

Yttrande från styrelsen (art. 64)



Yttrande 28/2024 om vissa dataskyddsaspekter vid behandling av personuppgifter i samband med AI-modeller

Antaget den 17 december 2024

Translations proofread by EDPB Members.

This language version has not yet been proofread.

Sammanfattning

AI-teknik skapar många möjligheter och fördelar inom en rad olika sektorer och sociala verksamheter.

Genom att skydda den grundläggande rätten till dataskydd stöder den allmänna dataskyddsförordningen (nedan kallad *dataskyddsförordningen*) dessa möjligheter och främjar andra grundläggande rättigheter i EU, inklusive rätten till tankefrihet, yttrandefrihet och informationsfrihet samt rätten till utbildning och näringsfrihet. I så måtto är dataskyddsförordningen en rättslig ram som uppmuntrar till ansvarsfull innovation.

I detta sammanhang, och med beaktande av de dataskyddsfrågor som denna teknik ger upphov till, begärde den irländska tillsynsmyndigheten att Europeiska dataskyddsstyrelsen (EDPB) skulle avge ett yttrande om frågor av allmän räckvidd i enlighet med artikel 64.2 i dataskyddsförordningen. Begäran avser behandling av personuppgifter i samband med faserna för utveckling och driftsättning av modeller för artificiell intelligens (AI-modeller). I begäran ställdes följande frågor: 1) När och hur kan en AI-modell betraktas som "anonym"? 2) Hur kan personuppgiftsansvariga visa att berättigat intresse är en lämplig rättslig grund i utvecklingsfasen respektive 3) i driftsättningsfasen? 4) Vilka är konsekvenserna av olaglig behandling av personuppgifter i en AI-modells utvecklingsfas för efterföljande behandling eller drift av AI-modellen?

När det gäller den första frågan nämns i yttrandet att påståenden om en AI-modells anonymitet bör bedömas av behörig tillsynsmyndighet från fall till fall, eftersom EDPB anser att AI-modeller som tränats med personuppgifter inte i samtliga fall kan anses vara anonyma. För att en AI-modell ska anses vara anonym ska både 1) sannolikheten för direkt (inklusive probabilistisk) utvinning av personuppgifter tillhörande personer vars personuppgifter användes för att utveckla modellen och 2) sannolikheten för att sådana personuppgifter, avsiktligt eller inte, erhålls genom sökningar/frågor, vara obetydlig, med beaktande av alla hjälpmedel "*som rimligen kan komma att användas*" av den personuppgiftsansvarige eller någon annan person.

För att genomföra sin bedömning bör tillsynsmyndigheten granska den dokumentation som den personuppgiftsansvarige tillhandahåller för att visa att modellen är anonym. I detta avseende innehåller yttrandet en icke-bindande och icke uttömmande förteckning över metoder som kan användas av personuppgiftsansvariga för att påvisa anonymitet, och som därmed kan beaktas av tillsynsmyndigheten när den bedömer en personuppgiftsansvarigs påstående om anonymitet. Detta omfattar de metoder som de personuppgiftsansvariga använder under utvecklingsfasen för att förhindra eller begränsa insamlingen av personuppgifter som används för att träna AI-modellen, för att minska uppgifternas identifierbarhet, förhindra att de utvinns eller för att ge garantier om att den senaste tekniken för att avvärja hot och angrepp används.

När det gäller fråga 2 och fråga 3 innehåller yttrandet allmänna överväganden som tillsynsmyndigheterna bör beakta vid bedömningen av huruvida personuppgiftsansvariga kan åberopa berättigat intresse som en lämplig rättslig grund för behandling som utförs i samband med utveckling och driftsättning av AI-modeller.

I yttrandet framhålls att det inte finns någon hierarki mellan de rättsliga grunder som anges i dataskyddsförordningen, och att det är de personuppgiftsansvarigas ansvar att fastställa lämplig rättslig grund för sin behandlingsverksamhet. I yttrandet erinras sedan om det trestegstest som bör utföras vid bedömning av berättigat intresse som rättslig grund, nämligen 1) identifiering av det berättigade intresse som den personuppgiftsansvarige eller en tredje part eftersträvar, 2) analys av

behandlingens nödvändighet för det eller de berättigade intressen som eftersträvas (även kallat nödvändighetsprövning, "necessity test") och 3) bedömning av att de berättigade intressena inte åsidosätts av de registrerades intressen eller grundläggande rättigheter och friheter (även kallat avvägningstest, "balancing test").

När det gäller det första steget erinrar yttrandet om att ett intresse kan anses vara berättigat om följande tre kumulativa kriterier är uppfyllda: intresset är 1) lagligt, 2) tydligt och precist formulerat och 3) verkligt och aktuellt (dvs. inte spekulativt). Vid utvecklingsfasen av en AI-modell kan ett sådant intresse till exempel omfatta framtagande av en samtalsagent som en hjälp till användarna, och i driftsättningsfasen kan det omfatta förbättring av hotidentifieringen i ett informationssystem.

När det gäller det andra steget erinras det i yttrandet om att bedömningen av nödvändighet kräver att man överväger 1) huruvida behandlingen av personuppgifter kommer att medge att det berättigade intresset tillgodoses, och 2) om det inte finns något mindre inkräktande sätt att tillgodose detta intresse. Vid bedömningen av huruvida villkoret om nödvändighet är uppfyllt bör tillsynsmyndigheterna särskilt beakta den mängd personuppgifter som behandlas och huruvida den är proportionerlig med det berättigade intresse som eftersträvas, även mot bakgrund av principen om uppgiftsminimering.

När det gäller det tredje steget erinras det i yttrandet om att avvägningstestet bör utföras med beaktande av de särskilda omständigheterna i varje enskilt fall. Därefter ges en översikt över de faktorer som tillsynsmyndigheterna kan ta hänsyn till när de bedömer om en personuppgiftsansvarigs eller tredje parts intresse väger tyngre än de registrerades intressen, grundläggande rättigheter och friheter.

Som en del av det tredje steget belyser yttrandet särskilda risker med avseende på de grundläggande rättigheterna som kan uppstå under utvecklings- eller driftsättningsfasen av AI-modeller. Det klargörs också att den behandling av personuppgifter som sker under utvecklingen och driftsättningen av AI-modeller kan påverka de registrerade på olika sätt, såväl positivt som negativt. För att bedöma denna påverkan kan tillsynsmyndigheten överväga vilken typ av uppgifter som behandlas av modellerna, sammanhanget för behandlingen och de möjliga ytterligare konsekvenserna av behandlingen.

Yttrandet belyser dessutom den roll som de registrerades rimliga förväntningar spelar i avvägningstestet. Detta kan vara viktigt på grund av komplexiteten hos den teknik som används i AI-modeller och det faktum att det kan vara svårt för registrerade att förstå de olika möjliga användningarna av dem, liksom de olika behandlingsaktiviteter som detta inbegriper. I detta avseende kan både den information som lämnas till de registrerade och sammanhanget för behandlingen vara faktorer som bör beaktas för att bedöma om de registrerade rimligen kan förvänta sig att deras personuppgifter behandlas. När det gäller sammanhanget kan detta omfatta följande: huruvida personuppgifterna var offentligt tillgängliga eller inte, typen av förhållande mellan den registrerade och den personuppgiftsansvarige (och huruvida det finns en koppling mellan de två), tjänstens art, det sammanhang i vilket personuppgifterna samlades in, källan från vilken uppgifterna samlades in (dvs. den webbplats eller tjänst som personuppgifterna hämtades från och de integritetsinställningar som de erbjuder), modellens potentiella ytterligare användningar och huruvida de registrerade över huvud taget är medvetna om att deras personuppgifter finns online.

I yttrandet erinras också om att den personuppgiftsansvarige, när de registrerades intressen, rättigheter och friheter förefaller åsidosätta de berättigade intressen som den personuppgiftsansvarige eller en tredje part har, kan överväga att införa riskreducerande åtgärder för att begränsa behandlingens inverkan på dessa registrerade. Riskreducerande åtgärder bör inte förväxlas med de åtgärder som den personuppgiftsansvarige enligt lag är skyldig att vidta för att

säkerställa efterlevnad av dataskyddsförordningen. Åtgärderna bör dessutom vara anpassade till omständigheterna i det enskilda fallet och AI-modellens egenskaper, inklusive dess avsedda användning. I detta avseende innehåller yttrandet en icke uttömmande förteckning över exempel på riskreducerande åtgärder i samband med utvecklingsfasen (även när det gäller webbinsamling) och driftsättningsfasen. De riskreducerande åtgärderna bör anpassas till omständigheterna i det enskilda fallet och kan behöva förändras. Därför bör tillsynsmyndigheterna bedöma lämpligheten hos de riskreducerande åtgärder som införts från fall till fall.

När det gäller fråga 4 framhålls det allmänt i yttrandet att tillsynsmyndigheterna har befogenhet att göra en skönsmässig bedömning av möjliga överträdelser och välja lämpliga, nödvändiga och proportionella åtgärder med beaktande av omständigheterna i varje enskilt fall. Därefter behandlas tre scenarier i yttrandet.

Scenario 1 bygger på att personuppgifterna kvarhålls i AI-modellen (vilket innebär att modellen inte kan anses vara anonym, såsom anges under fråga 1) och därefter behandlas av samma personuppgiftsansvarige (till exempel i samband med driftsättningen av modellen). I yttrandet anges att det i varje enskilt fall – och beroende på sammanhanget i detta enskilda fall – bör bedömas huruvida utvecklingsfasen och driftsättningsfasen har skilda ändamål (och därmed utgör separata behandlingsaktiviteter) och i vilken utsträckning avsaknaden av rättslig grund för den inledande behandlingsaktiviteten påverkar lagligheten av den efterföljande behandlingen.

I scenario 2 kvarhålls personuppgifterna i modellen och behandlas av en annan personuppgiftsansvarig i samband med driftsättningen av modellen. I detta avseende anges i yttrandet att tillsynsmyndigheterna bör ta hänsyn till huruvida den personuppgiftsansvarige som använder modellen har utfört en lämplig bedömning som en del av sin ansvarsskyldighet för att visa efterlevnad av artikel 5.1 a och artikel 6 i dataskyddsförordningen, för att säkerställa att AI-modellen inte togs fram med hjälp av olaglig behandling av personuppgifter. Denna bedömning bör ta hänsyn till exempelvis personuppgifternas ursprung och huruvida behandlingen under utvecklingsfasen var föremål för en konstaterad överträdelse, särskilt om den fastställdes av en tillsynsmyndighet eller en domstol, och bedömningens grad av detaljrikedom bör styras av de risker som behandlingen under driftsättningsfasen medför.

I scenario 3 behandlar en personuppgiftsansvarig olagligen personuppgifter för att utveckla AI-modellen och ser sedan till att de anonymiseras. Därefter påbörjar samma eller en annan personuppgiftsansvarig en annan behandling av personuppgifter i samband med driftsättningen. I detta avseende anges i yttrandet att EDPB anser att om det kan visas att den efterföljande driften av AI-modellen inte medför behandling av personuppgifter, bör dataskyddsförordningen inte vara tillämplig. Att den inledande behandlingen varit olaglig bör därför inte påverka den efterföljande användningen av modellen. Vidare anser EDPB att när personuppgiftsansvariga behandlar personuppgifter efter det att de samlats in under driftsättningsfasen och efter det att modellen har anonymiserats, bör dataskyddsförordningen vara tillämplig på denna behandling. När det gäller dataskyddsförordningen anser EDPB i dessa fall att lagligheten av den behandling som utförs i driftsättningsfasen inte bör påverkas av att den inledande behandlingen var olaglig.

Innehåll

1	Inledning.....	6
1.1	Sammanfattning av omständigheterna	6
1.2	Tillåtlighet för begäran om ett yttrande enligt artikel 64.2 i dataskyddsförordningen	8
2	Omfattning och viktiga begrepp	9
2.1	Yttrandets omfattning.....	9
2.2	Nyckelbegrepp	11
2.3	AI-modeller inom ramen för detta yttrande.....	12
3	Huruvida begäran är välgrundad	13
3.1	Om AI-modellernas egenskaper i förhållande till definitionen av personuppgifter	13
3.2	Om de omständigheter under vilka AI-modeller kan betraktas som anonyma och tillhörande bevis.....	14
3.2.1	Allmänna överväganden om anonymisering i det aktuella sammanhanget	15
3.2.2	Element för att utvärdera den kvarvarande sannolikheten för identifiering	17
3.3	Om huruvida ett berättigat intresse är en lämplig rättslig grund för behandling av personuppgifter i samband med utveckling och driftsättning av AI-modeller.....	20
3.3.1	Allmänna synpunkter	20
3.3.2	Överväganden avseende de tre stegen i bedömningen av berättigat intresse i samband med utvecklingen och driftsättningen av AI-modeller	22
3.4	Om den möjliga inverkan som olaglig behandling vid utvecklingen av en AI-modell kan ha på lagligheten av den efterföljande behandlingen eller användningen av AI-modellen	32
3.4.1	Fall 1. En personuppgiftsansvarig behandlar olagligen personuppgifter för att utveckla modellen. Personuppgifterna kvarhålls i modellen och behandlas därefter av samma personuppgiftsansvarige (till exempel i samband med att modellen tas i bruk).	34
3.4.2	Fall 2. En personuppgiftsansvarig behandlar olagligt personuppgifter för att utveckla modellen. Personuppgifterna sparas i modellen och behandlas sedan av en annan personuppgiftsansvarig i samband med driftsättningen av modellen.	34
3.4.3	Fall 3. En personuppgiftsansvarig behandlar olagligen personuppgifter för att utveckla modellen och ser sedan till att modellen anonymiseras innan samma eller en annan personuppgiftsansvarig påbörjar en annan behandling av personuppgifter i samband med driftsättningen.	36
4	Avslutande anmärkningar	37

Europeiska dataskyddsstyrelsen har antagit detta yttrande

med beaktande av artiklarna 63 och 64.2 i 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 (nedan kallad dataskyddsförordningen),

med beaktande av EES-avtalet, särskilt bilaga XI och protokoll 37, ändrat genom gemensamma EES-kommitténs beslut nr 154/2018 av den 6 juli 2018¹,

med beaktande av artikel 10 och artikel 22 i arbetsordningen, och

av följande skäl:

1) Den viktigaste uppgiften för Europeiska dataskyddsstyrelsen (nedan kallad *dataskyddsstyrelsen* eller EDPB) är att säkerställa en enhetlig tillämpning av dataskyddsförordningen inom hela Europeiska ekonomiska samarbetsområdet (EES). I artikel 64.2 i dataskyddsförordningen föreskrivs att varje tillsynsmyndighet, styrelsens ordförande eller kommissionen, i syfte att utarbeta ett yttrande får begära att en fråga med allmän räckvidd eller som har följder i mer än en EES-medlemsstat ska granskas av dataskyddsstyrelsen. Syftet med detta yttrande är att granska en fråga med allmän räckvidd eller som har följder i mer än en EES-medlemsstat.

