet.¹² Forskarna bakom dessa studier framhåller att AI-systemen inte tagits fram för att ersätta mänskliga domare, utan för att assistera dem.¹³ Det är dock inte omöjligt att liknande system i framtiden kommer att utvecklas i syfte att ersätta mänskliga domare i vissa fall. I Estland utforskar justitieministeriet till exempel möjligheten att låta AI-system avgöra enklare civilrättsliga tvister om max 7 000 euro.¹⁴ En annan typ av mål som skulle kunna avgöras av AI i framtiden är mål om äktenskapsskillnad där båda parter är överens och det inte finns någon rättsligt komplicerad fråga.¹⁵

I USA har domstolarna använt algoritmer för att bedöma risken för återfall i brott.¹⁶ Just det verktyget har dock kritiserats i media för att det bedömer att svarta personer, i större utsträckning än vita personer, löper ”hög risk” att återfalla i brott.¹⁷ I England används ett liknande system av polisen.¹⁸ I Sverige får risken för upprepad brottslighet läggas till grund för valet av påföljd enligt 30 kap. 7 § brottsbalken. Det innebär att AI-system som bedömer risk kan påverka påföljdsbestämningen om de introduceras i svenska domstolar.

AI kan också användas för att samla in och sortera relevanta data snabbt och på ett mer effektivt sätt än människor kan.¹⁹ Det finns till exempel AI-system som används på advokatbyråer för att utföra vissa av de arbetsuppgifter en bolagsjurist ägnar sig åt. AI-systemen kan bland annat göra rättsutredningar, granska dokument och sammanfatta rättsfall.²⁰ Med tanke på de effektivitetsvinster som kan uppnås genom att använda AI är det tänkbart att

12 Katz, D.M., Bommarito II, M.J. & Blackman, J., *A general approach for predicting the behaviour of the Supreme Court of the United States*, PLoSOne, 12(4): e0174698, 2017, s. 8. Jfr även Bylander, 2019, s. 1.

13 Jfr Medvedeva m.fl., 2020, s. 263; Aletras m.fl., 2016, s. 3.

14 Niiler, E., *Can AI be a Fair Judge in Court? Estonia Thinks So*, Wired, 25 mars 2019, tillgänglig på <https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/> (besökt 1 april 2021).

15 Jfr Öster, U. & Ekholm, E., *Robotdomaren ännu långt bort*, Advokaten, nr. 4, 2019, årgång 85. Som rubriken indikerar bedöms det inte troligt att AI kommer ansvara för avgöranden i domstolarna inom en snar framtid, men exemplet med mål om äktenskapsskillnad lyfts som ett eventuellt framtidsscenario.

16 *Loomis mot Wisconsin*, p. 13–15.

17 Angwin, J., Larson, J., Mattu, S. & Kirchner, L., *Machine Bias*, ProPublica, 23 maj 2016, tillgänglig på <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (besökt 1 februari 2021).

18 Babuta, A., *Innocent Until Predicted Guilty? Artificial Intelligence and Police Decision-Making*, RUSI Newsbrief, 38(2), 2018, s. 1.

19 Dymitruk, M., *The Right to a Fair Trial in Automated Civil Proceedings*, Masaryk University Journal of Law and Technology, 13(1), 2019, s. 37.

20 Se t.ex. <https://www.bakermckenzie.com/en/newsroom/2020/02/tech-updates-in-australia-and-china> (besökt 9 april 2021) och <https://rossintelligence.com/case-studies> (besökt 12 februari 2021).

även domstolarna kan komma att använda liknande system framöver.²¹ AI kan även användas för att utföra administrativa uppgifter i domstolarna, som att boka salar och översätta dokument.²²

4. AI OCH RÄTTEN TILL EN RÄTTVIS RÄTTEGÅNG

4.1 ALLMÄNT OM RÄTTEN TILL EN RÄTTVIS RÄTTEGÅNG

Av 2 kap. 11 § st. 2 RF följer att en rättegång ska genomföras rättvist och inom skälig tid. Principerna för en rättvist genomförd rättegång preciseras i förarbetena.²³ Rätten till en rättvis rättegång stadgas även i art. 6 EKMR. ED menar att rätten utgör en hörnpelare i ett demokratiskt samhälle och att dess innebörd inte ska tolkas restriktivt.²⁴ Rätten får inte heller vara teoretisk eller illusorisk.²⁵

Genom lagen (1994:1219) om den europeiska konventionen angående skydd för de mänskliga rättigheterna och de grundläggande friheterna har EKMR status som lag i Sverige. Trots att EKMR inte har grundlagsstatus och att kravet på en rättvis rättegång kan härledas direkt ur grundläggande rättsstatliga principer i Sverige, har art. 6 EKMR haft en stor betydelse för svensk lagstiftning och praxis.²⁶ Av 2 kap. 19 § RF följer att svensk lag inte får meddelas i strid med EKMR. Därigenom säkerställs att EKMR får genomslag i Sverige såväl i samband med normgivning som vid rättstillämpning genom tolkning och utfyllnad av nationella rättsregler.²⁷ I utredningen inför grundlagsreformen 2010, när rätten till en rättvist genomförd rättegång skrevs in i regeringsformen, framhövdes det att art. 6 EKMR borde vara vägledande för utformningen av den svenska bestämmelsen.²⁸ Samtidigt påpekas i propositionen att art. 6 EKMR inte utan vidare ska påverka innehållet i svensk rätt.²⁹ Det kan konstateras att 2 kap. 11 § st. 2 RF inte kan ses isolerad från art. 6 EKMR, även om tillämpnings- och skyddsområdet som följer av

21 Jfr Europarådets kommission för effektivisering av rättsväsendet (CEPEJ), *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, av den 2–4 december 2018.

22 Jfr Öster, U. & Ekholm, 2019.

23 Prop. 2009/10:80, *En reformerad grundlag*, s. 160 och s. 252; SOU 2008:125, *En reformerad grundlag*, s. 426; Bylander, E., *Regeringsformens krav på alla rättegångars genomförande rättvist och inom skälig tid*, Svensk Juristtidning, 2017, s. 370 ff.

24 *Moreira de Azavedo mot Portugal*, nr. 11296/84, dom meddelad 23 oktober 1990, p. 66.

25 *Airey mot Island*, nr. 6289/73, dom meddelad 9 oktober 1979, p. 24.

26 Prop. 2009/10:80, s. 159; SOU 2008:125, s. 401 och s. 422.

27 Prop. 2009/10:80, s. 158.

28 SOU 2008:125, s. 425 f.

29 Prop. 2009/10:80, s. 159 f.

bestämmelserna i viss utsträckning skiljer sig åt.³⁰ Detta speglas även i svensk praxis. Generellt beaktar Högsta domstolen (HD) båda bestämmelserna i större eller mindre mån.³¹

Fortsättningsvis diskuteras hur användning av AI i domstolarna kan komma att påverka några av de rättssäkerhetskrav som följer av rätten till en rättvis rättegång enligt både 2 kap. 11 § st. 2 RF och art. 6.1 EKMR. De aspekter som kommer att behandlas är kravet på att rättegången ska genomföras inom skälig tid, den kontradiktoriska principen, principen om parternas likställdhet och kravet på en oavhängig och opartisk domstol.³² Genom att fokusera på dessa särskilt utvalda rättssäkerhetskrav tillåts en djupare diskussion om några möjligheter och utmaningar med att använda AI i domstolarna. Dessutom illustrerar urvalet möjligheter och utmaningar av sådan karaktär att de kan göra sig gällande även i relation till andra krav som ligger inom ramen för en rättvis rättegång. Eftersom artikeln utgår från ett hypotetiskt användande av AI kommer detaljskillnader i skyddet som följer av 2 kap. 11 § st. 2 RF och art. 6.1 EKMR inte att diskuteras.

