Riksarkivet

Yttrande
2021-06-24

Dnr RA-KS 2021/01284
Ert Dnr I2021/01304

Infrastrukturdepartementet
103 33 Stockholm

Riksarkivets yttrande över Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens, Proposal for a Regulation laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (COM(2021) 206)

Riksarkivet har tagit del av rubricerad remiss och vill lämna följande synpunkter.

Övergripande synpunkter

Det är rimligt och lämpligt att kommissionen genom förordningen om harmoniserade regler för artificiell intelligens (AI-förordningen) valt att begränsa regleringen av utveckling och användning inom AI och datadriven utveckling i EU till s.k. hög-risk-system. Detta minskar den negativa påverkan av regleringen och påverkar inte konkurrenskraften hos europeiska företag eller minskar viljan för utomeuropeiska företag att vilja sälja AI-system som ska användas inom EU eller av EUs medborgare i lika hög grad som ett mer långtgående förslag som reglerar all AI-användning..

Det är positivt att man har valt att komplettera dokumentationen och tillsynen för hög-risk-systemen med en code of conduct som uppmuntrar till att även andra system ska dokumenteras. Denna transparens är positiv inför upphandlingar och inför konsumenter och medborgare som påverkas av användningen av AI-system eftersom det ger en möjlighet till insyn och kvalitetssäkring.

Det är bra att man bygger vidare och samordnar AI-förordningen med existerande lagstiftning, såsom dataskyddsförordningen och andra regleringar på den inre marknaden. Vidare är det en bra ambition att försöka göra AI-förordningen så framtidssäkrad och teknikneutral som möjligt, även om det i realiteten är mycket svårt att göra.

Den föreslagna dokumentationen kan förväntas skapa en viss transparens gentemot tillsynsmyndigheter och användare att utvecklandet av systemet

Postadress: Riksarkivet, Box 7223, 187 13 Täby · **Organisationsnr:** 202100-1074
Telefon: 010-476 70 00 · **E-post:** riksarkivet@riksarkivet.se · **Internet:** riksarkivet.se

håller en viss kvalitet/klass och att uppenbara fel med snedfördelning och skevhet i dataset kan upptäckas i ett tidigt skede. Dokumentationen kan dock komma att invagga användarna i en falsk trygghet om att man förstår den output som kommer från AI-systemet. Även om ”explainable AI” är ett stort forskningsfält just nu så kommer man inte ifrån att det inte har så lovande resultat ännu och att när man använder data för att skapa en AI-modell så har man i dagsläget och sannolikt ett bra tag framöver ingen kontroll eller förståelse för hur modellen når de slutsatser som den kommer fram till. Därför kommer den här dokumentationen inte att gå att använda som ett sätt att förstå hur AI-systemet genererar en viss output. De som sätter AI-system i bruk på den inre marknaden måste ha en förståelse för detta och att om de använder datadrivna metoder som maskininlärning för att skapa system för automatiserade beslut så är det den givna konsekvensen. Av det följer att datadrivna system inte kan agera helt autonomt inom vissa områden och att en datadriven approach till systemutveckling inte alltid kommer att löna sig.

Det finns anledning att anta att genomdrivandet av AI-förordningen kommer att hämma utvecklingen inom maskininlärningsdriven språkteknologi för de europeiska språken. De ledande företagen inom AI finns i USA och Kina och det har fram tills nu varit framförallt de amerikanska företagen som utarbetat lösningar för europeiska språk. Utökade regleringar på den europeiska marknaden kommer att fördyra och minska deras vilja till att samla in och processa data från EU, vilket kommer att göra att teknikutvecklingen inom det här området kommer att gå långsammare än för andra delar av världen. Sämre språkförståelse hos AI-systemen innebär i sin tur en ökad risk för missförstånd och felaktiga beslut. Detta kan därför leda till att system som används inom EU kommer att prestera sämre och därför vara sämre för användarna än system i andra delar av världen.