2) Styrelsens yttrande ska antas i enlighet med artikel 64.3 i dataskyddsförordningen, jämförd med artikel 10.2 i dataskyddsstyrelsens arbetsordning, inom åtta veckor från det datum då ordföranden och den behöriga tillsynsmyndigheten har beslutat att handlingarna i ärendet är fullständiga. Efter beslut av ordförande får denna period förlängas med ytterligare sex veckor med hänsyn till sakfrågans komplexitet.

HÄRIGENOM FRAMFÖRS FÖLJANDE.

1 Inledning

1.1 Sammanfattning av omständigheterna

1. Den 4 september 2024 begärde den irländska tillsynsmyndigheten (nedan kallad *den begärande myndigheten*) att dataskyddsstyrelsen skulle avge ett yttrande i enlighet med artikel 64.2 i dataskyddsförordningen avseende AI-modeller och behandling av personuppgifter (nedan kallad *begäran*).
2. Styrelsens ordförande och den irländska tillsynsmyndigheten ansåg att handlingarna i ärendet var fullständiga den 13 september 2024. Följande arbetsdag, den 16 september 2024, sändes filen av dataskyddsstyrelsens sekretariat. Med hänsyn till ärendets komplexitet beslutade ordföranden att förlänga den rättsliga tidsfristen i enlighet med artikel 64.3 i dataskyddsförordningen och artikel 10.4 i EDPB:s arbetsordning.

¹ De hänvisningar till "medlemsstater" som görs i hela detta yttrande ska förstås som hänvisningar till "EES-medlemsstater". De hänvisningar till "unionen" som görs i detta yttrande ska förstås som hänvisningar till "EES".

3. Begäran avser vissa delar av träningen, uppdateringen, utvecklingen och driften av AI-modeller där personuppgifter utgör en del av det relevanta datasetet. Den irländska tillsynsmyndigheten betonar att begäran rör viktiga frågor som har stor inverkan på registrerade och personuppgiftsansvariga i EES, och att det i detta skede inte finns något harmoniserat synsätt bland de nationella tillsynsmyndigheterna². Den terminologi som används i detta yttrande anges i avsnitt 2.2 och 2.3 nedan.

4. Följande frågor ställdes av den irländska tillsynsmyndigheten:

Fråga 1: Anses den slutliga AI-modellen, som har tränats med hjälp av personuppgifter, i samtliga fall inte motsvara definitionen av personuppgifter (enligt artikel 4.1 i dataskyddsförordningen)?

Om svaret på fråga 1 är "ja":

- i. I vilket skede av den behandlingsverksamhet som leder fram till en AI-modell behandlas inte längre personuppgifter?
 - a) Hur kan det bevisas att AI-modellen inte behandlar personuppgifter?
- ii. Finns det några faktorer som skulle leda till att driften av den slutliga AI-modellen inte längre anses vara anonym?
 - a) Om så är fallet, hur kan de åtgärder som vidtagits för att reducera riskerna eller förhindra eller skydda mot dessa faktorer (för att säkerställa att AI-modellen inte behandlar personuppgifter) bevisas?

Om svaret på fråga 1 är "nej":

- i. Vilka är omständigheterna under vilka detta kan uppstå?
 - a) Om så är fallet, hur kan man bevisa vilka åtgärder som har vidtagits för att säkerställa att AI-modellen inte behandlar personuppgifter?

Fråga 2: Om en personuppgiftsansvarig förlitar sig på berättigade intressen som rättslig grund för behandling av personuppgifter för att skapa, uppdatera och/eller utveckla en AI-modell, hur bör den personuppgiftsansvarige visa att de berättigade intressena är lämpliga som rättslig grund, både när det gäller behandlingen av uppgifter från tredje part och första part?

- i. Vilka överväganden bör denna personuppgiftsansvarige i detta fall beakta för att säkerställa att intressena hos de registrerade vars personuppgifter behandlas på lämpligt sätt balanseras mot den personuppgiftsansvariges intressen i samband med
 - a) uppgifter från tredje part, respektive
 - b) uppgifter från första part?

Fråga 3: Om en personuppgiftsansvarig efter träningen av AI-modellen förlitar sig på berättigade intressen som rättslig grund för den behandling av personuppgifter som sker inom AI-modellen eller inom ett AI-system som en AI-modell utgör en del av, hur ska en personuppgiftsansvarig visa att de berättigade intressena är en lämplig rättslig grund?

² Begäran, s. 1.

Fråga 4: Om det har konstaterats att en AI-modell har skapats, uppdaterats eller utvecklats med hjälp av olagligt behandlade personuppgifter, vilken inverkan har detta, om någon, på lagligheten av den fortsatta eller efterföljande behandlingen eller driften av AI-modellen, antingen i sig eller som en del av ett AI-system, där

- i. AI-modellen, antingen ensam eller som en del av ett AI-system, behandlar personuppgifter?
- ii. varken AI-modellen i sig eller AI-modellen som en del av ett AI-system behandlar personuppgifter?

1.2 Tillåtlighet för begäran om ett yttrande enligt artikel 64.2 i dataskyddsförordningen

5. I artikel 64.2 i dataskyddsförordningen föreskrivs specifikt att varje tillsynsmyndighet, i syfte att erhålla ett yttrande, får begära att en fråga med allmän räckvidd eller som har följder i mer än en medlemsstat ska granskas av dataskyddsstyrelsen.
6. Den begärande tillsynsmyndigheten ställde frågor till EDPB om dataskyddsaspekter i samband med AI-modeller. I begäran specificerades att även om många organisationer nu använder AI-modeller, inbegripet stora språkmodeller (LLM), ger deras verksamhet, träning och användning upphov till *"ett antal omfattande dataskyddsproblem"*³, som *"påverkar registrerade i hela EU/EES"*⁴.
7. I begäran ställs i huvudsak frågor om i) tillämpningen av begreppet personuppgifter, ii) principen om laglighet, med särskilt beaktande av den rättsliga grunden för berättigat intresse, i samband med AI-modeller, samt om iii) konsekvenserna vid olaglig behandling av personuppgifter i AI-modellers utvecklingsfas för den efterföljande behandlingen eller användningen av modellen.
8. Dataskyddsstyrelsen anser att begäran avser en *fråga med allmän räckvidd* i den mening som avses i artikel 64.2 i dataskyddsförordningen. Frågan gäller särskilt tolkningen och tillämpningen av artiklarna 4.1, 5.1 a och 6 i dataskyddsförordningen när det gäller behandling av personuppgifter i samband med utveckling och driftsättning av AI-modeller. Som den begärande tillsynsmyndigheten påpekat väcker tillämpningen av dessa bestämmelser på AI-modeller nya, systemiska och abstrakta frågor⁵. Den snabba utvecklingen och användningen av AI-modeller av allt fler organisationer ger upphov till särskilda frågor och som påpekas i begäran *"kommer det att vara till stor nytta för EDPB om en gemensam ståndpunkt kan nås när det gäller de frågor som tas upp i begäran, eftersom dessa frågor är centrala för EDPB:s planerade arbete på kort och medellång sikt"*⁶. Vidare skapar AI-tekniken många möjligheter och fördelar inom en rad olika sektorer och sociala verksamheter. Dessutom är dataskyddsförordningen en rättslig ram som främjar ansvarsfull innovation. Av detta följer att det finns ett allmänt intresse av att göra denna bedömning i form av ett yttrande från EDPB, för att säkerställa en enhetlig tillämpning av vissa bestämmelser i dataskyddsförordningen när det gäller AI-modeller.
9. Det andra villkoret som nämns i artikel 64.2 i dataskyddsförordningen avser frågor *"som har följder i mer än en medlemsstat"*. EDPB erinrar om att begreppet *"följder"* ska tolkas i vid bemärkelse och

³ Begäran, s. 1.

⁴ Se föregående fotnot.

⁵ Begäran, s. 2.

⁶ Begäran, s. 1. Såsom anges i EDPB:s arbetsprogram för 2024–2025, som antogs den 8 oktober 2024 och som finns tillgängligt på https://www.edpb.europa.eu/system/files/2024-10/edpb_work_programme_2024-2025_en.pdf, planerar EDPB bland annat att utfärda riktlinjer om anonymisering, pseudonymisering och webbinsamling (dataskrapning) i samband med generativ AI.

således inte är begränsat till enbart rättsverkningar⁷. Allt eftersom fler AI-modeller tränas och används av ett växande antal organisationer inom EES påverkar de ett allt större antal registrerade i hela EES, varav vissa redan har framfört farhågor till sin behöriga tillsynsmyndighet⁸. EDPB anser därför att den fråga som tagits upp av den begärande tillsynsmyndigheten även uppfyller detta villkor.

10. Begäran innehåller ett skriftligt resonemang rörande bakgrunden och orsakerna till att dessa frågor ställs till dataskyddsstyrelsen, inklusive det relevanta rättsliga ramverket. Därför anser dataskyddsstyrelsen att begäran är motiverad i enlighet med artikel 10.3 i EDPB:s arbetsordning.
11. Enligt artikel 64.3 i dataskyddsförordningen⁹ ska EDPB inte avge ett yttrande om den redan har avgett ett yttrande i ärendet. EDPB har inte avgett något yttrande i detta ärende och har ännu inte besvarat de frågor som begäran innehåller.
12. Av dessa skäl anser dataskyddsstyrelsen att begäran är tillåtlig och att de frågor som uppstår till följd av den bör analyseras i detta yttrande (nedan kallat *yttrandet*) som antas i enlighet med artikel 64.2 i dataskyddsförordningen.

2 Omfattning och viktiga begrepp

2.1 Yttrandets omfattning

13. Dataskyddsstyrelsen håller med den begärande tillsynsmyndigheten om att utvecklingen och användningen av AI-modeller, ur ett dataskyddsperspektiv, ger upphov till grundläggande dataskyddsfrågor. Frågorna gäller särskilt följande: i) När och hur kan en AI-modell betraktas som "anonym" (fråga 1 i begäran)? ii) Hur kan personuppgiftsansvariga visa att berättigat intresse är en lämplig rättslig grund i utvecklingsfasen (fråga 2 i begäran) respektive i driftsättningsfasen (fråga 3 i begäran)? iii) Vilka är konsekvenserna av olaglig behandling av personuppgifter i en AI-modells utvecklingsfas för lagligheten i efterföljande behandling eller drift av AI-modellen (fråga 4 i begäran)?
14. EDPB påminner om att tillsynsmyndigheterna är ansvariga för att övervaka tillämpningen av dataskyddsförordningen och bör bidra till dess enhetliga tillämpning i hela unionen¹⁰. Det ingår därför i de tillsynsmyndigheternas ansvarsområde att undersöka specifika AI-modeller och bedöma varje enskilt fall individuellt.
15. Detta yttrande tillhandahåller en ram som tillsynsmyndigheterna kan vägledas av när de står i begrepp att bedöma specifika fall där en eller flera av de frågor som tas upp i begäran uppstår. Syftet med detta yttrande är inte att vara uttömmande, utan snarare att tillhandahålla allmänna överväganden om tolkningen av de relevanta bestämmelserna, vilka i hög grad bör beaktas av de behöriga tillsynsmyndigheterna när de använder sina utredningsbefogenheter. Detta yttrande riktar sig till behöriga tillsynsmyndigheter och avser deras verksamhet och befogenheter, men det påverkar inte personuppgiftsansvarigas och personuppgiftsbiträdens skyldigheter enligt dataskyddsförordningen. I synnerhet ska personuppgiftsansvariga, i enlighet med den princip om ansvarsskyldighet som fastställs

⁷ EDPB:s interna dokument 3/2019 avseende intern vägledning om artikel 64.2 i dataskyddsförordningen, antaget den 8 oktober 2019, punkt 15, tillgängligt på https://www.edpb.europa.eu/system/files/2022-07/internaledpb_document_201903_art64.2_en.pdf.