4.2 RÄTTEGÅNGENS GENOMFÖRANDE INOM SKÄLIG TID

Ett krav på att rättegången ska genomföras inom skälig tid följer av ordalydelsen i både 2 kap. 11 § st. 2 RF och art. 6.1 EKMR.³³ Kravet syftar bland annat till att skydda parterna mot en alltför långsam handläggning och säkerställa att rättvisa tillhandahålls utan sådant dröjsmål som kan påverka rättssystemets effektivitet och trovärdighet negativt.³⁴ Möjligheten att använda AI för att sammanställa och undersöka praxis och doktrin skulle kunna ha en positiv inverkan på en rättegångs genomförande inom skälig tid, eftersom AI kan bearbeta information snabbare och mer effektivt än människor. Samma

30 För diskussion om relationen mellan 2 kap. 11 § st. 2 RF och art. 6 EKMR, se Bylander, 2017, s. 370 ff.

31 Se t.ex. NJA 2015 s. 222 där HD huvudsakligen hänvisar till art. 6 EKMR och praxis från ED, även om domstolen också hänvisar till 2 kap. 11 § RF inom parentes. I HFD 2017 ref. 53 använder domstolen praxis från ED som tolkningsdata för att bestämma innehållet i 2 kap. 11 § RF. I NJA 2017 s. 503 hänvisar HD parallellt till rättskällestödet i svensk rätt och i EKMR. I NJA 2017 s. 256 hänvisar HD för första gången till 2 kap. 11 § st. 2 RF utan att explicit hänvisa till EKMR.

32 Prop. 2009/10:80, s. 160 och s. 252; SOU 2008:125, s. 426; *Donadze mot Georgien*, nr. 74644/01, dom meddelad den 7 mars 2006, p. 35; *Kraska mot Schweiz*, nr. 13942/88, dom meddelad 19 april 1993, p. 30; *Regner mot Tjeckien [GC]*, nr. 35289/11, dom meddelad 19 september 2017, p. 146; *Ruiz-Mateos mot Spanien*, nr. 12952/87, dom meddelad 23 juni 1993, p. 63.

33 Av 2 kap. 11 § st. 2 RF följer att "en rättegång ska genomföras inom skälig tid". Av art. 6 EKMR följer att "var och en skall [...] vara berättigad till en rättvis och offentlig rättegång inom skälig tid".

34 SOU 2008:125, s. 423; NJA 2012 s. 1038; *Bottazzi mot Italien*, nr. 34884/97, dom meddelad den 28 juli 1999, p. 22.

effektivitetsresultat kan uppnås om AI används för att automatisera själva beslutsprocessen.³⁵ Viss effektivitetsvinst kan också uppnås genom att använda AI för att utföra administrativa uppgifter.³⁶

På grund av en ökad tillströmning av mål i de svenska domstolarna finns risk för förlängda handläggningstider.³⁷ Är det här ett problem som kan lösas med hjälp av AI? Är det till exempel möjligt att identifiera områden där AI kan användas för att minska arbetsbördan för domare, andra jurister och handläggare och skulle det då kunna leda till minskad risk för oskäligt långa handläggningstider? Det här är frågor som bör undersökas vidare framöver. För de fall AI kan användas för att effektivisera rättsprocesser och säkerställa att rättegångar genomförs inom skälig tid, kan argumenteras för att det finns skäl att använda AI för det ändamålet. Samtidigt bör det beaktas att överdrivna effektivitetsåtgärder kan ske på bekostnad av både rättssäkerheten och rättsväsendets legitimitet.³⁸ ED framhäver i detta avseende att viljan att effektivisera en process inte får ske på bekostnad av rätten till ett kontradiktoriskt förfarande.³⁹

4.3 DEN KONTRADIKTORISKA PRINCIPEN

4.3.1 PRINCIPENS INNEBÖRD

Den kontradiktoriska principen innebär att ingen får dömas ohörd och att båda parter ska få möjlighet att få ta del av den andres yrkanden, grunder och argument.⁴⁰ HD framhåller i NJA 2017 not. 7 att principen ökar möjligheten för materiellt riktiga domar. Till exempel kan en part vilja kommentera existensen av, innehållet i eller riktigheten hos de grunder och den bevisning som är avsedda att påverka ett domslut.⁴¹ Bristande möjlighet att ta del av bevisning kan kränka rätten till en rättvis rättegång om det hindrar en parts möjlighet att överklaga ett beslut eller påverka utgången i ett mål.⁴² Omfattningen av den kontradiktoriska principens innebörd kan dock variera

35 Dymitruk, 2019, s. 36 ff. Jfr även Lundqvist, U., *Artificiell Intelligens – rättsordning och rättstillämpning*, Svensk Juristtidning, 2020, s. 390.

36 Jfr Öster, U. & Ekholm, E., 2019.

37 Domstolsverket, pressmeddelande från Sveriges domstolar, 28 februari 2020, <https://www.domstol.se/nyheter/2020/02/allt-fler-mal-till-landets-domstolar/> (besökt 14 januari 2021).

38 Jfr avsnitt 4.3–4.5 om bristande transparens som ett hot mot förtroendet för rättssystemet.

39 *Nideröst-Huber mot Schweiz*, nr. 18990/91, dom meddelad 18 februari 1997, p. 30.

40 Ekelöf, P.O., Edelstam, H. & Heuman, L., *Rättegång fjärde häftet*, Norstedts Juridik, Stockholm, 2009, s. 30.

41 *Ruiz-Mateos mot Spanien*, p. 63; *Krčmář m.fl. mot Tjeckien*, nr. 35376/97, dom meddelad den 3 mars 2000, p. 42.

42 *McMichael mot Förenade Konungariket*, nr. 16424/90, dom meddelad den 24 februari 1995, p. 80.

beroende på omständigheterna i det enskilda fallet.⁴³ Även om principen inte uttrycks direkt i lag anses den ha stöd i art. 6.1 EKMR, och den utgör därmed en del av rätten till en rättvis rättegång.⁴⁴

Nedan återges det amerikanska rättsfallet *Loomis mot Wisconsin*, där domstolen använde ett algoritmiskt verktyg för att bedöma den tilltalades risk för att återfalla i brottslighet. Målet berörs i syfte att belysa när användning av AI i domstolarna kan resultera i att en process inte lever upp till de krav som den kontradiktoriska principen ställer på processen.

