Integritetsskyddet i AI-lagstiftningen

JOHANNA CHAMBERLAIN*

*Prisföreläsning vid högtidlighållandet av Stig Strömholms 93-årsdag den 16 september 2024***

Kära åhörare, bäste Stig!

För tre år sedan – på dagen – hade jag äran att få vara med och tala vid Stig Strömholms 90-årsfirande. Ämnet den gången var utvecklingen av det svenska integritetsskyddet under de senaste 50 åren.[1] Idag känns det oerhört hedrande att få uppmärksamma Stigs bemärkelsedag igen, nu som mottagare av Stig Strömholms pris för humanistisk rätts- och samhällsvetenskaplig forskning år 2024. Mycket har hänt under de tre år som har passerat. Vårt samhälle har gått från pandemi till andra kriser och stora utmaningar för integritetsskyddet, såväl på global nivå som i svenskt sammanhang. Idag tänkte jag tala om en av dessa utmaningar, och på så vis tar jag också vid där jag slutade sist. Från EU:s omtalade dataskyddsförordning[2] som kom år 2018 har utvecklingen nämligen gått snabbt till EU-förordningen om artificiell intelligens (AI)[3] som trätt i kraft redan nu och kommer att börja gälla fullt ut under 2026 och 2027, den så kallade AI-akten.

I det här anförandet tänkte jag i tur och ordning uppmärksamma vilka risker mot rätten till privatliv och dataskydd som AI kan sägas skapa, hur dessa problem har hanterats i AI-akten och slutligen förhållandet mellan dataskyddsregleringen och AI-akten. Det här är frågor som jag har arbetat med de senaste åren

* Docent och universitetslektor i civilrätt vid Juridiska institutionen, Uppsala universitet.

** Föreläsningsmanuset publiceras här såsom det framfördes, dock har några noter lagts till. Dessa innehåller hänvisningar främst till EU-lagstiftning, men även till några texter där jag och medförfattare utvecklat teman som endast nämns kort i föreläsningen.

1 Se Johanna Chamberlains, Therése Fridström Montoyas & Johan Hirschfeldts omarbetade föreläsningar från 90-årsfirandet i SvJT nr 7 2022.

2 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).

3 Europaparlamentets och rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens och om ändring av förordningarna (EG) nr 300/2008, (EU) nr 167/2013, (EU) nr 168/2013, (EU) 2018/858, (EU) 2018/1139 och (EU) 2019/2144 samt direktiven 2014/90/EU, (EU) 2016/797 och (EU) 2020/1828 (förordning om artificiell intelligens).

under min tid som postdoktor, inom ramen för ett större projekt som har handlat om risk- och ansvarsfrågor kopplade till AI.[4] I detta projekt har mitt perspektiv varit främst skadeståndsrättsligt, vilket också kommer att märkas i det följande.

AI-teknologin är här för att stanna och den för med sig många fördelar. Men vilka utmaningar för vår privata sfär och personliga data orsakar AI? Jämfört med de senaste decenniernas intensiva diskussion kring behandling av personuppgifter, där det rättsliga skyddet också har ökat i takt med samhällets digitalisering, kan man tycka att dataskyddsperspektivet har tagit förvånansvärt lite plats i AI-utvecklingen. Alla AI-system bygger ju på enorma mängder just personliga data. Desto mer uppmärksamhet har ägnats åt övervakningsaspekterna med AI. Ett återkommande tema är den ofta problematiskt diffusa gränsdragningen mellan användning av AI för exempelvis brottsuppklarande verksamhet, vilket med de senaste årens ökande fokus på gränsöverskridande kriminalitet ofta anses motiverat, och ett övervakningssamhälle som européerna inte önskar.

I slutversionen av AI-akten som godkändes i maj i år nämns ordet ”privatliv” 16 gånger och ”dataskydd” 51 gånger. Betydelsen av dessa värden har förstärkts under lagstiftningsprocessen; i kommissionens ursprungliga förslag till AI-akt nämndes ”privatliv” fyra gånger och ”dataskydd” 30. Konkreta hot och risker som identifieras av unionslagstiftaren är bland annat AI-system för ansiktsigenkänning och för övervakning på arbetsplatser.[5] På ett mer principiellt plan kopplas hot mot privatlivet också till hot mot demokratin.[6] Detta utvecklas inte i förordningen, men känns igen från fördigital tid som ett tema med nära anknytning till viktiga värden som annars ofta ställs *mot* rätten till privatliv: yttrandefrihet och transparens. Den som känner sig övervakad kan inte uttrycka sig och bete sig fritt, till exempel genom att skapa konst och litteratur, demonstrera eller kommunicera med andra människor.