⁸ Begäran, s. 1–2.

⁹ Artikel 64.3 i dataskyddsförordningen och artikel 10.4 i EDPB:s arbetsordning.

¹⁰ Artikel 51.1 och artikel 51.2 i dataskyddsförordningen.

i artikel 5.2 i dataskyddsförordningen, vara ansvariga för och kunna visa att de följer alla principer som rör deras behandling av personuppgifter.

16. I vissa fall ges exempel i yttrandet, men med tanke på den breda omfattningen av de frågor som ingår i begäran och de olika typer av AI-modeller som omfattas, kan inte alla potentiella scenarier beaktas i yttrandet. Tekniken som används i AI-modeller utvecklas snabbt, vilket bör beaktas när EDPB:s överväganden i detta yttrande tolkas.
17. **I detta yttrande analyseras inte bestämmelserna nedan, vilka dock ändå kan spela en viktig roll vid bedömningen av de dataskyddskrav som är tillämpliga på AI-modeller.**
- **Behandling av särskilda kategorier av uppgifter:** Dataskyddsstyrelsen påminner om förbudet i artikel 9.1 i dataskyddsförordningen mot behandling av särskilda kategorier av uppgifter och de begränsade undantagen i artikel 9.2 i samma förordning¹¹. I detta hänseende har Europeiska unionens domstol (EU-domstolen) vidare klargjort att *”då en datamängd med personuppgifter som innehåller både känsliga och icke-känsliga uppgifter samlas in som en helhet, utan att uppgifterna kan skiljas från varandra vid tidpunkten för insamlingen, ska behandling av denna datamängd anses vara förbjuden, i den mening som avses i artikel 9.1 i GDPR, om den innehåller åtminstone en känslig uppgift och inte något av undantagen i artikel 9.2 i GDPR är tillämpligt”*¹². EU-domstolen betonade vidare att det *”vid tillämpningen av undantaget i artikel 9.2 e i GDPR ska kontrolleras huruvida den registrerade, uttryckligen och genom en entydig aktiv handling, avsett att göra de aktuella personuppgifterna tillgängliga för allmänheten”*¹³. Dessa överväganden bör beaktas när behandling av personuppgifter inom ramen för AI-modeller omfattar särskilda kategorier av uppgifter.
 - **Automatiserat beslutsfattande, inbegripet profilering:** De behandlingar som utförs i samband med AI-modeller kan omfattas av artikel 22 i dataskyddsförordningen, som ålägger personuppgiftsansvariga ytterligare skyldigheter och inför ytterligare skyddsåtgärder för att värna om de registrerade. EDPB erinrar i detta avseende om sina riktlinjer om automatiserat individuellt beslutsfattande och profilering enligt förordning (EU) 2016/679¹⁴.
 - **Förenlighet i fråga om ändamål:** I artikel 6.4 i dataskyddsförordningen föreskrivs – för vissa rättsliga grunder – kriterier som en personuppgiftsansvarig ska beakta för att fastställa huruvida behandling för andra ändamål är förenlig med det ändamål för vilket personuppgifterna ursprungligen samlades in. Denna bestämmelse kan vara relevant i samband med utveckling och driftsättning av AI-modeller, och dess tillämplighet bör bedömas av tillsynsmyndigheterna.

¹¹ Se även punkt 18 i EDPB:s rapport om det arbete som utförts av dess arbetsgrupp om ChatGPT, antagen den 23 maj 2024: *”När det gäller behandling av särskilda kategorier av personuppgifter måste dessutom något av undantagen i artikel 9.2 vara tillämpligt för att behandlingen ska vara laglig. Ett av dessa undantag kan i princip vara artikel 9.2 e i dataskyddsförordningen. Enbart det faktum att personuppgifter är allmänt tillgängliga innebär dock inte att ”den registrerade uppenbart har offentliggjort sådana uppgifter [...]”*.

¹² EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 89.

¹³ EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 77.

¹⁴ Artikel 29-arbetsgruppens riktlinjer för automatiserat individuellt beslutsfattande och profilering enligt förordning (EU) 2016/679, senast reviderade och antagna den 6 februari 2018 och godkända av EDPB den 25 maj 2018. Se även EU-domstolens dom av den 7 december 2023, mål C-634/21, *SCHUFA Holding m.fl.* (ECLI:EU:C:2023:957).

- **Konsekvensbedömningar avseende dataskydd** (artikel 35 i dataskyddsförordningen): Konsekvensbedömningar avseende dataskydd är en viktig del av ansvarsskyldigheten, där behandlingen i samband med AI-modeller sannolikt kommer att innebära en hög risk för fysiska personers rättigheter och friheter¹⁵.
- **Princip om inbyggt dataskydd** (artikel 25.1 i dataskyddsförordningen): Inbyggt dataskydd är en viktig skyddsåtgärd som ska bedömas av tillsynsmyndigheterna i samband med att en AI-modell utvecklas och driftsätts.

2.2 Nyckelbegrepp

18. Som en inledande anmärkning vill EDPB förtydliga de termer och begrepp som används i detta yttrande och för de syften som avses i detsamma:

- **"Uppgifter från första part"** avser de personuppgifter som den personuppgiftsansvarige har samlat in från de registrerade.
- **"Uppgifter från tredje part"** avser personuppgifter som personuppgiftsansvariga inte har erhållit från de registrerade, utan har samlat in eller mottagit från en tredje part, till exempel från en datamäklare eller genom webbinsamling.
- **"Webbinsamling"** (även kallad webbskrapning) är en vanligt förekommande teknik för att samla in information från allmänt tillgängliga källor på internet. Exempelvis kan det ingå personuppgifter i information som hämtas från tjänster som nyhetsbyråer, sociala medier, forumdiskussioner och personliga webbplatser.
- I begäran hänvisas till **AI-modellers "livscykel"** samt till dess olika stadier, däribland "skapande", "utveckling", "träning", "uppdatering", "finjustering", "drift" eller "efterhandsträning" av AI-modeller. EDPB konstaterar att sådana steg, beroende på omständigheterna, kan förekomma i utvecklingen och driftsättningen av AI-modeller och kan inbegripa behandling av personuppgifter för olika behandlingsändamål. Inom ramen för detta yttrande anser dock EDPB att det är viktigt att inrikta sig på de stadier som med störst sannolikhet kommer att inträffa. I detta yttrande hänvisar EDPB därför till **utvecklingsfasen** och **driftsättningsfasen**. Fasen för utveckling av en AI-modell inbegriper alla stadier som föregår driftsättningen av AI-modellen och omfattar bland annat kodutveckling, insamling och förbehandling av de personuppgifter som ska användas för att träna modellen samt själva träningen av modellen. Driftsättningen av en AI-modell omfattar alla faser som rör användningen av AI-modellen och kan inkludera all verksamhet som bedrivs efter utvecklingsfasen. EDPB är medveten om de många olika användningsscenarier som finns och deras potentiella konsekvenser när det gäller behandling av personuppgifter. Därför bör tillsynsmyndigheterna överväga om iakttagelserna i detta yttrande är relevanta för den behandling som de ska bedöma.
- EDPB framhåller också att termen **"träning"** avser den del av utvecklingsfasen där AI-modeller använder data för att "lära sig" att utföra sin avsedda uppgift (vilket förklaras i nästa avsnitt i detta yttrande).

¹⁵ Artikel 29-arbetsgruppens riktlinjer om konsekvensbedömning avseende dataskydd och fastställande av huruvida behandlingen "sannolikt leder till en hög risk" i den mening som avses i förordning 2016/679, reviderade och antagna den 4 oktober 2017, godkända av EDPB den 25 maj 2018.

- Begreppet och omfattningen av **AI-modeller**, såsom det förstås av EDPB inom ramen för detta yttrande, anges närmare i följande särskilda avsnitt.

2.3 AI-modeller inom ramen för detta yttrande

19. I EU:s förordning om artificiell intelligens (nedan kallad AI-förordningen)¹⁶ definieras ett AI-system som *"ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som kan uppvisa anpassningsförmåga efter införande och som, för uttryckliga eller underförstådda mål, drar slutsatser härledda från den indata det tar emot, om hur utdata såsom förutsägelser, innehåll, rekommendationer eller beslut som kan påverka fysiska eller virtuella miljöer ska genereras"*¹⁷. I skäl 12 i AI-förordningen ges en mer ingående förklaring av begreppet AI-system. Av detta följer att en viktig egenskap hos AI-system är deras slutledningsförmåga. De tekniker som möjliggör slutledning i ett AI-system omfattar maskininlärning samt logik- och kunskapsbaserade metoder.
20. "AI-modeller" definieras å sin sida endast indirekt i AI-förordningen: *"Även om AI-modeller är väsentliga komponenter i AI-system utgör de inte AI-system i sig. AI-modeller kräver tillägg av ytterligare komponenter, till exempel ett användargränssnitt, för att bli AI-system. AI-modeller är vanligtvis integrerade i och utgör en del av AI-system"*¹⁸.
21. EDPB tolkar definitionen av AI-modeller som föreslås i begäran som snävare än den i AI-förordningen, eftersom den beskriver AI-modeller som *"omfattande den produkt som är resultatet av de träningsmekanismer som tillämpats på en uppsättning träningsdata, i samband med artificiell intelligens, maskininlärning, djupinlärning eller andra relaterade behandlingssammanhang"* och vidare specificerar att *"Termen gäller både för AI-modeller som är avsedda att genomgå ytterligare träning, finjustering och/eller utveckling, samt AI-modeller som inte är det."*¹⁹
22. EDPB antog därför detta yttrande på grundval av antagandet att ett AI-system använder sig av en AI-modell för att uppnå sitt avsedda mål genom att införliva modellen i en större ram (ett AI-system för kundtjänst kan till exempel använda en AI-modell som tränats med hjälp av samtalshistorik för att ge svar på användarnas frågor).
23. Vidare är de AI-modeller (nedan även kallade *modeller*) som är relevanta för detta yttrande sådana som utvecklats genom en träningsprocess. En sådan träningsprocess är en del av utvecklingsfasen, där modellerna använder data för att lära sig att utföra sin avsedda uppgift. För träningsprocessen krävs det därför att det finns tillgång till ett dataset som modellen kan använda för att identifiera mönster och "lära sig" utifrån dessa. I dessa fall kommer modellen att använda olika tekniker för att kartlägga den kunskap som erhållits från träningsdatasetet. Detta är särskilt fallet vid maskininlärning.
24. I praktiken är varje AI-modell en algoritm vars funktion bestäms av en uppsättning element. Djupinlärningsmodeller utgörs till exempel ofta av neurala nätverk (även kallade neuronnät) med flera lager som är anslutna via samarbetande neuroner (eller noder) med en eller flera "ingångar" som var och en har en vikt. Dessa vikter ändras vid inlärning, vilket sker genom en koppling från utdatan av en neuron till indatan av en annan neuron. En enkel djupinlärningsmodell skulle ha följande egenskaper: i) typ av och storlek på varje lager, ii) den vikt som tillskrivs varje utgång (ibland kallade "parametrar"),

¹⁶ 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 (EU:s förordning om artificiell intelligens).

¹⁷ Artikel 3.1 i AI-förordningen.

¹⁸ Skäl 97 i AI-förordningen.

¹⁹ Begäran, s. 3.

iii) aktiveringsfunktionerna²⁰ mellan lagren, och eventuellt iv) andra processer som kan äga rum mellan lagren. När man till exempel tränar en enkel djupinlärningsmodell för bildklassificering kommer indata (bildpixlar) att kopplas till utdata, och vikterna kan justeras så att rätt utdata oftast erhålls.

25. Andra exempel på djupinlärningsmodeller är stora språkmodeller (LLM) och generativ AI, som används för att t.ex. skapa nya data och innehåll som liknar människoskapat innehåll.
26. **På grundval av ovanstående överväganden är tillämpningsområdet för detta yttrande, i linje med begäran, endast den undergrupp av AI-modeller som tränats med personuppgifter.**

3 Huruvida begäran är välgrundad

3.1 Om AI-modellernas egenskaper i förhållande till definitionen av personuppgifter

27. I artikel 4.1 i dataskyddsförordningen definieras personuppgifter som *"varje upplysning som avser en identifierad eller identifierbar fysisk person"* (dvs. en registrerad). Vidare föreskrivs i skäl 26 i dataskyddsförordningen att dataskyddsprinciperna inte ska gälla för anonym information, dvs. information som inte avser en identifierad eller identifierbar fysisk person, med beaktande av alla hjälpmedel som *"rimligen kan komma att användas"* av den personuppgiftsansvarige eller av en annan person. Detta innefattar i) uppgifter som aldrig avsett en identifierad eller identifierbar enskild person, och ii) personuppgifter som har anonymiserats på ett sådant sätt att den registrerade inte längre är identifierbar.
28. Fråga 1²¹ i begäran kan således besvaras genom en analys av om en AI-modell som tränats genom behandling av personuppgifter i samtliga fall bör anses vara anonym. På grundval av formuleringen i frågan använder EDPB begreppet "träning" (av AI-modeller) i detta avsnitt.
29. Till att börja med vill EDPB framföra följande allmänna överväganden. Oavsett om de tränats med personuppgifter eller inte är AI-modeller vanligtvis utformade för att göra prognoser/förutsägelser eller dra slutsatser, dvs. de är utformade för att ha slutledningsförmåga. AI-modeller som tränats med personuppgifter är dessutom ofta utformade för att dra slutsatser om andra individer än de vars personuppgifter använts för att träna AI-modellen. Vissa AI-modeller är dock specifikt utformade för att tillhandahålla personuppgifter för enskilda personer vars personuppgifter använts för att träna modellen, eller på något sätt göra sådana uppgifter tillgängliga. I dessa fall kommer sådana AI-modeller på grund av sin natur (och som regel nödvändigtvis) att innehålla information som rör en identifierad eller identifierbar fysisk person, och kommer därmed att medföra behandling av personuppgifter. Därför kan dessa typer av AI-modeller inte betraktas som anonyma. Detta skulle till exempel vara fallet med i) en generativ modell som finjusterats med hjälp av en individs röstinspelningar för att efterlikna personens röst, eller ii) modeller som utformats för att svara med personuppgifter som den tränats med när den uppmanas att lämna information om en viss person.
30. Mot bakgrund av ovanstående överväganden inriktar sig EDPB i sitt svar på fråga 1 i begäran på AI-modeller som inte är utformade för att tillhandahålla personuppgifter relaterade till de data som den tränats med.