4.3.2 LOOMIS MOT WISCONSIN

När det gäller frågan om en parts möjlighet att få kännedom om och kommentera de grunder och den bevisning som är avsedda att ligga till grund för ett beslut eller en dom, finns det ett rättsfall från USA som är särskilt intressant att belysa. I målet *Loomis mot Wisconsin* prövades det om en tilltalad hade haft möjlighet att kommentera uppgifter som tagits fram av ett algoritmiskt verktyg. Klaganden i målet, Loomis, hade åtalats och dömts för brott i en domstol i Wisconsin. Domstolen hänvisade vid påföljdsbedömning bland annat till riskbedömningsverktyget ”Correctional Offender Management Profiling for Alternative Sanctions” (COMPAS), som bedömde att Loomis löpte ”hög risk” för att återfalla i brottslighet. Loomis åberopade i delstatens högsta instans att underrättens användning av COMPAS kränkte hans rätt till ”due process”⁴⁵ (jfr rättvis rättegång). Som grund för detta angav Loomis bland annat att han inte kunde bedöma korrektheten i informationen som COMPAS grundade sitt beslut på, då algoritmen som styr COMPAS är immaterialrättsligt skyddad och därför inte transparent.⁴⁶

Högsta domstolen i Wisconsin påpekade inledningsvis att en tilltalad har rätt att få se över, kontrollera och förklara sin inställning till den information som en domstol grundar sitt beslut på.⁴⁷ Domstolen fann däremot inte att Loomis rättighet hade kränkts. Den angav att Loomis kunde påverka hur han svarade på de frågor som algoritmen sedan använde som underlag för sitt beslut, samt att han hade möjlighet att korrigera eventuella felaktigheter i svaren på ett senare stadium. Dessutom menade domstolen att han kunde påverka underrättens bedömning genom att lägga fram ny bevisning som kunde motbevisa riskgraderingen.⁴⁸

43 *Hudáková m.fl. mot Slovakien*, nr. 23083/05, dom meddelad den 27 april 2010, p. 26–27.

44 Prop. 2009/10:80 s. 160; *Ruiz-Mateos mot Spanien*, p. 63.

45 Se den amerikanska konstitutionens femte tillägg och fjortonde tillägg sektion ett.

46 *Loomis mot Wisconsin*, p. 34.

47 *Loomis mot Wisconsin*, p. 47–48.

48 *Loomis mot Wisconsin*, p. 53–56.

4.3.3 ICKE-TRANSPARENT AI OCH DEN KONTRADIKTORISKA PRINCIPEN

Fallet *Loomis mot Wisconsin* illustrerar utmaningar med att använda AI för att bedöma risken för återfall i brottslighet, men det visar även den problematik som kan uppstå när AI används för andra ändamål, som att skapa bevisning eller assistera i beslutsprocesser. Låt oss anta att icke-transparent AI skulle användas för att ta fram sakkunnigbevisning i domstolen.⁴⁹ En part kan då berövas möjligheten att kommentera riktigheten av de uppgifter som ligger till grund för AI-systemets slutsats. I förlängningen kan det utgöra ett hinder mot att kommentera riktigheten av slutsatsen. Med andra ord finns det en risk för att en part inte kan kommentera det material som domstolen avser lägga till grund för sitt avgörande, vilket kan kränka rätten till en rättvis rättegång.⁵⁰

Hur vanligt är det då med icke-transparent AI? Förutom att källkoden kan vara immaterialrättsligt skyddad, kan maskininlärda algoritmer beskrivas som en "svart låda". Även om det går att få information om vilka indata ett AI-system delges och vilket utfallet av dess arbetsprocess blir kan det vara omöjligt att få tillgång till information om hur AI-systemet väger olika indata mot varandra.⁵¹ Det innebär att det kan vara svårt eller omöjligt, även för den som programmerat ett AI-system, att förstå de resonemang som ligger till grund för en slutsats systemet når. I förlängningen kan det göra det svårt eller omöjligt att bemöta eller kommentera riktigheten av de uppgifter som AI-systemet tar fram.

Användningen av AI för att ta fram bevisning som kan ligga till grund för ett beslut eller en dom kan jämföras med annan expertbevisning. En sorts expertbevisning är intyg som utfärdas av rättsmedicinskt sakkunniga. Det är möjligt att ett sådant intyg inte återger vilka ingångsdata som den sakkunnige i fråga haft tillgång till. Det finns också en risk för att den rättsmedicinska ingångsdata som legat till grund för intyget är missvisande, otillräcklig eller felaktig. Vidare kan olika omständigheter ha påverkat den sakkunniges bedömning, medvetet eller omedvetet, utan att det framkommer i intyget. Utmaningar med bristande transparens är alltså inget nytt fenomen som uppstått i och med utvecklingen av AI-system. Det som gör frågan om AI speciell är att medan det är möjligt att be den rättsmedicinskt sakkunnige att motivera slutsatserna i intyget, finns (än så länge) inte alltid samma möjlighet att få slutsatser från AI-system motiverade. Med andra ord kastar teknikens utveckling nytt ljus på frågan om transparens, och problematiken ställs på sin spets i fråga om AI.

49 Jfr avsnitt 4.4.

50 För fler överväganden om AI och bevisrätten, se Lundqvist, U., 2020, s. 391 ff.

51 Wisser, L., *Pandora's Algorithmic Black Box: The Challenges of Using Algorithmic Risk Assessment in Sentencing*, *American Criminal Law Review*, 56(4), 2019, s. 1816.

Det kan tänkas att HD och ED skulle framföra samma argument som högsta instansen i Wisconsin och finna att användningen av icke-transparent AI inte kränker rätten till en rättvis rättegång. Frågan är om de argumenten är hållbara. En möjlighet att ifrågasätta indata (svaren på intervjufrågorna) behöver nämligen inte per automatik medföra en möjlighet att bemöta och påverka utdata (gradering av risk). För att kunna bemöta riskgraderingen kan det hävdas att den tilltalade, förutom att kunna göra en kvalitetsbedömning av indata, också måste få information om hur algoritmen överväger och resonerar kring data för att nå en slutsats.⁵² Först när den tilltalade vet hur algoritmen väger information om exempelvis tidigare brottslighet mot andra fakta, såsom den tilltalades bostadsort, kan denne bemöta utdata genom att granska indata.⁵³ Detta resonemang bör även ses i ljuset av att juridisk argumentation karaktäriseras av möjligheten att nå olika slutsatser beroende på hur olika omständigheter vägs mot varandra.⁵⁴ Alltså kan det hävdas att användning av icke-transparent AI för att ta fram exempelvis sakkunnigutlåtanden trots allt riskerar att kränka rätten till en rättvis rättegång.

För det fall AI skulle användas för att förutse, föreslå eller meddela domar kan det vid första anblick tänkas att situationen är något annorlunda. Eventuella kommentarer om motpartens grunder och bevisning kan då delges AI-systemet som indata, så att det kan beakta kommentarerna i dess materiella bedömning. Det kan dock ifrågasättas om det faktum att en part kan kommentera motpartens grunder och bevisning så att de beaktas som indata skapar en reell eller endast illusorisk möjlighet att påverka utgången i målet.⁵⁵ Mot bakgrund av att HD i NJA 2017 not. 7 påpekar att det kontradiktoriska förfarandet bland annat främjar materiellt riktiga utgångar samt tilltron till det rättsliga förfarandet, och att ED i målet *Airey mot Island* framhäver vikten av att rätten till en rättvis rättegång inte blir illusorisk,⁵⁶ kan det hävdas att en part åtminstone måste få möjlighet att kontrollera viss mellanliggande informationsbearbetning för att efterlevnad av den kontradiktoriska principen ska säkerställas.

Om vi utgår från att AI främst kommer användas för att assistera i rättsprocesser och inte för att meddela domar på egen hand så blir frågan om när en part bör få möjlighet att kontrollera AI-systemets informationsbearbetning intressant. Bör möjligheten ges endast när domstolen beaktar AI-systemets

52 Washington, A., *How to Argue with an Algorithm: Lessons from the COMPAS ProPublica Debate*, *The Colorado Technology Law Journal*, 17(1), 2019, s. 134.