Mot den här bakgrunden är det intressant att man inom EU (och med Sverige i spetsen) samtidigt har drivit på för införandet av den så kallade ”Chat control”-lagstiftningen,[7] som skulle öppna för massövervakning av privat digital kommunikation med motiveringen att motarbeta barnpornografi. Det ursprungliga förslaget fick nej av Europaparlamentet bland annat med hänsyn till att det gick emot rätten till privatliv och dataskydd enligt EU:s rättighetsstadga. Förhandlingar pågår nu om en omarbetad version – där AI ska användas för att scanna av och avgöra om bilder som skickas mellan människor är semesterbilder från stranden eller olaglig barnpornografi. Chat control-debatten kan ses som ännu

[4] Projektet, ”AI and the Financial Markets: Accountability and Risk Management with Legal Tools”, finansierades av WASP-HS.

[5] Se AI-aktens beaktandesatser 43–44.

[6] AI-aktens beaktandesats 8.

[7] Förslag till Europaparlamentets och rådets förordning om fastställande av regler för att förebygga och bekämpa sexuella övergrepp mot barn, COM/2022/209 final. Lagstiftningsärendet initierades av den f.d. svenska EU-kommissionären Ylva Johansson.

en illustration av privatlivets status som ett individuellt värde som ofta får ge vika för kollektiva motintressen.

Om vi går tillbaka till AI-akten kan det konstateras att unionslagstiftaren har formulerat ett narrativ kring AI som bygger på *kontroll av risker* knutna till AI.[8] Det här är ett rimligt sätt att angripa problemet eftersom risker enligt en etablerad vetenskaplig bild går att beräkna, försäkra och just kontrollera – samtidigt som man kan fråga sig om det är just risker det handlar om och inte ren osäkerhet, eftersom kunskapen om vilka typer av skador som AI kan leda till ännu är så begränsad. Osäkerhet är betydligt svårare att hantera juridiskt än risk. Valet av terminologi är däremot förståeligt ur perspektivet att man vill skapa tillit för den nya teknologin bland konsumenter och EU-medborgare, och på så vis främja den inre marknaden.

AI-akten grundar sig på en uppdelning av AI-system i fyra olika riskkategorier, och illustreras ofta med en pyramid bestående av fyra nivåer. Regleringen riktar främst in sig på den näst högsta nivån – där man finner de så kallade *högrisksystemen*. System med *minimal* eller *begränsad* risk anses inte (eller knappt) behöva regleras. System med minimal risk faller utanför regleringen, medan system med begränsad risk endast behöver uppfylla vissa transparenskrav. Men högst upp i pyramiden finns det också AI-system som anses skapa *oacceptabla* risker och därför som utgångspunkt ska vara förbjudna. Dessa system anses så riskfyllda att förbudet träder i kraft redan i februari 2025, alltså betydligt tidigare än den vanliga tvååriga implementeringstiden som gäller för nya EU-direktiv och förordningar. I AI-aktens inledande satser betonas att AI-utvecklingen har öppnat för nya sätt att manipulera, utnyttja och utöva social kontroll. Sådana skadliga verksamheter ska förbjudas eftersom de strider mot grundläggande unionsvärderingar om respekt för mänsklig värdighet, frihet, jämlikhet, demokrati, rättsstatsprincipen och grundläggande rättigheter som kommer till uttryck i EU-stadgan.[9] Rättigheter som här nämns specifikt är rätten till privatliv och dataskydd, vilka alltså direkt har påverkat vilka AI-system som ska vara tillåtna inom EU. Men som bekant är dessa rättigheter inte absoluta utan relativa. De kan begränsas och i AI-akten listas omständigheter där annars förbjudna AI-system trots allt är tillåtna, framför allt gällande brottsuppklarning och terroristbekämpning.[10]

Vad gäller högrisksystemen beskrivs de av unionslagstiftaren som system som utgör betydande risk för fysiska personers hälsa, säkerhet eller grundläggande rättigheter, såsom rätten till privatliv och dataskydd.[11] Trots detta ska de

[8] Temat utvecklas i Johanna Chamberlain & Andreas Kotsios, ”Defining Risk and Promoting Trust in AI Systems”, och i Maria Bergström & Valsamis Mitsilegas (red), EU Law in the Digital Age: Swedish Studies in European Law (London, Hart/Bloomsbury Publishing, kommande 2025).

[9] AI-aktens beaktandesats 28.

[10] Se AI-aktens beaktandesats 33, samt förordningens artikel 5.

[11] AI-aktens beaktandesats 48.

vara tillåtna, men endast i enlighet med AI-aktens omfattande kontrollsystem bestående av förhandsbedömningar av riskfaktorer, säkerhetsåtgärder, registrering av system i en databas som ska byggas upp, tillsyn i flera olika nivåer – mänsklig tillsyn, nationella tillsynsmyndigheter och en tillsynsmyndighet på unionsnivå – utvärderingsmekanismer och inte minst kraftiga sanktionsavgifter som kan beslutas vid konstaterade överträdelser. Här rör det sig uppenbarligen om en balansakt som unionslagstiftaren utför, mellan AI-systemens effektivitet och samhällsnytta och de nackdelar som samtidigt framträder.