²⁰ Det vill säga funktioner som utifrån indata och vikter beräknar en neurons utdata, vilken sedan överförs till nästa lager i det neurala nätverket.

²¹ "Ska det i samtliga fall anses att slutliga AI-modeller som har tränats med hjälp av personuppgifter inte uppfyller definitionen av personuppgifter (enligt artikel 4.1 i dataskyddsförordningen)?"

31. EDPB anser att även om en AI-modell inte avsiktligt har utformats för att alstra information som rör en identifierad eller identifierbar fysisk person från träningsuppgifterna, kan information från träningsuppgifterna, inklusive personuppgifter, fortfarande finnas upptagna i modellens parametrar i form av matematiska objekt. De kan skilja sig från de ursprungliga träningsdatapunkterna, men kan fortfarande behålla den ursprungliga information som ingick i dessa data, som i slutändan kan extraheras eller på annat sätt erhållas, direkt eller indirekt, från modellen. Om information som rör identifierade eller identifierbara personer vars personuppgifter använts för att träna en AI-modell kan erhållas från denna modell med hjälpmedel som rimligen kan komma att användas, kan man dra slutsatsen att en sådan modell inte är anonym.
32. I detta avseende anges i begäran att *”Befintliga forskningspublikationer belyser vissa potentiella sårbarheter som kan finnas i AI-modeller och som kan leda till att personuppgifter behandlas²² samt till att behandling av personuppgifter kan ske när modeller driftsätts för användning med andra uppgifter, antingen genom gränssnitt för applikationsprogrammering (”API”) eller prompt-baserade gränssnitt²³.”*
33. På samma sätt är forskningen på området insamling av träningsdata särskilt dynamisk²⁴. Den visar att det i vissa fall är möjligt att använda medel som rimligen kan tänkas utvinna personuppgifter från vissa AI-modeller, eller att oavsiktligt få tillgång till personuppgifter genom interaktion med en AI-modell (till exempel som en del av ett AI-system). Kontinuerliga forskningsinsatser inom detta område kommer att bidra till att ytterligare bedöma de kvarstående riskerna med återflöde/återskapande²⁵ och utvinning av personuppgifter i varje enskilt fall.
34. **På grundval av ovanstående överväganden anser EDPB att AI-modeller som tränats med personuppgifter inte i samtliga fall kan anses vara anonyma. I stället bör fastställandet av huruvida en AI-modell är anonym bedömas från fall till fall, baserat på specifika kriterier.**

3.2 Om de omständigheter under vilka AI-modeller kan betraktas som anonyma och tillhörande bevis

²² Till exempel Membership Inference Attacks ([OWASP](#)) och Model Inversion Attacks ([OWASP](#) och [Veale m.fl., 2018](#)).

²³ Begäran, s. 1–2.

²⁴ Se till exempel följande i detta avseende: i) Veale M., Binns R., Edwards L., 2018, *Algorithms that remember: model inversion attacks and data protection law*. Phil. Trans. R. Soc. A 376: 20180083, tillgänglig på <http://dx.doi.org/10.1098/rsta.2018.0083>; ii) Brown H., Lee K., Mireshghallah F., Shokri R., och Tramèr F., *What Does it Mean for a Language Model to Preserve Privacy?*, 2022, ACM Digital Library, FAccT '22, 20 juni 2022, Seoul, Republiken Korea, tillgänglig på <https://dl.acm.org/doi/abs/10.1145/3531146.3534642>; iii) Vassilev A., Oprea A., Fordyce A., Anderson H., *Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations*, januari 2024, National Institute of Standards and Technology, tillgänglig på <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.pdf>; iv) Carlini N., Tramèr F., Wallace E., Jagielski M., Herbert-Voss A., Lee K., Roberts A., Brown T., Song D., Erlingsson U., Oprea A., Raffel C., *Extracting Training Data from Large Language Models*, arXiv:2012.07805v2 [cs.CR] 15 juni 2021, tillgänglig på <https://arxiv.org/pdf/2012.07805>; v) Fredrikson M., Jha S., Ristenpart T., *Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures*, ACM Digital Library, 12 oktober 2015, tillgänglig på <https://dl.acm.org/doi/abs/10.1145/2810103.2813677>; vi) Zhang Y., Jia R., Pei H., Wang W., Li B., Song D., *The Secret Revealer: Generative Model-Inversion Attacks Against Deep Neural Networks*, arXiv:1911.07135v2 [cs.LG], 18 april 2020, tillgänglig på <https://arxiv.org/pdf/1911.07135>.

²⁵ I ett AI-system som bygger på generativ AI används på engelska ordet ”regurgitation” (regurgiterande), vilket avser ett återflöde/återskapande där utdata relaterar direkt till träningsdata.

35. När det gäller fråga 1 i begäran²⁶ uppmanas EDPB att klargöra under vilka omständigheter en AI-modell som har tränats med hjälp av personuppgifter kan anses vara anonym. När det gäller fråga 1 i (a) i begäran²⁷ ombeds EDPB att klargöra vilka bevis och/eller vilken dokumentation som tillsynsmyndigheterna ska beakta vid bedömningen av om en AI-modell är anonym.

3.2.1 Allmänna överväganden om anonymisering i det aktuella sammanhanget

36. Användningen av uttrycket "varje upplysning" i definitionen av "personuppgifter" i artikel 4.1 i dataskyddsförordningen speglar avsikten att ge detta begrepp ett brett tillämpningsområde som omfattar alla typer av upplysningar förutsatt att de "avser" den registrerade, som identifieras eller är identifierbar direkt eller indirekt.
37. Information kan avse en fysisk person även om den är tekniskt organiserad eller kodad (till exempel i ett format som endast kan läsas av maskiner, oavsett om det är skyddat eller öppet) på ett sätt som inte gör förhållandet till den fysiska personen omedelbart uppenbart. I sådana fall kan programvaruapplikationer användas för att enkelt identifiera, känna igen och utvinna specifika data. Detta gäller särskilt för AI-modeller där parametrarna representerar statistiska samband mellan träningsuppgifterna och där det kan vara möjligt att utvinna korrekta eller felaktiga (på grund av inferens) personuppgifter, antingen direkt från sambanden mellan de uppgifter som ingår i modellen, eller genom att fråga modellen.
38. Eftersom AI-modeller vanligtvis inte innehåller data som kan isoleras eller länkas direkt, utan snarare parametrar som representerar probabilistiska relationer mellan de data som ingår i modellen, kan det vara möjligt att härleda²⁸ information från modellen, till exempel genom medlemskapsinferens, i realistiska scenarier. För att en tillsynsmyndighet ska komma överens med den

²⁶ "Under vilka omständigheter kan detta inträffa?"

²⁷ "Om så är fallet, hur kan man visa vilka åtgärder som har vidtagits för att säkerställa att AI-modellen inte behandlar personuppgifter?"

²⁸ (i) Carlini N., Chien S., Nasr M., Song S., Terzis A., Tramer F., *Membership Inference Attacks From First Principles*, arXiv:2112.03570, tillgänglig på <https://arxiv.org/abs/2112.03570>;

(ii) Crețu A.M., Guépin F., and De Montjoye Y.A., *Correlation inference attacks against machine learning models*. Sci. Adv.10, ead9260(2024). DOI:10.1126/sciadv.ad9260 tillgänglig på <https://www.science.org/doi/10.1126/sciadv.ad9260>;

(iii) Dana L., Pydi M. S., Chevalere Y., *Memorization in Attention-only Transformers* arXiv:2411.10115v1 [cs.AI], 15 november 2024, tillgänglig på: <https://arxiv.org/abs/2411.10115>;

(iv) Gehrke M., Liebenow J., Mohammadi E. & Braun T. m.fl. *Lifting in Support of Privacy-Preserving Probabilistic Inference*. Künstl Intell, 13 juni 2024, tillgänglig på: <https://doi.org/10.1007/s13218-024-00851-y>;

(v) Hu H., *Membership Inference Attacks and Defenses on Machine Learning Models Literature*, tillgänglig på: <https://github.com/HongshengHu/membership-inference-machine-learning-literature>;

(vi) Nasr M., Carlini N., Hayase J., Jagielski M., Cooper A. F., Ippolito D., Choquette-Choo C. A., Wallace E., Tramèr F., and Lee K., *Scalable Extraction of Training Data from (Production) Language Models*, arXiv:2311.17035, 28 november 2023, tillgänglig på: <https://arxiv.org/abs/2311.17035>;

(vii) Shokri R., Stronati M., Song C., Shmatikov V., *Membership Inference Attacks against Machine Learning Models* arXiv:1610.05820v2 [cs.CR], 31 mars 2017, tillgänglig på <https://arxiv.org/abs/1610.05820>;

(viii) Staab R., Vero M., Mislav Balunović, Martin Vechev, 2024, *Beyond Memorization: Violating Privacy Via Inference with Large Language Models*, arXiv:2310.07298v2, 6 maj 2024, tillgänglig på <https://arxiv.org/abs/2310.07298>;

(ix) Wu F., Cui L., Yao S., Yu S., *Inference Attacks in Machine Learning as a Service: A Taxonomy, Review, and Promising Directions* arXiv:2406.02027v1 [cs.LG], 27 juni 2024, tillgänglig på <https://arxiv.org/abs/2406.02027v1>;

(x) Zhang J., Das D., Kamath G., Tramèr F., *Membership Inference Attacks Cannot Prove that a Model Was Trained On Your Data* arXiv:2409.19798v1, [cs.LG], 29 september 2024, tillgänglig på <https://arxiv.org/abs/2409.19798>;

(xi) Zhou Z., Xiang J., Chen C., and Su S., *Quantifying and Analyzing Entity-Level Memorization in Large Language Models*, arXiv:2308.15727v2 [cs.CL], 5 november 2023, tillgänglig på: <https://arxiv.org/abs/2308.15727>.

personuppgiftsansvarige om att en viss AI-modell kan betraktas som anonym bör den därför åtminstone kontrollera om den har mottagit tillräckliga bevis som innebär att i) personuppgifter relaterade till träningsuppgifterna inte med rimliga hjälpmedel kan extraheras²⁹ ur modellen, och ii) inga utdata som produceras när modellen används avser de registrerade vars personuppgifter använts för att träna modellen.

39. Tillsynsmyndigheterna bör beakta tre faktorer vid bedömningen av huruvida dessa villkor är uppfyllda.
40. För det första bör tillsynsmyndigheterna beakta de faktorer som anges i de senaste yttrandena från artikel 29-arbetsgruppen och/eller EDPB:s riktlinjer på temat. När det gäller anonymisering vid tidpunkten för detta yttrande bör tillsynsmyndigheterna beakta de element som ingår i artikel 29-gruppens yttrande 05/2014 om anonymiseringstekniker (nedan kallat *artikel 29-gruppens yttrande 05/2014*), där det anges att om det inte är möjligt gallra ut, länka eller härleda information från den förmodat anonyma uppsättningen data, kan uppgifterna betraktas som anonyma³⁰. Där anges också att *"när ett förslag inte uppfyller något av kriterierna bör en noggrann bedömning av identifieringsriskerna göras"*³¹. **Med tanke på den ovannämnda sannolikheten för utvinning och inferens anser EDPB att det är mycket sannolikt att AI-modeller kräver en sådan grundlig utvärdering av riskerna för identifiering.**
41. För det andra bör denna bedömning göras med beaktande av alla hjälpmedel som *"rimligen kan komma att användas"* av den personuppgiftsansvarige eller av någon annan för att identifiera enskilda personer³², och fastställandet av dessa hjälpmedel bör baseras på objektiva faktorer, såsom förklaras i skäl 26 i dataskyddsförordningen, vilka kan omfatta följande:
 - a. Egenskaperna hos själva träningsuppgifterna, AI-modellen och träningsprocessen³³.
 - b. Det sammanhang i vilket AI-modellen släpps och/eller bearbetas³⁴.
 - c. Den ytterligare information som skulle möjliggöra identifiering och som kan tänkas vara tillgänglig för en given person.
 - d. De kostnader och den tid som personen skulle behöva för att få sådan ytterligare information (om den inte redan är tillgänglig för dem)³⁵.
 - e. Den teknik som fanns tillgänglig vid tidpunkten för behandlingen samt den tekniska utvecklingen³⁶.
42. För det tredje bör tillsynsmyndigheterna överväga huruvida personuppgiftsansvariga har bedömt risken för identifiering av den personuppgiftsansvarige och av olika typer av *"andra personer"*,

²⁹ Utvinning eller extrahering omfattar särskilt de fall där personuppgifter härleds från själva AI-modellen, med liten eller ingen användning av gränssnittet för förfrågningar.