53 Washington, A., s. 146.

54 CEPEJ, 2018, s. 36 f.

55 Jfr Washington, A., s. 146.

56 *Airey mot Island*, p. 24.

slutsatser i praktiken? Eller bör en sådan möjlighet ges oberoende av domstolens inställning till AI-systemets slutsatser? I målet *McMichael mot Förenade Konungariket* konstaterade ED att rätten till en rättvis rättegång hade kränkts när föräldrar inte hade delgetts en rapport från socialtjänsten som hade legat till grund för ett beslut om vårdnad av barn. ED angav att föräldrarnas bristande möjlighet att få se ett för processen så viktigt dokument kunde påverka deras möjlighet att överklaga beslutet.⁵⁷ Det kan tala för att det värde domstolen tillskriver material som tas fram av ett AI-system kan påverka vilka krav som ställs på transparens. Samtidigt anger ED i andra fall att bevisningens faktiska effekt på domstolens bedömning är av mindre betydelse.⁵⁸ Att domstolen redovisar andra skäl som stöd för en dom, vid sidan av ett AI-systems slutsats, bör därför inte per automatik betyda att kraven enligt den kontradiktoriska principen efterlevs.

4.4 PRINCIPEN OM PARTERNAS LIKSTÄLLDHET

Ytterligare en aspekt av rätten till en rättvis rättegång är principen om parternas likställdhet.⁵⁹ Principen har ett nära samband med rätten till ett kontradiktoriskt förfarande.⁶⁰ HD anger i NJA 2013 s. 503 att principen innebär att en part inte får gynnas på den andres bekostnad. Samtidigt hänvisar HD till ED, som anger att staten inte är skyldig att garantera total likställdhet mellan parterna så länge de har en rimlig möjlighet att lägga fram sin sak under förhållanden som inte innebär en betydande nackdel i förhållande till motparten.⁶¹ I praktiken innebär det att parterna ska få en rimlig möjlighet att presentera sitt fall, inklusive bevisning, under sådana förhållanden som inte försätter en part i en sämre ställning än den andra parten.⁶²

Principen om parternas likställdhet kan bland annat anses överträdd om parterna har olika möjlighet att få tillgång till relevant information eller material.⁶³ Om en myndighet eller en domstol inte motiverar ett beslut eller en dom på ett sätt som ger en part en reell möjlighet att effektivt utnyttja en eventuell rätt att överklaga avgörandet kan det också kränka rätten till en rättvis

57 *McMichael mot Förenade Konungariket*, p. 80.

58 *Nideröst-Huber mot Schweiz*, p. 27; *Ziegler mot Schweiz*, nr. 33499/96, dom meddelad den 21 februari 2002, p. 38.

59 Prop. 2009/10:80 s. 160; *Regner mot Tjeckien*, p. 146.

60 *Regner mot Tjeckien*, p. 146.

61 *Handölsdalen Saami Village m.fl. mot Sverige*, nr. 39013/04, dom meddelad den 30 mars 2010, p. 51.

62 *Regner mot Tjeckien*, p. 146; *Dombo Beheer B.V mot Nederländerna*, nr. 14448/88, dom meddelad den 27 oktober 1993, p. 33.

63 Jfr *Yvon mot Frankrike*, nr. 44962/98, dom meddelad den 24 april 2003, p. 37.

rättegång.⁶⁴ Diskussionen om bristande transparens i avsnitt 4.3.3 är därmed aktuell även här. Till exempel kan domstolens möjlighet att motivera ett domslut påverkas om AI används för att skapa underlag till domen. Det kan nämligen vara svårt att skriva utvecklade domskäl om det inte går att få tillräcklig information om vilka data AI-systemet baserar sin slutsats på och hur systemet övervägt olika data för att nå slutsatsen. Det gör i sin tur att det kan bli svårt för en part att överklaga en dom när domarna beaktat slutsatser från ett AI-system. Vad gäller domstolarnas skyldighet att motivera avgöranden bör det dock påpekas att ED anger att skyldigheten inte omfattar ett krav på att avge detaljerade skäl för alla argument.⁶⁵ Domstolarna ska dock utreda parternas huvudargument.⁶⁶ Dessutom ska de grunder som rör grundläggande fri- och rättigheter i EKMR och dess tillhörande protokoll undersökas med särskild noggrannhet.⁶⁷

Huruvida icke-transparent AI hindrar domstolarna från att motivera avgöranden i tillräcklig grad beror därmed på omständigheterna i det enskilda fallet och torde ha samband med vilka frågor processen rör och hur stort inflytande ett AI-systems slutsatser får på beslutet eller domen. I detta avseende kan en ytterligare jämförelse göras med *Loomis*-målet. Högsta domstolen i Wisconsin beaktade nämligen att underrätten inte grundade sina domskäl *enbart* på riskbedömningsverktyget.⁶⁸ Det är möjligt att även HD eller ED skulle anse att det finns en reell möjlighet att överklaga och föra bevisning mot ett avgörande trots att det inte kan formuleras hur AI-systemet nått en slutsats som beaktas av domstolen, så länge domstolen också hänvisar till andra skäl för sitt beslut eller sin dom.

I relation till principen om parternas likställdhet lyfter en av Europarådets expertkommittéer, MSI-NET⁶⁹, risken för att algoritmer präglas av bias och fördomar.⁷⁰ MSI-NET utvecklar inte resonemanget, men eventuellt syftar expertkommittén på att användning av AI som förstärker fördomar kan medföra att en part diskrimineras och försätts i en sämre ställning än motparten. I kombination med bristande transparens i ett AI-system kan det dessutom

64 Jfr *Hentrich mot Frankrike*, nr. 13616/88, dom meddelad den 22 september 1994, p. 9, 15 och 56. Jfr även Lundqvist, U., 2020, s. 402 om domskrivning och den enskildes rättsskydd.

65 *Van de Hurk mot Nederländerna*, nr. 16034/90, dom meddelad den 19 april 1994, p. 61; *García Ruiz mot Spanien*, p. 26.

66 *Buzescu mot Rumänien*, nr. 61302/00, dom meddelad den 24 maj 2005, p. 67.

67 *Hirvisaari mot Finland*, nr. 49684/99, dom meddelad den 27 september 2001, p. 53.

68 *Loomis mot Wisconsin*, p. 68–74.

69 MSI-NET har bland annat arbetat med en rekommendation om mellanhänder på internet. Mellanhänder avser till exempel tjänsteleverantörer och sökmotorer.

70 MSI-NET, 2017, s. 11.

vara svårt att bevisa fördomens existens, vilket kan utgöra ytterligare ett hinder mot likställdhet.⁷¹

AI kan däremot också tänkas öka möjligheterna till likställdhet i processen. Katz m.fl. diskuterar i detta avseende möjligheten att använda AI som expertvittne i USA. De menar att AI kan söka och jämföra vetenskapliga data, utreda tillförlitligheten hos dessa och därefter komma med egna expertutlåtanden. Introduktion av AI i domstolarna på det här sättet anses kunna öka möjligheten till enhetliga och rättvisa beslut. Dessutom bedöms höga kostnader för parterna kunna undvikas vilket också kan öka likställdheten mellan dem.⁷²

På ett liknande sätt skulle AI kunna användas i stället för sakkunnigbevisning i svenska domstolar. En sakkunnig kan antingen kallas av parterna enligt 40 kap. 19 § rättegångsbalken (RB) eller förordnas av rätten i enlighet med 40 kap. 1 § RB. Tillkallandet av sakkunniga medför ofta höga kostnader. Det kan innebära att en ekonomiskt starkare part får större möjlighet att inhämta bevis för sin ståndpunkt. Det är dock tveksamt om en möjlighet för parterna att använda AI som sakkunnig skulle minska den eventuella maktobalansen, eftersom parterna själva måste stå för kostnaden att använda tekniken. Tillgång till AI-system för detta ändamål lär, åtminstone till en början, varken vara särskilt billig eller lättillgänglig.