Skälen

Nedan följer Riksarkivets synpunkter på specifika punkter i skälen, s. 17 och framåt.

s. 18 punkt 5: AI-förordningen syftar till att skapa balans mellan säkerheten och integriteten hos EUs medborgare samtidigt som man ska skapa goda förutsättningar för företag som utvecklar tjänster med AI. Det är inte uppenbart att man har hittat den bästa avvägningen mellan dessa igenom det föreliggande förslaget. En tvingande inre marknad för AI-utveckling där medlemsländerna inte kan välja att även reglera system som inte är hög-risk inte nödvändigtvis är en rimlig avvägning där medborgarnas integritet prioriteras jämfört med företagens intressen.

s. 19 punkt 8 och 9: det är bra att det tas upp att det finns skillnader mellan ”real-time” och ”post” biometriska identifikationssystem, men det är inte uppenbart att skillnaden dem emellan i funktionalitet kommer att vara framtidssäkert (”future proof”) vilket annars är ett tydligt uttryckt önskemål

med hela AI-förordningen. Detta eftersom beräkningsmöjligheterna och dataaggregationen kan antas öka.

s. 20 punkterna 10 och 11: I dessa punkter beskrivs olika tankar om hur AI-förordningen ska gälla även för företag inom EU som skapar system som inte ska användas inom den inre marknaden utan i tredje land, men som kvalificeras som hög-risk och vars effekter påverkar personer inom EU. Det är väldigt svårt att se hur det skulle kunna övervakas och hur tillsynen kring detta ska kunna genomföras eftersom den sista punkten - kan påverka personer inom EU - är väldigt öppen för tolkning. Kommissionen försöker även hitta sätt för att täppa till möjligheter att kringgå AI-förordningen så att de som skapar och använder system utanför unionen ”*to the extent the output produced by those systems is used in the Union.*” Hur ska medlemsstaterna kunna följa upp det?

Det kommer också att innebära att möjligheterna att dela dataset, öppna data och öppen källkod kringskärs globalt, särskilt med tanke på att företagen och användarna av AI-system riskerar att dra på sig stämningar mm från aktörer inom EU.

s. 21 punkt 17: I punkten förbjuds system från offentlig verksamhet (eller å dess vägnar) som betygsätter användare med ”sociala poäng”. Samma restriktion verkar emellertid inte föreslås för privat sektor. Givet att vissa företag har närmast monopolställning inom vissa områden så kan man fråga sig om det är rimligt eller om det rent av inte behövs restriktioner även för dessa aktörer.

s. 21-23 punkterna 18-24: Biometri, anges på flera ställen i dokumentet. I dessa punkter i anslutning till polisiär verksamhet. Det vore kanske bättre att samlade alla punkter som rör biometri tillsammans och inte sprider ut dem i listan så som nu är fallet. Men överlag är resonemanget kring användning av biometri inom polisiär verksamhet rimlig som en lägstanivå.

s. 26, punkt 33: Biometri en gång till och på ett område som berörs på sidan 21. Hade än en gång nog varit bättre att samla alla punkter på ett ställe samt skapa underrubriker för olika kategorier.

s. 26 punkt 34: Kommunikation och IT-infrastruktur som fiber och 4G, 5G borde stå med som kritisk teknisk infrastruktur.

s. 26-27 punkt 35-37: Punkterna tar upp frågan om att AI system som rör bedömningar om rätt till utbildningar, arbete eller kreditvärdighet bör klassas som hög-risk eftersom de kan ha svåra konsekvenser för den enskilde om det finns okänd/okontrollerad bias i systemet. Detta är inte orimligt, men samtidigt bör man ha i åtanke att denna typ av tjänsteutövning idag utförs av oerhört subjektiva människor. Det enda sättet att komma runt detta ter sig vara att skapa system som är helt regelbaserat algoritmiska och inte datadrivna. Att försöka komma åt problemet med detta genom att ålägga utvecklingen av AI-system en hög dokumentationsbörda är inte uppenbart den bästa lösningen.

s. 28 punkt 38: Det är märkligt att man undantar AI-system som ska arbeta administrativt med skatter och tull från att vara högrisk eftersom man kan argumentera för deras påverkan för den enskilde och risker för bias i lika hög grad som de andra typerna av myndighets- och rättsutövning som nämnts.

s. 29 punkt 44: I punkten står det

”*In order to protect the right of others from the discrimination that might result from the bias in AI systems, the providers shouldbe (sic!) able to process also special categories of personal data, as a matter of substantial public interest, in order to ensure the bias monitoring, detection and correction in relation to high-risk AI systems.*”

Detta behöver förtydligas. Vad innebär “special categories of personal data”?

s. 30 punkt 46: Rent logiskt krävs den här typen av dokumentation och uppföljning för att kontrollera regelefterlevnad av IT-system generellt, men det är viktigt att de som sätter systemen i bruk och de som utför tillsynen förstår att det inger en falsk trygghet att ha så väldokumenterade system eftersom AI-modeller ändå förblir ”black boxes” vars funktionalitet inte till 100% kan förstås eller kontrolleras.