³⁰ Artikel 29-gruppens yttrande 05/2014, s. 24.

³¹ Artikel 29-gruppens yttrande 05/2014, s. 24.

³² EU-domstolens dom av den 19 oktober 2016 i mål C-582/14, *Breyer mot Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), punkt 43.

³³ Detta inkluderar faktorer som unikhet hos datapunkterna i träningsuppgifterna, graden av noggrannhet i informationen, aggregering, randomisering och i synnerhet hur allt detta påverkar sårbarheten i fråga om identifieringsrisk.

³⁴ Detta inkluderar kontextuella element, såsom att begränsa åtkomsten till vissa personer och rättsliga skyddsåtgärder.

³⁵ EU-domstolens dom av den 7 mars 2024, mål C-479/22 P, *OC mot Europeiska kommissionen* (ECLI:EU:C:2024:215), punkt 50.

³⁶ EU-domstolens dom av den 7 mars 2024, mål C-479/22 P, *OC mot Europeiska kommissionen* (ECLI:EU:C:2024:215), punkt 50.

inklusive oavsiktliga tredje parter som får tillgång till AI-modellen, även med hänsyn till huruvida de rimligen kan anses kunna få tillgång till eller behandla uppgifterna i fråga.

43. **Sammanfattningsvis anser EDPB att för att en AI-modell ska anses vara anonym bör båda följande risker vid användning med rimliga hjälpmedel vara obetydliga³⁷: i) sannolikheten för direkt (inklusive probabilistisk) utvinning av personuppgifter om personer vars personuppgifter använts för att träna modellen, och ii) sannolikheten för att, avsiktligt eller inte, erhålla sådana personuppgifter vid förfrågningar, för alla registrerade. Som standard bör tillsynsmyndigheterna ta hänsyn till att AI-modeller sannolikt kommer att kräva en grundlig utvärdering av risken för identifiering för att fastställa deras eventuella anonyma karaktär. Denna sannolikhet bör bedömas med hänsyn till alla hjälpmedel "*som rimligen kan komma att användas*" av den personuppgiftsansvarige eller någon annan person, och bör även beakta oavsiktligt (vidare)utnyttjande eller utlämnande av modellen.**

3.2.2 Element för att utvärdera den kvarvarande sannolikheten för identifiering

44. Även om åtgärder kanske vidtas både under utvecklings- och driftsättningsstadierna för att minska risken för att AI-modellen ska röja personuppgifter bör utvärderingen av anonymitet för AI-modellen också överväga direkt åtkomst till modellen.
45. Dessutom bör tillsynsmyndigheterna i varje enskilt fall bedöma om de åtgärder som den personuppgiftsansvarige har vidtagit för att säkerställa och bevisa att en AI-modell är anonym är lämpliga och effektiva.
46. I synnerhet kan resultatet av en tillsynsmyndighets bedömning skilja sig åt mellan en offentligt tillgänglig AI-modell, som är tillgänglig för ett okänt antal personer med ett okänt antal metoder för att utvinna personuppgifter, och en intern AI-modell som endast är tillgänglig för anställda inom ett visst företag. I båda fallen bör tillsynsmyndigheterna kontrollera att de personuppgiftsansvariga har uppfyllt sin ansvarsskyldighet enligt artikel 5.2 och artikel 24 i dataskyddsförordningen. De hjälpmedel "*som rimligen kan komma att användas*" av andra personer kan dock påverka omfattningen och arten av de möjliga scenarier som bör beaktas. Därför kan tillsynsmyndigheterna överväga olika nivåer av testning och motståndskraft mot attacker, beroende på sammanhanget för utvecklingen och driftsättningen av modellen.
47. I detta avseende tillhandahåller EDPB nedan en icke-bindande och icke uttömmande förteckning över möjliga element som tillsynsmyndigheterna kan beakta vid bedömningen av en personuppgiftsansvarigs påstående om anonymitet. Andra tillvägagångssätt kan vara möjliga om de erbjuder en likvärdig skyddsnivå, särskilt med hänsyn till hur tekniken utvecklas.
48. Förekomsten eller frånvaron av de element som anges nedan är inte avgörande kriterier för att fastställa anonymiteten hos en AI-modell.

3.2.2.1 AI-modellens utformning

49. När det gäller hur AI-modeller utformats bör tillsynsmyndigheterna utvärdera de metoder som de personuppgiftsansvariga har använt under utvecklingsfasen. Tillämpningen av och effektiviteten hos fyra nyckelområden (som anges nedan) bör övervägas i detta avseende.

³⁷ EU-domstolens dom av den 19 oktober 2016 i mål C-582/14, *Breyer mot Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), punkt 46, och EU-domstolens dom av den 7 mars 2024 i mål C-479/22 P, *OC mot Europeiska kommissionen* (ECLI:EU:C:2024:215), punkt 51.

Urval av källor

50. Det första steget innebär att granska valet av källor som använts för att träna AI-modellen. Detta omfattar en utvärdering, utförd av tillsynsmyndigheterna, av alla åtgärder som vidtagits för att undvika eller begränsa insamlingen av personuppgifter, däribland i) urvalskriteriernas lämplighet, ii) de valda källornas relevans och lämplighet med hänsyn till det eller de avsedda ändamålen, och iii) huruvida olämpliga källor har uteslutits.

Förbearbetning och minimering av uppgifter

51. Det andra utvärderingsområdet gäller förberedelserna av data för träningsfasen. Tillsynsmyndigheterna bör särskilt undersöka i) huruvida användningen av anonyma uppgifter och/eller personuppgifter som har genomgått pseudonymisering har övervägts, och ii) ifall det beslutats att inte använda sådana åtgärder, skälen till detta, med beaktande av det avsedda ändamålet, iii) de strategier och tekniker för uppgiftsminimering som använts för att begränsa den mängd personuppgifter som ingår i träningsprocessen och iv) eventuella datafiltreringsprocesser som genomförts före modellträningen i syfte att rensa bort irrelevanta personuppgifter.

Metodologiska val avseende träningen

52. Det tredje utvärderingsområdet gäller huruvida robusta metoder använts i utvecklingen av AI-modellen. Tillsynsmyndigheterna bör bedöma om sådana metoder som kan minska eller eliminera identifierbarheten avsevärt har använts, bland annat i) om det ingår regleringsmetoder för att förbättra modellgeneraliseringen och minska överanpassning, och, vilket är avgörande, ii) om den personuppgiftsansvarige har använt lämpliga och effektiva integritetsskyddande tekniker (t.ex. differentiell integritet).

Åtgärder avseende modellens resultat

53. Det sista utvärderingsområdet gäller eventuella metoder eller åtgärder som integrerats i AI-modellen och som kanske inte påverkar risken för direkt utvinning av personuppgifter från modellen av någon som har direkt åtkomst till den, men som kan minska risken för att personuppgifter relaterade till träningsdata erhålls via förfrågningar.

3.2.2.2 AI-modellanalys

54. För att tillsynsmyndigheterna ska kunna bedöma den utformade AI-modellens robusthet när det gäller anonymisering är ett första steg att säkerställa att designen har utvecklats som planerat och är föremål för effektiv teknisk styrning. Tillsynsmyndigheterna bör utvärdera om de personuppgiftsansvariga har genomfört några dokumentbaserade revisioner (interna eller externa) som omfattar en utvärdering av de valda åtgärderna och av deras inverkan för att begränsa risken för identifiering. Detta kan omfatta analys av rapporter från kodgranskningar samt en teoretisk analys som dokumenterar lämpligheten av de åtgärder som valts för att minska risken för återidentifiering i den berörda modellen.

3.2.2.3 Testning av AI-modeller och deras motståndskraft mot attacker

55. Slutligen bör tillsynsmyndigheterna ta hänsyn till omfattningen, frekvensen, antalet och kvaliteten på de tester som den personuppgiftsansvarige har utfört på modellen. Tillsynsmyndigheterna bör i synnerhet ta hänsyn till att endast framgångsrika tester som omfattar allmänt kända, toppmoderna attackmetoder kan utgöra bevis på motståndskraft mot dessa attacker. Vid tidpunkten för detta yttrande kan detta bland annat innefatta strukturerade tester mot i) attribut- och medlemsinferens, ii) exfiltrering, iii) återflöde/återskapande av träningsdata, iv) modellinvertering eller v) rekonstruktionsattacker.

3.2.2.4 Dokumentation

56. Artiklarna 5, 24, 25 och 30 i dataskyddsförordningen och, i fall där de registrerades rättigheter och friheter sannolikt utsätts för hög risk, artikel 35 i dataskyddsförordningen, kräver att

personuppgiftsansvariga på ett tillfredsställande sätt dokumenterar sin behandling av personuppgifter. Detta gäller också för all behandling som omfattar träning av en AI-modell, även om syftet med behandlingen är anonymisering. Tillsynsmyndigheterna bör beakta sådan dokumentation och eventuella regelbundna bedömningar av riskerna av den behandling som utförs av personuppgiftsansvariga, eftersom dessa steg är grundläggande för att visa att personuppgifter inte behandlas.

57. **EDPB anser att tillsynsmyndigheterna alltid bör ta hänsyn till dokumentationen när ett påstående om anonymitet för en viss AI-modell ska utvärderas. EDPB noterar att om en tillsynsmyndighet, efter att ha bedömt påståendet om anonymitet med beaktande av dokumentationen, inte kan bekräfta att effektiva åtgärder har vidtagits för att anonymisera AI-modellen, kan denna tillsynsmyndighet anse att den personuppgiftsansvarige inte har uppfyllt sina ansvarsskyldigheter enligt artikel 5.2 i dataskyddsförordningen. Därför bör även efterlevnaden av andra bestämmelser i dataskyddsförordningen beaktas.**
58. Helst bör tillsynsmyndigheterna kontrollera om den personuppgiftsansvariges dokumentation omfattar följande:
- All information som rör konsekvensbedömning avseende dataskydd, inklusive bedömningar och beslut om att en konsekvensbedömning inte varit nödvändig.
 - Rådgivning eller återkoppling som tillhandahållits av dataskyddsombudet (om ett dataskyddsombud har eller borde ha utnämnts).
 - Information om de tekniska och organisatoriska åtgärder som vidtagits vid utformningen av AI-modellen för att minska identifieringsrisken, inbegripet den hotmodell och de riskbedömningar som dessa åtgärder grundar sig på. Detta bör inbegripa de specifika åtgärderna för varje källa till träningsdatasetet, inklusive relevanta webbadresser och beskrivning av åtgärder som vidtagits (eller som vidtogs redan av tredjepartsleverantören av datasetet).
 - De tekniska och organisatoriska åtgärder som vidtagits i alla skeden av modellens livscykel och som antingen bidragit till eller bekräftat frånvaron av personuppgifter i modellen.
 - Dokumentation som visar AI-modellens teoretiska motståndskraft mot tekniker för återidentifiering, samt de kontroller som utformats för att begränsa eller bedöma huvudattacker framgång och verkan (återflöde/återskapande, angrepp genom medlemskapsinferens, exfiltrering osv.). I detta kan särskilt följande ingå: i) Förhållandet mellan mängden träningsdata och antalet parametrar i modellen, inbegripet analysen av dess inverkan på modellen³⁸, ii) mätvärden avseende sannolikheten för återidentifiering baserat på den aktuella tekniska utvecklingen, iii) rapporter om hur modellen har testats (av vem, när, hur och i vilken utsträckning) och iv) resultaten av testerna.

³⁸ Ricciato F., *A Cautionary Reflection on (Pseudo-)Synthetic Data from Deep Learning on Personal Data*, Privacy in Statistical Databases conference (PSD 2024), Antibes, Frankrike, september 2024, bilder finns på https://cros.ec.europa.eu/system/files/2024-10/20240926_PSD2024_Ricciato_v6_1.pdf och Belkin M., Hsu D., Ma S. och Mandal S. (2019), *Reconciling modern machine-learning practice and the classic bias-variance trade-off*. Proceedings of the National Academy of Sciences, 24 juli 2019, 116(32) 15849-15854, tillgänglig på: <https://www.pnas.org/doi/10.1073/pnas.1903070116>

- f. Den dokumentation som lämnats till den eller de personuppgiftsansvariga som använder modellen och/eller till de registrerade, särskilt den dokumentation som rör de åtgärder som vidtagits för att minska risken för identifiering och som rör eventuella kvarstående risker.