Om domstolarna skulle få möjlighet att förordna användning av AI som sakkunnig skulle AI-systemet visserligen kunna utnyttjas till båda parternas fördel. Att domstolarna ges en sådan möjlighet innebär dock inte nödvändigtvis att parternas likställdhet i en rättsprocess ökar. För det första är huvudregeln i svensk rätt att parterna själva ansvarar för bevisningen enligt 30 kap. 6 § RB. Domstolarnas möjlighet att inhämta bevisning ex officio är därmed begränsad.⁷³ För det andra är domstolarnas möjlighet till materiell processledning också ytterst begränsad i svensk rätt.⁷⁴ Domstolarna kan visserligen föreslå bevisning så att parterna begär att rätten förordnar en sakkunnig.⁷⁵ På samma sätt skulle parterna, efter förslag från domstolen, då möjligen kunna begära att rätten förordnar att AI-system används som sakkunnig. Frågan om vilken rätt parterna måste ha att nyttja domstolens AI-system för att likställdheten mellan dem ska öka i praktiken kvarstår dock. För att garantera likställdhet mellan parterna i samtliga rättsprocesser kan

71 För en mer ingående diskussion om bias och fördomar, se avsnitt 4.5.

72 Katz, D.M. m.fl., 2017, s. 39 ff.

73 Ekelöf, P.O. m.fl., 2009, s. 55 f.

74 Ekelöf, P.O. & Edelstam, H., *Rättegång femte häftet*, Norstedts Juridik, Stockholm, 2011, s. 44 ff. och s. 200 ff.

75 Ekelöf, P.O. m.fl., 2009, s. 55 f., s. 287 f.

tänkas att det krävs en generell rätt att använda domstolarnas AI-system för sakkunnigbevisning.

4.5 KRAVET PÅ EN OAVHÄNGIG OCH OPARTISK DOMSTOL

Inom ramen för en rättvis rättegång ligger en rätt till en förhandling inför en oavhängig och opartisk domstol.⁷⁶ Dessutom följer av 1 kap. 9 § RF att domstolar ska iaktta saklighet och opartiskhet. Begreppet oavhängighet tar sikte på domstolens oberoende i förhållande till yttre påverkan från till exempel den verkställande makten, parterna i målet och lagstiftaren.⁷⁷ Kravet på opartiskhet innebär att varje domare ska döma opartiskt mellan parterna.⁷⁸ Frågan om opartiskhet bedöms utifrån ett subjektivt och ett objektiva test. Det subjektiva testet tar sikte på om en specifik domare de facto dömer opartiskt. Presumtionen är dock att domaren är opartisk tills motsatsen bevisats. Det objektiva testet är en bedömning av om en oberoende tredje part hade uppfattat domaren som opartisk eller inte.⁷⁹ En domare kan vara eller uppfattas som partisk av flera anledningar. Till exempel kan en domares fördomar och bias påverka domaren så att denne agerar partiskt.⁸⁰

Bias kan beskrivas som en systematisk snedvridning. Bias kan förekomma både hos människor och i data. Mänsklig bias kan vara medveten eller omedveten. Den kan till exempel grundas i okunskap, oförståelse, eller en tendens att söka efter och tolka information så att den bekräftar övertygelser en person har sedan tidigare.⁸¹ Mänsklig bias kan bland annat resultera i partiska beslut som ger en felaktig fördel eller nackdel till en part. Bias i data kan uppstå när den indata ett AI-system delges inte är representativ för den mänskliga befolkningen, vilket kan leda till att AI-systemet förstärker fördomar och missgynnar vissa grupper eller personer.⁸² En viktig fråga är hur risken för bias i data kan påverka domstolarnas oavhängighet och opartiskhet om AI-system skulle komma att användas i rättsprocesser. Mot bakgrund av risken för mänsklig bias är en annan relevant fråga om AI, trots risken för bias i data, kan

76 Art. 6.1 EKMR; Prop. 2009/10:80, s. 252.

77 *Beaumont mot Frankrike*, nr. 15287/89, dom meddelad den 24 november 1994, p. 38; *Sramek mot Österrike*, nr. 8790/79, dom meddelad den 22 oktober 1984, p. 42.

78 *Micallef mot Malta [GC]*, nr. 17056/06, dom meddelad den 15 oktober 2009, p. 93.

79 *Micallef mot Malta*, p. 93–96.

80 *Wettstein mot Schweiz*, nr. 33958/96, dom meddelad den 21 december 2000, p. 43.

81 Trevisky, A. & Kahneman, K., *Judgement Under Uncertainty: Heuristics and Biases*, Science, New Series, Vol. 185, No. 4157 (September 27, 1974), s. 1124–1131.

82 Ntoutsis, E., Fafalios, P., Gadiraju, U., m.fl., *Bias in Data-driven Artificial Intelligence Systems – An Introduction Survey*, Wiley Interdisciplinäre Reviews. Data Mining and Knowledge Discovery, 10(3), 2020, 10:e1356, s. 2.

användas för att minska risken för partiska beslut i domstol genom att minska risken för att mänsklig bias hos mänskliga domare får genomslag.

Till att börja med kan konstateras att AI programmeras och skapas av människor, varför AI oundvikligen kommer att spegla skaparens värderingar och fördomar. Om den data som ett AI-system delges inte är neutral och representativ för samhället, kommer AI-systemets slutsats också att präglas av bias.⁸³ Det är visserligen möjligt att programmera ett AI-system så att det inte beaktar vissa faktorer som kön, religion och etnisk tillhörighet, men sådana åtgärder verkar inte alltid kunna garantera att systemet inte når slutsatser som är diskriminerande.⁸⁴

Möjligen kan inte ens en enkel, än mindre en maskininlärd, algoritm vara helt neutral. Det kan då ifrågasättas om det alls är möjligt att garantera domstolens oavhängighet och opartiskhet om beslut eller domar skulle grundas på slutsatser från ett AI-system. Till exempel kan en domare som förlitar sig på ett AI-system kritiseras för att vara partisk om den indata som AI-systemet fått tillgång till präglas av bias. Beroende på vad AI skulle användas till kan det faktum att AI programmeras och skapas av människor som inte är domare möjligtvis också innebära ett hot mot domstolens oavhängighet, eftersom det indirekt kan öppna upp för en yttre påverkan på domstolen. Risk för AI-manipulerad information torde också kunna innebära ett hot mot domstolens oavhängighet.⁸⁵

Samtidigt kanske risken för bias i data inte ska överdrivas.⁸⁶ AI kan endast reproducera fördomar som finns i samhället, eftersom det är därifrån AI-systemet får sitt dataunderlag.⁸⁷ Det innebär att mänskliga domare kan ställas inför utmaningar som liknar de som gäller för AI. Människor har, liksom AI, en tendens att präglas av bias och fördomar.⁸⁸ Om mångfalden hos domarkåren är bristande kan AI till exempel tänkas användas för att minska risken för en partisk domstol, genom att minska risken för mänsklig bias.⁸⁹ Dessutom är

83 Gerards, J., 2019, s. 206; Allen, R. & Masters, D. *Artificial Intelligence: The Right to Protection from Discrimination Caused by Algorithms, Machine Learning and Automated Decision-Making*, ERA Forum, 20(4), 2020, s. 589 f.