Hur ska då AI-hoten mot privatliv och dataskydd hanteras? Det här lär visa sig gradvis, i takt med att hot och risker konkretiseras samtidigt som systemen blir alltmer sofistikerade. Men med tanke på att det rör sig om så pass betydande risker är ambitionen i AI-akten ändå att sätta upp ett proaktivt regelverk, bland annat genom de nyss nämnda åtgärderna som förhandsbedömning och registrering. I AI-akten slås det fast att rätten till privatliv och dataskydd måste garanteras genom hela AI-systemets livscykel. Detta ska göras genom att principerna om dataminimering och inbyggt dataskydd iakttas.[12] Så lite data som möjligt ska alltså användas, och dataskydd ska vara en del av processen när AI-system designas. Dataskyddsreglerna – där främst dataskyddsförordningen kan antas bli aktuell – ska givetvis följas vid utformandet av AI-system. Privatlivsskyddande åtgärder nämns också flera gånger i AI-akten, där ett exempel är pseudonymisering av uppgifter.[13] Utvecklingen på detta område går också mot mer avancerade sätt att skapa artificiella data (så kallad syntetisk data) med hjälp av algoritmer, just i syfte att slippa hantera känsliga personuppgifter.

De tillsynsmyndigheter som ansvarar för att dataskyddsreglerna efterlevs kommer med AI-akten att få ett utökat ansvar. I Sverige kan man därmed vänta sig att Integritetsskyddsmyndigheten, som redan arbetar intensivt med skärningspunkten mellan dataskydd och AI, får ytterligare omfattande ansvarsområden kopplade till AI. Det gäller allt från hjälp med förhandsbedömningar av system till utvärdering och åläggande av de potentiellt mycket höga sanktionsavgifterna. Sanktionsavgifterna kan vara på upp till 35 miljoner euro,[14] vilket alltså är betydligt högre än dataskyddsförordningens maxnivå på 20 miljoner euro.[15] De europeiska tillsynsmyndigheternas beslut om sanktionsavgifter enligt dataskyddsförordningen har fått mycket uppmärksamhet de senaste åren, och införandet av kraftfulla sanktionsavgifter även i AI-förordningen kan ses som ett uttryck för att sanktionsavgifterna har ansetts effektiva i dataskyddskontex-

[12] AI-aktens beaktandesats 69.
[13] AI-aktens artikel 10 b.
[14] AI-aktens artikel 99.3.
[15] Dataskyddsförordningens artikel 83.5.

ten. Sett över flera rättsområden är det även en allmän tendens på EU-nivå att tillsynsmyndigheter får allt större makt, vilket vi bör vara medvetna om.[16]

Ett annat tillvägagångssätt för att hantera svårbedömda risker på ett nytt område är att använda sig av befintlig kunskap och lagstiftning. Under mitt postdok-projekt har jag tittat på möjligheterna att relatera riskstrukturen i AI-förordningen till skadeståndsrättsliga ansvarsformer och riskbedömningar.[17] Riskbedömningen, eller riskfördelningen, är central både när det gäller strikt ansvar och culpaansvar inom skadeståndsrätten. I det första fallet gäller det bedömning av vilka verksamheter som är så riskfyllda att strikt ansvar anses motiverat, och här finns det tydliga paralleller till diskussionen om AI-system med hög respektive oacceptabel risk. Inom culpaansvaret aktualiseras risk i den så kallade fria culpabedömningen, där risken för skada är en av faktorerna att ta hänsyn till – precis som vid klassificeringen av AI-system i högriskkategorin. Riskbegreppet inom de båda skadeståndsrättsliga ansvarsformerna har i den svenska kontexten hunnit undersökas inte bara i förarbeten och litteratur utan även i ett stort antal rättsfall, eftersom skadeståndslagen har varit i kraft i över 50 år. Genom att applicera de etablerade skadeståndsrättsliga principer, gränsdragningar och teorier som finns på AI-området kan de nya utmaningarna konkretiseras samtidigt som skadeståndsrätten uppdateras. På så sätt kan den traditionella skadeståndsrätten bidra med förståelse för EU:s framväxande regelverk.