3.3 Om huruvida ett berättigat intresse är en lämplig rättslig grund för behandling av personuppgifter i samband med utveckling och driftsättning av AI-modeller

59. För att besvara frågorna 2 och 3 i begäran kommer EDPB först att lämna allmänna synpunkter på vissa viktiga aspekter som tillsynsmyndigheterna bör beakta, oavsett den rättsliga grunden för behandlingen, när de bedömer hur personuppgiftsansvariga kan påvisa efterlevnad av dataskyddsförordningen i samband med AI-modeller. Med utgångspunkt i riktlinjerna 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i dataskyddsförordningen³⁹ kommer EDPB sedan att överväga de tre steg som krävs för bedömningen av berättigat intresse i samband med utveckling och driftsättning av AI-modeller.

3.3.1 Allmänna synpunkter

60. EDPB erinrar om att dataskyddsförordningen inte fastställer någon hierarki mellan de olika rättsliga grunder som anges i artikel 6.1 i dataskyddsförordningen⁴⁰.
61. I artikel 5 i dataskyddsförordningen fastställs principerna för behandling av personuppgifter. EDPB lyfter fram de som är av störst betydelse för detta yttrande och som åtminstone bör beaktas av tillsynsmyndigheterna vid bedömningen av specifika AI-modeller, samt de mest relevanta kraven från andra bestämmelser i dataskyddsförordningen, med beaktande av tillämpningsområdet för detta yttrande.
62. **Principen om ansvarsskyldighet** (artikel 5.2 i dataskyddsförordningen) – I denna princip föreskrivs att den personuppgiftsansvarige ska ansvara för och kunna visa att kraven i dataskyddsförordningen är uppfyllda. I detta avseende bör rollerna och ansvarsområdena för de parter som behandlar personuppgifter i samband med utvecklingen eller driftsättningen av en AI-modell bedömas innan behandlingen äger rum, så att personuppgiftsansvarigas, gemensamt personuppgiftsansvarigas och eventuella personuppgiftsbiträdens skyldigheter fastställs från början.
63. **Principerna om laglighet, korrekthet och öppenhet** (artikel 5.1 a i dataskyddsförordningen) – Vid bedömningen av lagligheten i behandlingen av uppgifter inom ramen för AI-modeller, mot bakgrund av artikel 6.1 i dataskyddsförordningen, anser EDPB att det är lämpligt att skilja mellan de olika stegen i behandlingen av personuppgifter⁴¹. Principen om korrekthet, som är nära kopplad till principen om insyn, kräver att personuppgifter inte behandlas med otilbörliga metoder eller genom bedrägeri, eller på ett sätt som är *”omotiverat missgynnande, olagligt diskriminerande, oväntat eller vilseledande för den registrerade”*⁴². Med tanke på komplexiteten i den teknik som används bör information om behandling av personuppgifter inom AI-modeller därför tillhandahållas i en lätt tillgänglig, begriplig

³⁹ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024.

⁴⁰ Ibidem, punkt 1.

⁴¹ EDPB:s rapport om det arbete som utförts av arbetsgruppen om ChatGPT, antagen den 23 maj 2024, punkt 14.

⁴² EDPB:s rapport om det arbete som utförts av arbetsgruppen om ChatGPT, antagen den 23 maj 2024, punkt 23; EDPB:s riktlinjer 4/2019 om inbyggt dataskydd och dataskydd som standard, version 2.0, antagna den 20 oktober 2020, punkt 69; artikel 29-gruppens riktlinjer om öppenhet enligt förordning 2016/679, reviderade och antagna den 11 april 2018 och godkända av EDPB den 25 maj 2018, punkt 2.

och användarvänlig form⁴³. Insyn i fråga om behandlingen av personuppgifter omfattar i synnerhet efterlevnad av informationsskyldigheterna enligt artiklarna 12–14 i dataskyddsförordningen⁴⁴, som vid automatiserat beslutsfattande, inbegripet profilering, också kräver meningsfull information om den bakomliggande logiken samt betydelsen och de förutsedda följderna av behandlingen för den registrerade⁴⁵. Med tanke på att AI-modellernas utvecklingsfaser kan inbegripa insamling av stora mängder data från offentligt tillgängliga källor (till exempel genom webbinsamlingsteknik) är tillämpningen av det undantag som föreskrivs i artikel 14.5 b i dataskyddsförordningen strikt begränsad till de fall då kraven i denna bestämmelse är helt uppfyllda⁴⁶.

64. **Principer för ändamålsbegränsning och uppgiftsminimering** (artikel 5.1 b och c i dataskyddsförordningen) – I enlighet med principen om uppgiftsminimering krävs det vid utvecklingen och driftsättningen av AI-modeller att personuppgifter är adekvata, relevanta och inte för omfattande i förhållande till de ändamål för vilka de behandlas. Detta kan omfatta behandling av personuppgifter för att undvika risken för potentiella snedvridningar och fel när detta tydligt och specifikt identifieras inom syftet och personuppgifterna är nödvändiga för detta ändamål (till exempel om ändamålet inte kan uppnås effektivt genom behandling av andra uppgifter, däribland syntetiska eller anonymiserade uppgifter)⁴⁷. Artikel 29-gruppen har redan betonat att *”syftet med insamlingen måste vara tydligt och specifikt angivet [...]”*⁴⁸. Vid bedömningen av om det eftersträvade syftet är berättigat, specifikt och uttryckligt och om behandlingen är förenlig med principen om uppgiftsminimering, bör man först identifiera den aktuella behandlingen. I synnerhet kan de olika stadierna inom utvecklings- eller driftsättningsfaserna inbegripa samma eller olika behandlingsaktiviteter och kan medföra flera på rad följande personuppgiftsansvariga eller gemensamt personuppgiftsansvariga. I vissa fall kan syftet med driftsättningen av AI-modellen fastställas i ett tidigt utvecklingsskede. Även om så inte är fallet bör det redan finnas ett tydligt sammanhang för denna driftsättning, och därför bör det övervägas hur detta sammanhang påverkar syftet med utvecklingen. När tillsynsmyndigheterna granskar syftet med behandlingen i ett visst utvecklingsskede bör de förvänta sig en viss grad av detaljrikedom från den eller de personuppgiftsansvariga och en förklaring av hur dessa detaljer förhåller sig till syftet med behandlingen. Detta kan till exempel omfatta information om den typ av AI-modell som utvecklats, dess förväntade funktioner och eventuella andra relevanta sammanhang som redan är kända i det skedet. Sammanhanget för driftsättningen kan till exempel också gälla om en modell utvecklas för internt bruk, om den personuppgiftsansvarige avser att sälja eller distribuera modellen till tredje parter efter att den har utvecklats, inbegripet om modellen i första hand är avsedd att användas för forskningsändamål eller kommersiella ändamål.
65. **Den registrerades rättigheter** (kapitel III i dataskyddsförordningen) – Trots att tillsynsmyndigheterna måste säkerställa att alla registrerades rättigheter respekteras när AI-modeller utvecklas och driftsätts av personuppgiftsansvariga, påminner EDPB om att rätten att göra invändningar enligt artikel 21 i

⁴³ Artikel 29-gruppens riktlinjer om öppenhet enligt förordning (EU) 2016/679, reviderade och antagna den 11 april 2018 och godkända av EDPB den 25 maj 2018, punkt 5.

⁴⁴ Se även skäl 39 i dataskyddsförordningen, där det anges att det *”bör vara klart och tydligt för fysiska personer hur personuppgifter som rör dem insamlas, används, konsulteras eller på annat sätt behandlas samt i vilken utsträckning personuppgifterna behandlas eller kommer att behandlas”*.

⁴⁵ Artikel 13.2 f och artikel 14.2 g i dataskyddsförordningen.

⁴⁶ EDPB:s rapport om det arbete som utförts av arbetsgruppen om ChatGPT, antagen den 23 maj 2024, punkt 27.

⁴⁷ I artikel 10.5 i AI-förordningen fastställs dessutom särskilda regler för behandling av specifika kategorier av personuppgifter i samband med AI-system med hög risk, i syfte att säkerställa upptäckt och korrigerande av bias.

⁴⁸ Artikel 29-gruppens yttrande 03/2013 om ändamålsbegränsning (WP203), s. 15–16.

dataskyddsförordningen gäller och bör säkerställas i samband med att berättigat intresse åberopas som rättslig grund av en personuppgiftsansvarig⁴⁹.

3.3.2 Överväganden avseende de tre stegen i bedömningen av berättigat intresse i samband med utvecklingen och driftsättningen av AI-modeller

66. För att avgöra om en viss behandling av personuppgifter kan grundas på artikel 6.1 f i dataskyddsförordningen bör tillsynsmyndigheterna kontrollera att de personuppgiftsansvariga noggrant har bedömt och dokumenterat huruvida följande tre kumulativa villkor är uppfyllda: i) den personuppgiftsansvariges eller en tredje part har ett berättigat intresse, ii) behandlingen är nödvändig för att tillgodose det berättigade intresset, och iii) det berättigade intresset åsidosätts inte av de registrerades intressen eller grundläggande rättigheter och friheter⁵⁰.

3.3.2.1 Första steget – En personuppgiftsansvarigs eller en tredje parts berättigade intresse

67. Ett intresse är den bredare drivkraft eller fördel som en personuppgiftsansvarig eller tredje part kan ha av att utföra en viss behandling⁵¹. Samtidigt som dataskyddsförordningen och EU-domstolen erkänt flera intressen som legitima⁵² bör bedömningen av legitimiteten för ett visst intresse vara resultatet av en individuell analys av varje enskilt fall.
68. Såsom EDPB erinrar om i sina riktlinjer om berättigat intresse⁵³ kan ett intresse anses vara berättigat om följande tre kumulativa kriterier är uppfyllda:
- a. Intresset är lagligt⁵⁴.

⁴⁹ Enligt artikel 21 i dataskyddsförordningen gäller att om en registrerad, av skäl som hänför sig till dennes särskilda situation, invänder mot behandlingen av personuppgifter som rör honom eller henne, ska den personuppgiftsansvarige inte längre behandla personuppgifterna, såvida inte den personuppgiftsansvarige kan uppvisa tvingande berättigade skäl för behandlingen som väger tyngre än den registrerades intressen, rättigheter och friheter eller för fastställande, utövande eller försvar av rättsliga anspråk. De två aspekter som tillsynsmyndigheterna bör ta hänsyn till är därför huruvida den personuppgiftsansvarige kan påvisa sådana tvingande, överordnande berättigade skäl och huruvida rätten att göra invändningar kan utövas.

⁵⁰ EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 106; EU-domstolens dom av den 11 december 2019 i mål C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), punkt 40. Se även EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 12 och följande punkter. Som påpekas i dessa riktlinjer bör denna ”bedömning göras i början av behandlingen, med medverkan av dataskyddsombudet (om ett sådant utsetts), och bör dokumenteras av den personuppgiftsansvarige i enlighet med den princip om ansvarsskyldighet som anges i artikel 5.2 i dataskyddsförordningen”.

⁵¹ EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 14.

⁵² EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 16.

⁵³ EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 17.

⁵⁴ EU-domstolens dom av den 4 oktober 2024 i mål C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), punkt 49, där EU-domstolen betonade att ett berättigat intresse inte får strida mot lagen. I detta avseende betonar EDPB att rättsliga ramar i tillämpliga fall bör beaktas vid bedömningen av huruvida ett visst intresse är lagligt. Se till exempel artiklarna 26.3 och 28 i Europaparlamentets och rådets förordning (EU) 2022/2065 av den 19 oktober 2022 om en inre marknad för digitala tjänster och om ändring av direktiv 2000/31/EG (förordningen om digitala tjänster) om förbjuden riktad reklam för minderåriga, artikel 5.1 och 5.2 i AI-förordningen om förbjuden AI-praxis (manipulativ praxis och under gränsen för medvetande) samt behandling som strider mot immateriella rättigheter och bestämmelserna i direktiv (EU) 2019/790 om upphovsrätt och närstående rättigheter på den digitala inre marknaden.

- b. Intresset är klart och precist formulerat.
- c. Intresset är verkligt och aktuellt, inte spekulativt.

69. Med förbehåll för de två andra steg som krävs enligt bedömningen av berättigat intresse kan följande exempel utgöra berättigade intressen i samband med AI-modeller: i) utveckling av en tjänst som utförs av en samtalsagent för att hjälpa användarna, ii) utveckling av ett AI-system för att upptäcka bedrägligt innehåll eller bedrägligt beteende, och iii) förbättrade metoder för upptäckt av hot i ett informationssystem.

3.3.2.2 Andra steget – Analys av huruvida behandlingen är nödvändig för att tillgodose det berättigade intresset

70. Det andra steget i bedömningen består i att avgöra om behandlingen av personuppgifter är nödvändig för att tillgodose det eller de berättigade intressen som åberopas⁵⁵ ("nödvändighetsprövning").
71. I skäl 39 i dataskyddsförordningen anges att personuppgifter endast bör behandlas "*om syftet med behandlingen inte rimligen kan uppnås genom andra medel*". Enligt EU-domstolen och tidigare vägledning från EDPB bör villkoret avseende nödvändigheten av behandlingen granskas mot bakgrund av de registrerades grundläggande rättigheter och friheter och i förhållande till principen om uppgiftsminimering som fastställs i artikel 5.1 c i dataskyddsförordningen⁵⁶.
72. Den metod som EU-domstolen hänvisar till tar hänsyn till sammanhanget för behandlingen samt konsekvenserna för den personuppgiftsansvarige och de registrerade. Bedömningen av nödvändigheten omfattar därför två delar: i) huruvida behandlingsaktiviteten bidrar till att uppnå syftet⁵⁷ och ii) huruvida det kan finnas något mindre inkräktande sätt att uppnå detta syfte⁵⁸.