84 Allen, R. & Masters, D., 2020, s. 589 f.

85 Se Lundqvist, U., 2020, s. 390 och s. 398 för fler resonemang om domstolens oavhängighet och opartiskhet samt risken för AI-manipulerad information.

86 Jfr Skeem, J.L. & Lowenkamp, C.T., *Risk, Race and Recidivism: Predictive Bias and Disparate Impact: Risk, Race, and Recidivism*, Criminology, 54(4), 2016, s. 37 f.

87 Mayson, S.G., *Bias in, Bias Out*, Yale Law Journal, 128(8), 2019, s. 2251.

88 Se Lidén, M., *Confirmation Bias in Criminal Cases*, Uppsala: Department of Law, Uppsala universitet, 2018, s. 215 ff.

89 Završnik, A., *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, ERA Forum, 24(4), 2020, s. 580.

AI enhetlig i sina fördomar och påverkas, till skillnad från människor, inte av andra yttre faktorer än den data som AI-systemet delges.⁹⁰ Det kan innebära att det är lättare att identifiera bias i utdata från ett AI-system än vad det är att identifiera mänsklig bias. Så länge en människa granskar slutsatser från ett AI-system kan det tänkas att människan kan uppmärksamma bias i data och säkerställa att fördomar inte reproduceras om AI används i domstolarna. I detta avseende finns dock återigen anledning att nämna något om problematiken med bristande transparens. Utmaningen är att bristande transparens i AI-system kan göra det svårt för en domare att avgöra om systemets slutsats baseras på diskriminerande grunder eller inte, även om AI-systemet är enhetligt i sina fördomar.⁹¹ Det bör också beaktas att det finns en risk för att domare på ett okritiskt sätt baserar beslut på ett AI-system som präglas av bias och att det kan vara svårt att bedöma om så är fallet eller inte, varför domaren kan uppfattas som partisk.⁹²

Risken för att indata och utdata i AI-system präglas av bias bör onekligen beaktas. Frågan är när bias i data, eller risken för bias i data, innebär att användning av AI i domstolarna skulle vara oförenligt med rätten att bli hörd inför en oavhängig och opartisk domstol. Detta är en fråga som bör utredas närmare utifrån de olika potentiella ändamål som AI skulle kunna användas för i domstolarna.

5. REGLERINGAR AV NÄR OCH HUR AI FÅR ANVÄNDAS

5.1 JURIDIKEN SOM VERKTYG FÖR ATT HINDRA RÄTTIGHETSKRÄNKNINGAR

Låt säga att det identifieras områden där AI kan användas för att stärka rätten till en rättvis rättegång. Det kan då finnas skäl att använda AI i domstolarna för de ändamålen. Samtidigt visar diskussionen i avsnitt 4.2–4.4 att användningen av AI-system kan innebära en risk för att rätten till en rättvis rättegång kränks. Reglering kan och bör användas som verktyg för att säkerställa att AI inte används på ett sätt som kränker rätten till en rättvis rättegång. Då kan AI även utnyttjas i de avseenden där AI kan bidra till att stärka rättigheten.

I dagsläget finns få bindande regler som uttryckligen begränsar möjligheten att använda AI i domstol, eller som garanterar vissa rättigheter för individer som möter tekniken i den offentliga förvaltningen. Det ska inte förstås som att AI får användas utan begränsningar i dag. Frågan är om de regler som redan finns är tillräckliga för att säkerställa att rätten till en rättvis rättegång inte kränks om

90 Babuta, A., 2018, s. 4.

91 För vidare diskussion om bristande transparens, se avsnitt 4.3 och 4.4.

92 Wisser, L., 2019, s. 1816.

AI används i domstolarna. Nedan diskuteras först några krav som redan ställs på när och hur AI får användas i offentlig verksamhet. Därefter lyfts behovet av nya regler för att säkerställa att användningen av AI inte kränker rätten till en rättvis rättegång.

5.2 BEFINTLIGA REGLERINGAR OCH RIKTLINJER OM NÄR OCH HUR AI FÅR ANVÄNDAS

5.2.1 DATASKYDDSFÖRORDNINGEN

Några regler som begränsar möjligheten att använda AI för automatiserat beslutsfattande finns i dataskyddsförordningen⁹³ (GDPR). I art. 22.1 GDPR finns till exempel en bestämmelse som anger att personer som har fått sina personuppgifter registrerade med stöd av GDPR har rätt att inte bli föremål för beslut som enbart grundas på automatiserat beslutsfattande, inbegripet profilering, vilket har rättsliga påföljder för honom eller henne, eller på ett liknande sätt i betydande grad påverkar honom eller henne.⁹⁴ Profilering definieras i art. 4.4 GDPR som automatisk behandling av personuppgifter för att bedöma vissa personliga egenskaper hos en person, i synnerhet för att förutsäga bland annat pålitlighet. Automatiserat beslutsfattande i domstolarna som innebär behandling av personuppgifter för att analysera eller förutse till exempel pålitlighet, beteende, arbetsprestation, hälsa eller dylikt kan därför vara otillåten.⁹⁵ Med andra ord kan GDPR till exempel begränsa möjligheten att använda AI för att bedöma risk för återfall i brottslighet.

Det finns också flera bestämmelser enligt vilka den enskilde ska få information om förekomsten av automatiserat beslutsfattande.⁹⁶ Vid sådant beslutsfattande, inbegripet profilering, ska den enskilde enligt samma bestämmelser också få meningsfull information om logiken bakom samt betydelsen och de förutsedda följderna av automatisk behandling av personuppgifter. Vad som utgör meningsfull information om logiken bakom behandlingen preciseras inte i förordningen. Viss ledning kan sökas i riktlinjer från den s.k. Artikel 29-gruppen⁹⁷, men de riktlinjerna är inte bindande.⁹⁸ För att informationen ska anses vara meningsfull bör den enligt riktlinjerna vara tillräcklig för att den

93 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 (GDPR).

94 Jfr även skäl 71 GDPR.

95 Jfr skäl 71 GDPR.

96 Se art. 13.2.f, art. 14.2.g och art. 15.1.h GDPR.

97 Artikel 29-gruppen är den oberoende europeiska arbetsgrupp som behandlade frågor om skydd av personuppgifter och integritetsskydd fram till den 25 maj 2018. Arbetsgruppen ersattes då av den europeiska dataskyddsstyrelsen.