s. 30 punkt 47: ”Users should be able to interpret the system output and use it appropriately”. Här utgår vi från att det som avses är att hög-risk-systemen ska ingå i en halv-automatiserad arbetsprocess där output från systemen ligger till grund för (myndighets-) beslut mm som fortfarande fattas av människor men där AI-systemet tar fram beslutsunderlag snarare än fattar beslut självständigt. Det är inte lämpligt att ha helautomatiserade beslutsprocesser som bygger på maskininlärning eftersom dessa inte kan leva upp till bl.a. GDPRs krav på att människor ska kunna få beslut som gäller dem förklarade för sig.

s. 30 punkt 48: I punkten står det ”*High-risk AI systems should be designed and developed in such a way that natural persons can oversee their functioning[…]*”. Här verkar kommissionen ställa krav på att hög-risk system ska övervakas kontinuerligt av personer med tillräckligt hög kompentens för att upptäcka och åtgärda om systemens beteenden börjar falla utanför de givna eller förväntade ramarna. Det verkar innebära att samtliga organisationer som upphandlar AI-system måste ha högt kvalificerade ingenjörer som övervakar systemen. Detta kan fördyra användningen av AI-systemen avsevärt och kommer i praktiken att sätta gränser för vilka företag och offentliga organ som kan implementera och använda den här typen av lösningar. Har detta tagits med i kostnadsberäkningen för genomförandet av AI-förordningen? Eftersom det redan råder brist på ingenjörer med maskininlärningskompetens så kommer detta att hämma utvecklingen inom EU ytterligare, förutom de aspekter som redan omnämnts ovan.

s. 30 punkter 50-51: Dessa punkter bör ligga mycket tidigare i texten. En väl etablerad och genomförd cybersäkerhet är en helt fundamental förutsättning för att man ska våga implementera hög-risk-AI-system. Det är därför svårt att förstå att frågan behandlas först som punkt 50 i förslaget.

s. 31 punkt 56: I punkten anges att *providers* av AI-system som tagits fram i tredje land måste se till att det finns en person inom EU att vända sig till om inte en importör kan identifieras för systemet. Det är svårt att se hur det ska kunna kontrolleras och vara robust över tid.

s. 32 punkt 60: Den här punkten är väldigt svävande samtidigt som den ställer krav på ”relevanta tredjeparter” som exempelvis är med och tar fram förtränade modeller och data. Detta är riskabelt eftersom en stor del av utvecklingen inom AI sker genom delandet av open source/data, särskilt dataset framtagna inom forskningsprojekt och förtränade modeller. Om de regleringar som kommer av den här förordningen även skulle omfatta aktörer inom akademin och/eller mer eller mindre ideella grupperingar som arbetar med öppen källkod och öppen data riskerar det att hämma den öppna oberoende utvecklingen inom maskininlärning, samtidigt som man styr utvecklingen mot proprietära betallösningar inom alla nivåer, hög-risk eller inte.

s. 32 punkt 61: Standardisering är inte lösningen för att underlätta AI-utveckling. Detta är fortfarande ett fält där man måste ha stora möjligheter att testa nya lösningar eftersom det finns väldigt mycket kvar att utveckla och upptäcka på en basal nivå. Risken finns att kraven på standardisering leder till att EU hamnar ännu längre efter USA och Kina i utvecklingen. Det stämmer inte med ett av kommissionens uttalade mål i förordningen.

Se exempelvis första sidan, sista stycket “*It supports the objective of the Union being a global leader in the development of secure, trustworthy and ethical artificial intelligence as stated by the European Council*”.

s. 34 punkt 70, sista stycket: Här finns det en risk i att man skapar lagstiftning inom EU som förutsätter att alla deep fakes mm ska märkas upp. Det kan leda till att användarna tror att allt som inte är uppmärkt är äkta. Det gör att tredje land som inte måste märka upp sina data alltid kommer att verka generera legitima bilder mm och att personer eller organisationer med illvilligt uppsåt inte kommer att märka upp manipulerade bilder och också få ett ytterligare sken av kredibilitet.

Förordningsförslagets artiklar

Title I

Article 2, (c) : Bestämmelsen medför att det blir svårare att dela dataset skapade från high-risk-system utanför EU med länder inom EU. Detta riskerar att hämma utvecklingen inom AI inom EU.