⁵⁵ EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkterna 28–30.

⁵⁶ EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkterna 108 och 109, även med hänvisning till EU-domstolens dom av den 11 december 2019 i mål C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), punkt 48; EU-domstolens dom av den 9 november 2010 i de förenade målen C-92/09 och C-93/09, *Volker und Markus Schecke* (ECLI:EU:C:2010:662), punkterna 85 och 86; EU-domstolens dom av den 22 juni 2021 i mål C-439/19, *Latvijas Republikas Saeima* (ECLI:EU:C:2021:504), punkterna 98, 109, 110 och 113. Se även till exempel: EDPB:s riktlinjer 3/2019 om behandling av personuppgifter genom videoenheter, version 2.0, antagna den 29 januari 2020, punkterna 24–26 och 73; EDPB:s riktlinjer 2/2019 om behandling av personuppgifter enligt artikel 6.1 b i den allmänna dataskyddsförordningen i samband med tillhandahållandet av onlinetjänster till registrerade, version 2.0, antagna den 8 oktober 2019, punkterna 23–25; EDPB:s yttrande 11/2024 om användning av ansiktsgenkänning för att rationalisera flödet av passagerare på flygplatser, version 1.1, antaget den 23 maj 2024, punkt 27.

⁵⁷ Se EU-domstolens dom av den 16 december 2008 i mål C-524/06, *Heinz Huber mot Bundesrepublik Deutschland* (ECLI:EU:C:2008:724), punkt 66. Se även i samma ärende förslag till avgörande av generaladvokat Poiares Maduro i mål C-524/06, *Heinz Huber mot Bundesrepublik Deutschland* (ECLI:EU:C:2008:194), punkt 16, där följande anges: "[...] det rätta kriteriet i detta mål är effektiviteten, och det är den nationella domstolen som ska tillämpa det. Den fråga som den ska ställa sig är om det finns andra sätt för behandling av uppgifter som invandrarmyndigheterna kan använda för att se till att bestämmelserna om uppehållsrättslig ställning efterlevs. Om den besvarar den frågan jakande ska den centraliserade datalagringen och behandlingen avseende unionsmedborgare anses vara rättsstridig. Det alternativa systemet är inte nödvändigtvis det mest effektiva eller det lämpligaste. Det räcker om det kan fungera på ett lämpligt sätt. Även om det centrala registret är mer effektivt eller passande eller användarvänligt än alternativen till det (exempelvis de decentraliserade lokala registren) är de sistnämnda klart att föredra om de kan användas för att visa unionsmedborgarnas uppehållsrättsliga ställning".

⁵⁸ Se EU-domstolens dom av den 27 september 2017 i mål C-73/16, *Peter Puškár* (ECLI:EU:C:2017:725), punkt 113: "Det ankommer således på den hänskjutande domstolen att pröva om åtgärden att upprätta den

73. Till exempel, och i tillämpliga fall, måste den avsedda volymen personuppgifter som ingår i AI-modellen bedömas mot bakgrund av mindre inkräktande alternativ som rimligen kan vara tillgängliga för att lika effektivt uppnå syftet med det berättigade intresse som eftersträvas. Om syftet också kan uppnås genom en AI-modell som inte innebär behandling av personuppgifter, bör behandling av personuppgifter inte anses vara nödvändig. Detta är särskilt relevant för utvecklingen av AI-modeller. Vid bedömningen av huruvida villkoret om nödvändighet är uppfyllt bör tillsynsmyndigheterna särskilt beakta den mängd personuppgifter som behandlas och huruvida den är proportionerlig med det berättigade intresse som eftersträvas, även mot bakgrund av principen om uppgiftsminimering.
74. Bedömningen av nödvändigheten bör också ta hänsyn till det bredare sammanhanget för den avsedda behandlingen av personuppgifter. Förekomsten av metoder som inkräktar mindre på de registrerades grundläggande rättigheter och friheter kan variera beroende på om den personuppgiftsansvarige har en direkt relation med de registrerade (förstapartsuppgifter) eller inte (tredjepartsuppgifter). EU-domstolen framförde vissa överväganden som man bör ta hänsyn till vid analysen av om det är nödvändigt att behandla uppgifter från första part för de berättigade intressen som eftersträvas (även i samband med utlämnande av sådana uppgifter till tredje part)⁵⁹.
75. Genomförande av tekniska skyddsåtgärder för att skydda personuppgifter kan också bidra till att uppfylla nödvändighetsprövningen. Detta skulle till exempel kunna omfatta genomförande av åtgärder såsom de som anges i avsnitt 3.2.2 på ett sätt som försvårar identifieringen av registrerade även om inte anonymisering uppnås. EDPB noterar att vissa av dessa åtgärder, även om de inte krävs för att följa dataskyddsförordningen, kan utgöra ytterligare skyddsåtgärder, vilket analyseras närmare i underavsnittet om riskreducerande åtgärder i avsnitt 3.3.2.3⁶⁰.

3.3.2.3 Tredje steget – Avvägningsprövningen

76. Det tredje steget i bedömningen av berättigat intresse är den så kallade **avvägningsprövningen**⁶¹. Detta steg består i att identifiera och beskriva de olika motstående rättigheter och intressen som riskerar att åsidosättas⁶², dvs. å ena sidan de registrerades intressen, grundläggande rättigheter och friheter och å andra sidan den personuppgiftsansvariges eller en tredje parts intressen. De konkreta omständigheterna i det enskilda fallet bör då beaktas för att visa att berättigat intresse är en lämplig rättslig grund för de aktuella behandlingsaktiviteterna⁶³.

*omtvistade förteckningen och att på denna uppföra namnet på de berörda personerna är ägnad att uppfylla de ändamål för vilka detta gjorts och om det inte är möjligt att uppnå dessa ändamål genom mindre ingripande åtgärder.” Se även till exempel förslag till avgörande av generaladvokat Rantos i mål C-252/21, *Meta mot Bundeskartellamt*, ECLI:EU:C:2022:704, punkt 61, där följande anges: “[...] Det krävs således att det finns en nära koppling mellan behandlingen och det intresse som eftersträvas och att det saknas alternativ som i större utsträckning iakttar skyddet av personuppgifter, eftersom det inte är tillräckligt att behandlingen helt enkelt är till nytta för personuppgiftsansvarig”.*

⁵⁹ EU-domstolens dom av den 4 oktober 2024 i mål C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), punkterna 51–53.

⁶⁰ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 57.

⁶¹ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkterna 31–60.

⁶² Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 32.

⁶³ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 32, med hänvisning till domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 110.

De registrerades intressen, grundläggande rättigheter och friheter

77. Av artikel 6.1 f i dataskyddsförordningen följer att den personuppgiftsansvarige vid bedömningen av de olika faktorerna i samband med avvägningstestet bör ta hänsyn till de registrerades intressen, grundläggande rättigheter och friheter. De registrerades intressen är i detta fall de som kan påverkas av behandlingen i fråga. I samband med utvecklingsfasen av en AI-modell kan dessa omfatta, utan att vara begränsade till, intresset i fråga om självbestämmande och behållande av kontrollen över de egna personuppgifterna (t.ex. de uppgifter som samlas in för utvecklingen av modellen). När det gäller driftsättningsfasen för en AI-modell kan de registrerades intressen omfatta, utan att vara begränsade till, intresse av att behålla kontrollen över de egna personuppgifterna (t.ex. de uppgifter som behandlas när modellen har införts), ekonomiska intressen (t.ex. om en AI-modell används av den registrerade för att generera intäkter eller används av en enskild person inom ramen för dennes yrkesverksamhet), intresse på grund av personliga fördelar (t.ex. om en AI-modell används för att förbättra tillgängligheten till vissa tjänster) eller socioekonomiska intressen (t.ex. om en AI-modell möjliggör tillgång till bättre hälso- och sjukvård eller underlättar utövandet av en grundläggande rättighet såsom tillgång till utbildning)⁶⁴.
78. Ju mer precist ett intresse definieras mot bakgrund av det avsedda syftet med behandlingen, desto bättre och tydligare kommer det att spegla verkligheten när det gäller de fördelar och risker som ska beaktas i avvägningstestet.
79. När det gäller de registrerades grundläggande rättigheter och friheter kan utvecklingen och driftsättningen av AI-modeller medföra allvarliga risker för rättigheter som skyddas av EU:s stadga om de grundläggande rättigheterna ("**EU-stadgan**"), bland annat rätten till respekt för privatlivet och familjelivet (artikel 7 i EU-stadgan) och rätten till skydd av personuppgifter (artikel 8 i EU-stadgan). Dessa risker kan uppstå under utvecklingsfasen, till exempel när personuppgifter erhålls genom webbinsamling mot de registrerades vilja eller utan deras vetskap. Dessa risker kan också uppstå under driftsättningsfasen, till exempel om personuppgifter behandlas av (eller som en del av) modellen på ett sätt som strider mot de registrerades rättigheter, eller om det är möjligt att av misstag eller genom angrepp (t.ex. medlemsinferens, extraktion eller modellinvertering) dra slutsatser om vilka personuppgifter som finns i träningsdatabasen. Sådana situationer utgör en risk för integriteten för registrerade vars uppgifter kan förekomma i AI-systemets driftsättningsfas (t.ex. ryktesrisk, identitetsstöld, bedrägeri eller annan säkerhetsrisk beroende på uppgifternas art).
80. Beroende på det specifika fallet kan även andra grundläggande rättigheter vara utsatta för risk. Exempelvis kan storskalig och godtycklig datainsamling av AI-modeller i utvecklingsfasen skapa en känsla av övervakning för de registrerade, särskilt med tanke på svårigheterna att förhindra att offentliga uppgifter erhålls genom webbinsamling. Detta kan leda till att personer börjar censurera sig själva, vilket riskerar att undergräva deras yttrandefrihet (artikel 11 i EU-stadgan). Driftsättningsfasen kan också medföra risker för yttrandefriheten om AI-modellen används för att blockera upplägg av innehåll från registrerade. Dessutom kan AI-modeller som rekommenderar olämpligt innehåll till sårbara personer utgöra en risk för deras psykiska hälsa (artikel 3.1 i EU-stadgan). I andra fall kan användningen av AI-modeller också ha negativa effekter på den enskildes rätt att arbeta (artikel 15 i EU-stadgan), till exempel när ett första urval av jobbansökningar görs med hjälp av en AI-modell. På samma sätt kan AI-modeller innebära risker för rätten till icke-diskriminering (artikel 21 i EU-stadgan) om de diskriminerar enskilda personer på grund av vissa personliga egenskaper (till exempel nationalitet eller kön). Dessutom kan användningen av AI-modeller innebära risker för individens

⁶⁴ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 38.

säkerhet (till exempel om AI-modellen används med ont uppsåt) och risker för individens fysiska och psykiska integritet⁶⁵.

81. Driftsättningen av AI-modeller kan också ha en positiv inverkan på vissa grundläggande rättigheter. Modellen kan till exempel stödja personens rätt till mental integritet (artikel 3 i stadgan), till exempel om en AI-modell används för att identifiera skadligt innehåll online, underlättar tillgången till vissa grundläggande tjänster eller underlättar utövandet av grundläggande rättigheter, såsom tillgång till information (artikel 11 i EU-stadgan) eller till utbildning (artikel 14 i EU-stadgan).

Behandlingens inverkan på registrerade

82. Den behandling av personuppgifter som sker under utvecklingen och driftsättningen av AI-modeller kan påverka de registrerade på olika sätt, såväl positivt som negativt⁶⁶. Om en behandlingsaktivitet till exempel medför fördelar för den registrerade kan dessa beaktas i avvägningstestet. Förekomsten av sådana fördelar kan leda till att en tillsynsmyndighet drar slutsatsen att den personuppgiftsansvariges eller en tredje parts intressen inte åsidosätts av de registrerades intressen, grundläggande rättigheter och friheter, men en sådan slutsats får endast vara resultatet av en analys av det specifika fallet med beaktande av alla relevanta faktorer.
83. Behandlingens inverkan på de registrerade kan påverkas av i) den typ av uppgifter som behandlas av modellerna, ii) sammanhanget för behandlingen och iii) de ytterligare konsekvenser som behandlingen kan ha⁶⁷.
84. När det gäller **typen av behandlade uppgifter** bör det erinras om att behandlingen av vissa kategorier av personuppgifter kan få betydande konsekvenser för de registrerade – med undantag för de särskilda kategorier av personuppgifter och andra uppgifter som rör fällande domar i brottmål och överträdelse, vilka åtnjuter förstärkt skydd enligt artiklarna 9 och 10 i dataskyddsförordningen. I detta sammanhang bör behandling av vissa typer av personuppgifter som rör mycket privat information (t.ex. finansiella uppgifter eller lokaliseringssuppgifter) i samband med utveckling och driftsättning av AI-modeller anses kunna ha en allvarlig inverkan på de registrerade. Under driftsättningsfasen kan konsekvenserna av sådan behandling för de registrerade till exempel vara ekonomiska (såsom diskriminering i anställningssammanhang) och/eller ryktesrelaterade (till exempel vid ärekränkning eller förtal).
85. När det gäller **behandlingens sammanhang**, är det först nödvändigt att identifiera de möjliga källorna till risk för de registrerade (till exempel det sätt på vilket modellen utvecklats, det sätt på vilket modellen kan användas och/eller huruvida de säkerhetsåtgärder som används för att skydda personuppgifterna är lämpliga). Modellens beskaffenhet och den avsedda operativa användningen spelar en viktig roll när det gäller att identifiera sådana potentiella källor till risk.
86. Det är också nödvändigt att bedöma hur allvarliga dessa risker är för de registrerade. Det kan bland annat övervägas hur personuppgifterna behandlas (t.ex. om de kombineras med andra dataset), omfattningen av behandlingen och mängden personuppgifter som behandlas⁶⁸ (till exempel den totala

⁶⁵ Riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 46.