98 Artikel 288 i fördraget om Europeiska unionens funktionssätt.

registrerade ska kunna förstå skälen bakom beslutet. Till exempel kan personen behöva delges information om vilka uppgifter ett AI-system beaktar för att nå sin slutsats.⁹⁹

Trots att GDPR i viss mån begränsar möjligheten att använda AI och garanterar ett visst skydd för den enskilde är det något osäkert hur långt skyddet sträcker sig och därmed hur omfattande begränsningen att använda AI är. Tre aspekter kan lyftas i detta sammanhang. För det första är det osäkert om art. 22.1 GDPR förbjuder domstolarna från att använda automatiserat beslutsfattande, inbegripet profilering, eller om bestämmelsen endast ger en person rätt att begära att domstolen inte använder automatiserat beslutsfattande. Artikel 29-gruppen tolkar bestämmelsen som att det rör sig om ett generellt förbud mot automatiserat beslutsfattande och att det inte behövs någon invändning från en part, men ger ingen närmare förklaring till varför den tolkar bestämmelsen så.¹⁰⁰ Andra menar att en sådan tolkning rimmar sämre med ordalydelsen i art. 22.1 GDPR som föreskriver just en ”rätt att inte bli föremål för ett beslut [...]” och att en läsning av artikeln mot bakgrund av andra artiklar i GDPR talar för att en enskild måste begära att aktören i fråga, i det här fallet domstolen, inte använder automatiserat beslutsfattande.¹⁰¹

För det andra kan ordalydelsen i artikeln läsas antingen som att den är tillämplig på alla former av automatiserat beslutsfattande, eller som att den endast är tillämplig på automatiserat beslutsfattande som inbegriper profilering. Med andra ord är det osäkert om bestämmelsen endast påverkar domstolarnas möjlighet att använda automatiserat beslutsfattande när det inbegriper en behandling av personuppgifter för att bedöma vissa personliga egenskaper hos en fysisk person, eller även i andra fall. I svenska förarbeten och doktrin antyds det att avsikten är att artikeln endast ska tillämpas på automatiserat beslutsfattande som inbegriper profilering, eftersom skäl 71 och 73 i GDPR endast tycks behandla profilering och inte andra automatiserade beslut.¹⁰² Integritetsskyddsmyndigheten gör en annan tolkning och påstår utan någon mer ingående argumentation att det framgår tydligt av art. 22.1 GDPR och skäl 71 att artikeln är tillämplig på allt automatiserat beslutsfattande.¹⁰³ Även om artikeln skulle anses tillämplig på allt automatiserat beslutsfattande

99 Artikel 29-gruppen, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 17 EN WP 251, antagen 3 oktober 2010, s. 25.

100 Artikel 29-gruppen, 2017, s. 19.

101 Öman, S., *Dataskyddsförordningen (GDPR) m.m.*, JUNO version 1A, uppdaterad 29 februari 2020, om art. 22.1 i GDPR.

102 Prop. 2017/18:115, *Anpassning till EU:s dataskyddsförordning av lagen om behandling av personuppgifter i verksamhet med val och folkomröstningar*, s. 30; Öman, S., 2020, om art. 22.1 i GDPR.

103 Datainspektionens yttrande 21 november 2017, Dnr. 2180–2017, s. 6(7)–7(8).

finns dock ett undantag i art. 22.1.b GDPR som ger medlemsstaterna möjlighet att i nationell lag godkänna automatiserat beslutsfattande.¹⁰⁴ Det innebär att förbudet kan begränsas till automatiserat beslutsfattande som inbegriper profilering, oberoende av avsikten bakom art. 22.1 GDPR.

Slutligen begränsas skyddet som följer av art. 22.1 GDPR genom att artikeln endast tillämpas på beslut som *enbart* grundas på automatisk behandling. Artikel 29-gruppen tolkar rekvisitet som att bestämmelsen inte ska tillämpas om en människa med befogenhet att ändra beslutet granskar det på ett meningsfullt sätt.¹⁰⁵ Om artikeln ska förstås på det här sättet innebär det att den ger ett svagt skydd för den enskilde att motsätta sig beslutsfattande som innefattar endast viss grad av automatisering. Det innebär också att GDPR inte begränsar möjligheten att använda AI för automatisering i domstolarna i någon större omfattning.

5.2.2 ETISKA RIKTLINJER FÖR AI I RÄTTSVÄSENDET

Europarådets kommission för effektivisering av rättsväsendet (CEPEJ) antog i december 2018 en stadga med etiska riktlinjer som bör gälla när rättsväsendet använder AI.¹⁰⁶ CEPEJ identifierar några huvudprinciper som kan vägleda lagstiftare, jurister och politiker när de hanterar frågan om AI i rättsväsendet. Stadgan är inte rättsligt bindande, men principerna om ansvarsfull tillämpning av AI signalerar förväntningar på konventionsstaterna. Till exempel anges att både privata och offentliga aktörer måste säkerställa att metoderna de använder inte reproducerar eller förstärker diskriminering som kan finnas i indata och att jurister som arbetar i rättsväsendet alltid ska ha möjlighet att granska och motsätta sig ett AI-systems slutsatser. Riktlinjerna kan bidra till att skapa garantier mot att AI används på ett sätt som kränker rätten till en rättvis rättegång, men flera av riktlinjerna är vaga och generellt utformade. De kan med fördel ses som indikationer på vad som behöver regleras, men för att säkerställa att rätten till en rättvis rättegång inte kränks krävs bindande regler som är anpassade för de olika situationer där AI kan tänkas användas i domstolarna. Detta särskilt mot bakgrund av att ED i målet *Airey mot Island* understryker vikten av att rätten till en rättvis rättegång inte blir teoretisk eller illusorisk.¹⁰⁷

¹⁰⁴ Jfr prop. 2017/18:95 s. 100 där det anges att 28 § FL kan omfattas av undantaget.

¹⁰⁵ Artikel 29-gruppen, 2017, s. 20 f.

¹⁰⁶ CEPEJ, 2018, s. 7 ff.

¹⁰⁷ *Airey mot Island*, p. 24; *Perez mot Frankrike [GC]*, nr. 47287/99, dom meddelad den 12 februari 2004, p. 80.

5.3 BEHOV AV FLER REGLERINGAR FRAMÖVER

För att AI ska kunna användas i domstolarna för att stärka rätten till en rättvis rättegång måste det också säkerställas att AI inte används på ett sätt som kränker samma rättighet. De regleringar som finns i dag synes inte vara tillräckliga för att garantera att rätten inte kränks om AI används i domstolarna. Till exempel identifieras i avsnitt 4.2–4.4 bristande transparens som ett potentiellt hinder mot respekten av rätten till en rättvis rättegång om AI skulle börja användas i domstolarna. Det bör därför upprättas regler som ställer krav på att AI-systemen uppnår en viss grad av transparens, bland annat för att säkerställa att en part får en reell möjlighet att påverka ett beslut eller en dom. Vilka krav på transparens som ska ställas vid användning av AI för olika ändamål bör utredas närmare. Samma krav behöver inte nödvändigtvis gälla vid all användning av AI i domstolarna. Tvärtom kan kraven på transparens utformas med beaktande av vad AI kan användas till samt om och hur den riskerar att påverka den enskildes rättigheter.

Fullt transparent källkod kan vara svårt att kräva, bland annat eftersom den kan vara skyddad som företagshemlighet.¹⁰⁸ En fullt transparent källkod behöver inte heller vara vare sig nödvändig eller önskvärd, eftersom överflödigt information kan vara svår att ta till sig.¹⁰⁹ I detta avseende kan inspiration hämtas från vad MSI-NET kallar ”effektiv transparens” och vilken information som kan behöva delges för att uppnå en sådan. Till exempel kan det röra sig om information om vilka mål en algoritm är optimerad för, vilka data den är upplärd på, vilka variabler den använder, vilken mängd och typ av data som den processar och vilka medelvärden samt standardavvikelser som observeras hos dess slutsatser.¹¹⁰ Vidare kan det krävas information som gör det möjligt att se vem eller vilka som skapat AI-systemet och kontrollera att det är skapat av någon som har tillräckliga meriter.¹¹¹ Det finns också anledning att kräva information om hur AI implementeras och används i rättsväsendet samt hur domarna tränas i tekniken.¹¹²

Vid utformningen av regler om transparens bör det beaktas att information om algoritmer kan komma att användas för att manipulera algoritmerna. Om algoritmerna manipuleras finns det bland annat en risk för att syftet

¹⁰⁸ Lundqvist, U., 2020, s. 390 och 395 f.