Article 3, (8): Vi tror att samlingsbegreppet ”operator” löper risk att skapa mer förvirring än vad det underlättar AI-förordningens läsbarhet och förståelsen av densamma.

Article 3, (9): Innebär detta att den organisation som först gör ett system tillgängligt ska uppbära högre kostnader för att göra det tillgängligt på inre marknaden (exempelvis ett företag i Sverige importerar och driftsätter ett system från USA och måste då stå för hela den administrativa kostnaden för att kontrollera systemet, medan andra organisationer från andra länder därefter inte behöver uppbära den kostnaden gentemot EU-kommissionen, AI advisory board och databasen där hög-risk-AI-system ska registreras).

Article 3, (16): ”lämna tillbaka AI-system” låter märkligt när det gäller IT-system. Det handlar snarare om att stänga ner systemet samt skicka tillbaka återkoppling till den som skapat den, exempelvis genom dokumentation eller dataset.

Article 3, (44): Även ekonomiska katastrofer/stor påverkan för den enskilde samt ”felaktiga” juridiska påföljder bör också klassas som ”serious incident”

Title II

Article 5, 1 (a): Man kan argumentera för att rankningsalgoritmerna i Facebook, Twitter, Snapchat m.fl. kan falla under den här kategorin, likaså under kategori 1(b).

Article 5, 1 (b): Hur bevisar man att ett AI system ”exploits vulnerabilities of a specific group of persons” mer än befolkningen i stort exempelvis. En sådan bedömning måste per definition bli subjektiv.

Article 5, 1 (c) (ii): Den här formuleringen kan uppfattas vara för subjektiv och tandlös för att skydda oppositionella i auktoritära stater, även om det ter sig som om att det delvis är det problemet som man försöker komma åt.

Title III

Chapter 2

Article 11, punkterna 1-3: Att det behövs teknisk dokumentation är otvivelaktigt då användningen av AI måste regleras eftersom datadrivna algoritmer inte är lämpliga för alla användningsområden. Men att ha just den här nivån av teknisk dokumentation som täcker de parametrar som tas upp behöver utvärderas i skarpt läge eftersom det finns en risk att man skapar en föreställning om att AI-systemen är mer säkra och kontrollerade än vad de i själva verket är. Samtidigt skapar man en oerhört kostsam kontrollapparat, från det enskilda företagets ansvariga representanter, till nationella tillsynsmyndigheter till expertmyndighet och ”board” hos

europakommissionen. Man bör således testa om den fördyring och tappade konkurrenskraft hos europeiska företag som uppkommer via den här regleringen ger någon påtaglig nytta och ökad säkerhet.

Article 12, 4 (b-c): Det är inte säkert att dessa parametrar kommer att vara applicerbara på alla system. Vidare är det inte uppenbart att de kommer att bidra i tillräckligt hög utsträckning till ökad transparens. Det kan dock skapa ett behov av att lagra enorma datamängder som kan vara oerhört känsliga för användarna om de skulle komma ut och spridas. Detta blir då viktigt att se över eventuell gallring av de loggar som skapas.

Chapter 3

Article 20, 1: Automatiskt genererade loggar som nämns här går in under samma regler som det som nämns i artikel 12, 4 (b-c) se ovan.

Article 21: Hur ska tillsynsorganen kunna kontrollera regelefterlevnaden av den här paragrafen om providers befinner sig i tredje land? Det är besvärligt att lägga ansvaret på enskilda users, distributers eller importers eftersom det ställer höga krav på kompetens.

Chapter 5

Article 50: I artikeln ställs krav om att den uppräknade dokumentationen ska hållas tillgänglig för tillsynsorganen för en period om 10 år. Att begränsa kravet på bevarande av sådan samhällsviktig information till 10 år känns väldigt snålt tilltaget. Normalt inom IT-utveckling bygger ny teknik på gammal tekning. Om kunskapen försvinner efter bara 10 år går samhällsviktig information förlorad.

Vidare saknar vi förslaget regler om hur man ska göra med dokumentationen om och när "the provider" upphör, fusioneras, köps upp, går i konkurs osv. Vem har då ansvaret för att dokumentationen bevaras och hålls tillgänglig?

Beslut i detta ärende har fattats av riksarkivarie Karin Åström Iko. Ärendet har handlagts av enhetschef Désirée Veschetti, föredragande. Vid ärendets slutliga beredning har enhetschef David Haskiya, utredare Catharina Dahlgren, samt juristen Ulrika Sturesdotter Andersson deltagit.

Karin Åström Iko

Désirée Veschetti