En fråga som ibland ställs i AI-sammanhang är om EU riskerar att ägna sig åt överreglering. Bör inte dataskyddsförordningen och andra närliggande regleringar i stort sett redan omfatta AI-system? AI kommer framöver inte endast att regleras i AI-akten på EU-nivå, utan även i den uppdaterade produktansvarsregleringen, i ett nytt direktiv om de processuella förutsättningarna för att hantera AI-ansvar, och som en del av många andra lagstiftningsakter för den digitala tidsåldern – vilket är ett stort lagstiftningstema inom EU just nu. Hur ska alla dessa regleringar förhålla sig till varandra? Är det någon som kommer att lyckas ha överblick över den alltmer komplexa lagstiftningsstrukturen? Mot bakgrund av dessa frågor är det intressant att titta närmare på förhållandet mellan dataskydds- och AI-reglerna. Tyvärr säger unionslagstiftaren inte mycket om detta, utan i stort sett bara att AI-akten kompletterar existerande lagstiftning såsom dataskyddsförordningen. Verksamheter som redan är förbjudna enligt till exempel dataskyddsreglering träffas inte av AI-akten.[18] Samtycke till testning inom ramen för AI-system särskiljs från samtycke till databehandling enligt

[16] Se för diskussion av detta Johanna Chamberlain & Jane Reichel, ”Supervision of Artificial Intelligence in the EU and the Protection of Privacy”, Florida International University Law Review, vol 17 no 2, s 267–285.

[17] Detta arbete redovisas i Johanna Chamberlain, ”The Risk-Based Approach of the European Union’s Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law Perspective”, European Journal of Risk Regulation, vol 14 no 1, s 1–13.

[18] AI-aktens beaktandesats 45.

dataskyddsförordningen.[19] Det måste ändå sägas att en alltmer komplicerad och svårnavigerad regelstruktur inte kan antas gynna vare sig utvecklare av AI-system eller unionsmedborgare som lidit skada genom felaktig personuppgiftsbehandling eller omotiverad övervakning.

Som nämndes inledningsvis utgör tillgång till högkvalitativ data en grundförutsättning för utveckling och användning av många AI-system. Unionslagstiftaren tydliggör i AI-akten att högkvalitativa dataset för testning och träning av system kräver implementering av lämpliga datastyrnings- och -hanteringsåtgärder.[20] För att underlätta efterlevnad av dataskyddsreglerna ska sådana åtgärder innefatta transparens avseende datainsamlingens ursprungliga ändamål. Vidare uppmärksammas att man i användningen av dataseten måste undvika att historiska data blir missvisande och därmed kan få negativa effekter för grundläggande rättigheter.[21] Det här betyder att man måste vara mycket försiktig med användningen av personuppgifter, så att inte snedvridna data leder till exempelvis diskriminering (vilket har skett vid AI-baserad rekrytering). Kravet på korrekta uppgifter ska dock inte påverka användningen av privatlivsskyddande åtgärder vid utformandet av AI-system, markerar unionslagstiftaren.[22] Intressant nog anas det alltså spänningsytor här mellan skyddet för privatlivet och dataskyddsreglerna, som ytterligare en följd av den komplicerade regelstrukturen och förhållandet mellan olika lagstiftningsakter.

Så har det blivit dags för en summering av de spaningar som nu har presenterats kring integritetsskyddet i AI-kontexten. Liksom vid varje nytt teknikgenombrott har vi under en tid levt i laglöst land, med medialt uppmärksammade följder som diskriminering genom AI-genererade verktyg för påföljdsbestämning, manipulering av barn och andra sårbara personer på nätet samt felaktigt utpekande av bidragsfuskare och orimliga skolplaceringar. Det finns ett behov av reglering och det är ett stort och betydelsefullt steg som EU har tagit med AI-akten. Kanske har den potential att få ett lika stort genomslag som dataskyddsförordningen har fått, och skapa en ny global standard. Samtidigt är det uppenbart att vi lever i kristider, och i sådana tider glöms rättigheter som privatlivsskydd och dataskydd lätt bort – fastän de är av största vikt för att upprätthålla demokratin. På ett personligt plan talar man ofta om att kris betyder utveckling. En tanke som slår mig i det här sammanhanget är istället att ordet kris med omkastade bokstäver blir risk. En klar fördel med AI-akten är just att den riskbaserade regleringsmodellen uppmärksammas. Risk är ett begrepp som studeras inom många olika vetenskaper och även förtjänar tvärvetenskaplig analys. Det är också ett begrepp som får allt större betydelse inom skadeståndsrätten, där man till exempel kan ställa frågan om läckta personuppgifter

[19] AI-aktens beaktandesats 141.
[20] AI-aktens beaktandesats 67.
[21] AI-aktens beaktandesats 67.
[22] AI-aktens beaktandesats 67.

och risk för identitetsstöld utgör ersättningsgill skada. Bland alla de risker som idag omger oss, inte minst de AI-genererade, har vi kanske alla ett behov av att få höra att de här riskerna kan kontrolleras. Om AI-akten med tiden kan åstadkomma detta återstår att se.