¹⁰⁹ MSI-NET, 2017, s. 38.

¹¹⁰ MSI-NET, 2017, s. 38. Jfr även New York City, Local law no. 49 of 2018 och New York City Automated Decision Systems Task Force Report, november 2019, tillgänglig på <https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf>, s. 19 och s. 23 (besökt 29 januari 2021).

¹¹¹ Washington, A., 2019, s. 154 f.

¹¹² Washington, A., 2019, s. 154 f.

med den automatiserade åtgärden motverkas. Ett mål från Kammarrätten i Stockholm illustrerar denna problematik. En kvinna begärde att få ta del av datorprogram eller algoritmer som genererade beslut om tandvårdsbidrag från Försäkringskassan. Kammarrätten angav att de efterfrågade uppgifterna indirekt kunde ge information om hur systemet kunde kringgås eller manipuleras. Ett utlämnande av uppgifterna riskerade enligt Kammarrätten att medföra att syftet med den automatiserade åtgärden, en korrekt användning av allmänna medel för tandvård, motverkades.¹¹³ Eventuella regler om transparens bör utformas med hänsyn till att viss information kan behöva förbli hemlig för att undvika manipulation av AI-systemet.

Utformningen av reglerna kommer att kräva en avvägning mellan flexibilitet och praktisk betydelse. Det gäller att förhindra att reglerna blir obsoleta snabbt och samtidigt undvika att de endast blir ett spel för galleriet. En möjlighet är att i lag reglera informationsskyldigheten i bredare termer och i förarbeten exemplifiera vilken typ av information som bör eller måste delges för att ett AI-system ska uppnå tillräcklig grad av transparens. De specifika kraven på informationsdelgivning kan då anpassas när tekniken utvecklas. För de fall att det inte går att garantera den transparens som krävs för att rätten till en rättvis rättegång ska respekteras i en given situation bör användningen av AI i den situationen uttryckligen förbjudas.

För att säkerställa att rätten till en rättvis rättegång inte kränks om AI används i domstolarna bör det också införas andra regler om hur och när AI får användas. Det bör krävas att en domstol specifikt anger när den använder AI och domstolarna bör åtminstone i ett första skede förbjudas att förlita sig fullt ut på AI i beslutsprocesser.¹¹⁴ Domstolarna bör också vara skyldiga att redogöra för risker kopplade till användningen av AI.¹¹⁵ Andra regler bör ställa krav på hur AI ska granskas på förhand för att få användas i domstolarna, hur bias i data ska hanteras, hur AI ska granskas och utvärderas löpande, i vilken grad domare får förlita sig på AI, vad som händer om AI som används inte uppfyller kraven som ställs på tekniken och i vilken mån utvecklare eller användare av AI-systemen ska hållas ansvariga när det visar sig att ett AI-system är behäftat med fel eller begår misstag. Mot bakgrund av att lägre och högre stående domstolar har olika funktioner i rättsväsendet kan olika regler gälla i olika instanser.

113 Kammarrätten i Stockholm, dom meddelad den 27 augusti 2019 i mål 4995–19.

114 Jfr t.ex. *Wisser*, L., 2019, s. 1827 ff. och *Loomis mot Wisconsin*, p. 98–100.

115 *Loomis mot Wisconsin*, p. 98–100.

I april 2021 presenterade EU-kommissionen ett förslag till en förordning om AI.¹¹⁶ Förslaget utgår från en riskbaserad metod där olika regler gäller beroende på vilka risker användningen av AI är förenad med. AI som är avsedd att användas i rättsväsendet för att göra riskbedömningar eller för att assistera domare genom att göra rättsutredningar samt tolka och tillämpa lag i konkreta fall är exempel på AI som bedöms vara av ”hög risk”.¹¹⁷ AI-system som utgör en ”hög risk” ska bland annat utformas och utvecklas på ett sätt som säkerställer att de är tillräckligt transparenta för att användaren ska kunna tolka AI-systemets slutsats (utdata).¹¹⁸ Data som används för att träna AI-system ska vidare uppfylla vissa kvalitetskrav. Till exempel ska man göra en bedömning om eventuella bias i data och dataseten ska vara relevanta, representativa, fullständiga och fria från felaktigheter.¹¹⁹ Det föreslås också regler om ändamålsenliga åtgärder för mänsklig översyn av AI-systemen och att AI ska utformas och utvecklas för att uppnå en ändamålsenlig nivå av noggrannhet, robusthet och cybersäkerhet.¹²⁰

EU-kommissionens förslag måste ses som ett viktigt steg på vägen mot harmoniserade regler om AI i EU. Om Europeiska unionens råd och EU-parlamentet kan enas och rösta igenom en slutlig AI-förordning kan reglerna möjligen säkerställa, eller åtminstone bidra till att säkerställa, att rätten till en rättvis rättegång inte kränks om AI skulle användas i svenska domstolar. Beroende på hur reglerna utformas i en eventuell AI-förordning kan dock tänkas att det också kommer att behövas en nationell lag som preciserar när och hur AI får användas i de svenska domstolarna.

6. AVSLUTANDE KOMMENTARER

Att den tekniska utvecklingen kommer att fortsätta påverka juristers arbete på olika sätt är nog många överens om. Om och när det är lämpligt att ersätta mänskliga bedömningar med AI är upp till oss människor att avgöra. Jag menar att det kan vara lämpligt att använda AI om tekniken kan stärka rätten till en rättvis rättegång, men AI får inte användas på ett sätt som kränker den rättigheten. För att garantera att AI introduceras i domstolarna på ett rättssäkert sätt är det viktigt att möjligheterna, utmaningarna och hoten med AI fortsätter att utredas. Den här artikeln kan ligga till grund för vidare diskussioner om när

116 EU-kommissionen, *Proposal for a Regulation of the European Parliament and of the Council: Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, Brussels 21.4.2021, COM(2021) 206 final, 2021/0106(COD).

117 Annex III, Artificial Intelligence Act, p. 6(a) och p. 8. Innebörden av begreppet ”hög risk” i detta avseende skiljer sig alltså från hur begreppet används i avsnitt 3 och 4.3.2.

118 Art. 13, Artificial Intelligence Act.

119 Art. 10, Artificial Intelligence Act.

120 Art. 14 och 15, Artificial Intelligence Act.

och hur AI kan eller till och med bör användas i svenska domstolar, samt vilka krav som då måste ställas för att garantera att rätten till en rättvis rättegång inte kränks. Det finns dock ett stort behov av fortsatt forskning på området. Vid sidan av de möjligheter, utmaningar och förslag på regleringar som lyfts fram i denna artikel finns säkerligen fler problemområden och juridiska lösningar att beakta. Jag ser fram emot vidare diskussioner och att få följa hur tekniken kommer att förändra arbetet i domstolarna framöver. 