⁶⁶ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 39.

⁶⁷ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 32.

⁶⁸ Se EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 43.

datavolymen, mängden data per registrerad, antalet registrerade som påverkas)⁶⁹, den registrerades status (till exempel om det rör sig om barn eller andra sårbara registrerade) och deras förhållande till den personuppgiftsansvarige (till exempel om den registrerade är en kund). Om det saknas tillräckliga skyddsåtgärder kan till exempel användningen av webbinsamling i utvecklingsfasen få betydande konsekvenser för enskilda personer på grund av den stora mängd uppgifter som samlas in, det stora antalet registrerade och den urskillningslösa insamlingen av personuppgifter.

87. De **ytterligare konsekvenser** som behandlingen kan få bör också beaktas vid bedömningen av behandlingens inverkan på de registrerade. De bör bedömas av tillsynsmyndigheterna i varje enskilt fall, med beaktande av de specifika fakta som föreligger.
88. Sådana konsekvenser kan innefatta (men är inte begränsade till) risker för kränkningar av de registrerades grundläggande rättigheter, såsom beskrivs i föregående underavsnitt⁷⁰. Riskerna kan variera i sannolikhets- och allvarlighetsgrad och kan uppstå till följd av behandling av personuppgifter som kan leda till fysisk, materiell eller immateriell skada, särskilt om behandlingen kan ge upphov till diskriminering⁷¹.
89. Om driftsättningen av en AI-modell innebär behandling av personuppgifter för både i) registrerade vars personuppgifter ingår i det dataset som används i utvecklingsfasen och ii) registrerade vars personuppgifter behandlas i driftsättningsfasen, bör tillsynsmyndigheterna skilja på och beakta de risker som påverkar intressen, rättigheter och friheter för var och en av dessa kategorier av registrerade när de bedömer ett avvägningstest som utförts av en personuppgiftsansvarig.
90. **Slutligen bör analysen av de möjliga ytterligare konsekvenserna av behandlingen också ta hänsyn till sannolikheten för att dessa ytterligare konsekvenser kommer att förverkligas.** Bedömningen av sannolikheten för detta bör göras med beaktande av de tekniska och organisatoriska åtgärder som vidtagits och de särskilda omständigheterna i det enskilda fallet. Tillsynsmyndigheterna kan till exempel överväga om åtgärder har vidtagits för att undvika ett potentiellt missbruk av AI-modellen. För AI-modeller som kan användas för en rad olika ändamål, till exempel generativ AI, kan detta omfatta kontroller som i möjligaste mån begränsar deras användning för skadliga syften, till exempel skapande av deepfakes, chatbottar som används för desinformation, nätfiske och andra typer av bedrägerier samt manipulativa AI/AI-agenter (särskilt när de är antropomorfa eller ger vilseledande information).

De registrerades rimliga förväntningar

91. I skäl 47 i dataskyddsförordningen anges att "[e]tt berättigat intresse kräver under alla omständigheter en noggrann bedömning, som inbegriper huruvida den registrerade vid tidpunkten för inhämtandet av personuppgifter och i samband med detta rimligen kan förvänta sig att en uppgiftsbehandling för detta ändamål kan komma att ske. Den registrerades intressen och grundläggande rättigheter skulle i synnerhet kunna väga tyngre än den personuppgiftsansvariges intressen, om personuppgifter

⁶⁹ EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 116.

⁷⁰ Se underavsnittet "Registrerades intressen, grundläggande rättigheter och friheter" ovan.

⁷¹ Se avsnitt 2.3 i EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024. Se även skäl 75 i dataskyddsförordningen för ytterligare exempel.

*behandlas under omständigheter där den registrerade inte rimligen kan förvänta sig någon ytterligare behandling*⁷².

92. Rimliga förväntningar är en mycket viktig aspekt i avvägningstestet, inte minst på grund av komplexiteten i den teknik som används i AI-modeller och det faktum att det kan vara svårt för registrerade att förstå de olika potentiella användningsområdena för en AI-modell och den databehandling som krävs⁷³. I detta syfte kan den information som lämnas till de registrerade beaktas för att bedöma om de registrerade rimligen kan förvänta sig att deras personuppgifter kommer att behandlas. Även om utelämnandet av information kan bidra till att de registrerade inte förväntar sig en viss behandling, är enbart uppfyllandet av de insynskrav som anges i dataskyddsförordningen i sig inte tillräckligt för att anse att de registrerade rimligen ska kunna förvänta sig en viss behandling⁷⁴. Att information om utvecklingsfasen för en AI-modell ingår i den personuppgiftsansvariges integritetspolicy innebär inte nödvändigtvis att de registrerade rimligen bör förvänta sig att behandlingen kan ske, utan detta bör analyseras av tillsynsmyndigheterna utifrån de specifika omständigheterna i det enskilda fallet och med beaktande av alla relevanta faktorer.
93. Vid bedömningen av de registrerades rimliga förväntningar i samband med behandling som sker i utvecklingsfasen är det viktigt att beakta de faktorer som nämns i EDPB:s riktlinjer om berättigat intresse⁷⁵. Inom ramen för detta yttrande är det dessutom viktigt att beakta behandlingens vidare sammanhang. Detta kan bland annat omfatta huruvida personuppgifterna var allmänt tillgängliga eller inte, typen av förhållande mellan den registrerade och den personuppgiftsansvarige (och huruvida det finns en koppling mellan dessa två), tjänstens art, det sammanhang i vilket personuppgifterna samlades in, källan från vilken uppgifterna samlades in (till exempel den webbplats eller tjänst från vilken personuppgifterna erhöles och deras sekretessregler), potentiella ytterligare användningsområden för modellen och huruvida de registrerade över huvud taget är medvetna om att deras personuppgifter finns online.
94. Under modellens utvecklingsfas kan de registrerades rimliga förväntningar variera beroende på om de uppgifter som behandlas för att utveckla modellen offentliggörs av de registrerade eller inte. Vidare kan de rimliga förväntningarna också variera beroende på om de tillhandahöll uppgifterna direkt till den personuppgiftsansvarige (t.ex. i samband med deras användning av tjänsten), eller om den personuppgiftsansvarige erhöles dem från en annan källa (till exempel via en tredje part eller webbinsamling). I båda fallen bör de åtgärder som vidtagits för att informera de registrerade om behandlingsaktiviteterna beaktas vid bedömningen av rimliga förväntningar.
95. Under AI-modellens driftsättningsfas är det också viktigt att beakta de registrerades rimliga förväntningar i förhållande till modellens specifika kapacitet. För AI-modeller som anpassar sig

⁷² Se även EU-domstolens dom av den 4 juli 2023 i mål C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 112; EU-domstolens dom av den 11 december 2019 i mål C-708/18, *Asociația de Proprietari bloc M5A-Scara A* (ECLI:EU:C:2019:1064), punkt 58, och EU-domstolens dom av den 4 oktober 2024 i mål C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), punkt 55.

⁷³ I dom av den 4 juli 2023 i mål C-252/21, *Meta mot Bundeskartellamt* (ECLI:EU:C:2023:537), punkt 123, konstaterade EU-domstolen visserligen att "produktförbättring" i princip inte kan uteslutas som ett berättigat intresse, men angav också att det är "[...] tvivelaktigt att produktutvecklingssyftet, med hänsyn till behandlingens omfattning, den väsentliga inverkan som behandlingen har på användaren och den omständigheten att användaren inte rimligen kan förvänta sig att personuppgifterna behandlas [...], skulle kunna väga tyngre än användarens intressen och grundläggande rättigheter – i synnerhet om användaren är ett barn".

⁷⁴ Riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 53.

⁷⁵ Riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkterna 50–54.

beroende på de indata som matats in kan det till exempel vara relevant att överväga om de registrerade varit medvetna om att de har lämnat personuppgifter som gör det möjligt för AI-modellen att anpassa sina svar till deras behov och erbjuda skräddarsydda tjänster. Det kan också vara relevant att överväga om denna behandling endast skulle påverka den tjänst som tillhandahålls de registrerade (till exempel anpassning av innehåll för en viss användare) eller om den skulle användas för att ändra den tjänst som tillhandahålls alla kunder (till exempel för att åstadkomma allmänna förbättringar av modellen). På samma sätt som under utvecklingsfasen kan det också vara särskilt relevant att överväga om det finns en direkt koppling mellan de registrerade och den personuppgiftsansvarige. En sådan direkt koppling kan till exempel göra det möjligt för den personuppgiftsansvarige att på ett enkelt sätt tillhandahålla information till de registrerade om behandlingsaktiviteten och modellen, vilket i sin tur kan påverka dessa registrerades rimliga förväntningar.

Riskreducerande åtgärder

96. När de registrerades intressen, rättigheter och friheter förefaller åsidosätta de berättigade intressen som den personuppgiftsansvarige eller en tredje part har, kan den personuppgiftsansvarige överväga att införa riskreducerande åtgärder för att begränsa behandlingens inverkan på dessa registrerade. Riskreducerande åtgärder är skyddsåtgärder som bör anpassas till omständigheterna i det enskilda fallet och bestäms av en rad faktorer, bland annat AI-modellens avsedda användning. Dessa riskreducerande åtgärder skulle syfta till att säkerställa att den personuppgiftsansvariges eller tredje partens intressen inte åsidosätts och göra det möjligt för den personuppgiftsansvarige att förlita sig på denna rättsliga grund.
97. Såsom påpekas i EDPB:s riktlinjer om berättigat intresse bör begränsande åtgärder inte förväxlas med de åtgärder som den personuppgiftsansvarige enligt lag ändå är skyldig att vidta för att säkerställa efterlevnad av dataskyddsförordningen, oavsett om behandlingen grundar sig på artikel 6.1 f i dataskyddsförordningen⁷⁶. Detta är särskilt viktigt för åtgärder som till exempel kräver att principerna i dataskyddsförordningen följs, såsom principen om uppgiftsminimering.
98. Förteckningen över åtgärder som anges nedan är icke uttömmande och icke-bindande, och vilka åtgärder som sätts in bör övervägas från fall till fall. Beroende på omständigheterna kan vissa av åtgärderna nedan vara nödvändiga för att uppfylla särskilda skyldigheter enligt dataskyddsförordningen, men när så inte är fallet kan de beaktas som möjliga ytterligare skyddsåtgärder. Dessutom rör vissa av nedan nämnda åtgärder områden som är föremål för snabb och ny utveckling och som tillsynsmyndigheterna bör ta hänsyn till när de hanterar varje specifikt fall.
99. **När det gäller utvecklingsfasen för AI-modeller** kan flera åtgärder vidtas för att minska riskerna med behandlingen av både förstaparts- och tredjepartsuppgifter (även för att minska riskerna med webbinsamling). På grundval av ovanstående ger EDPB några exempel på åtgärder som kan vidtas för att minska de risker som identifieras i avvägningstestet och som tillsynsmyndigheterna bör beakta vid bedömningen av specifika AI-modeller från fall till fall.
100. **Tekniska åtgärder:**
 - a. Åtgärder som nämns i avsnitt 3.2.2 och som är lämpliga för att minska de aktuella riskerna, såvida inte dessa åtgärder leder till anonymisering av modellen och inte krävs för att uppfylla andra skyldigheter enligt dataskyddsförordningen eller enligt nödvändighetsprövningen (andra steget i bedömningen av berättigat intresse).

⁷⁶ Riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 57.

101. Utöver dessa kan relevanta åtgärder omfatta följande:

- b. Pseudonymiseringsåtgärder: detta kan till exempel omfatta åtgärder för att förhindra kombinationer av uppgifter på grundval av individuella identifierare. Det kan hända att dessa åtgärder är olämpliga om tillsynsmyndigheten anser att den personuppgiftsansvarige har visat att det finns ett rimligt behov av att samla in olika uppgifter om en viss individ för utvecklingen av AI-systemet eller AI-modellen i fråga.
- c. Åtgärder för att maskera personuppgifter eller ersätta dem med falska personuppgifter i träningsdatasetet (till exempel ersättning av namn och e-postadresser med falska namn och falska e-postadresser). Denna åtgärd kan vara särskilt lämplig när det faktiska innehållet i uppgifterna inte är relevant för den övergripande behandlingen (t.ex. vid träning av stora språkmodeller).

102. **Åtgärder som underlättar utövandet av enskilda personers rättigheter:**

- a. Att låta en rimlig tidsperiod passera mellan insamlingen av ett träningsdataset och dess användning. Denna ytterligare skyddsåtgärd kan göra det möjligt för registrerade att utöva sina rättigheter under denna period. Vad som utgör en rimlig tidsperiod kan bedömas beroende på omständigheterna i varje enskilt fall.
- b. Att erbjuda en ovillkorlig rätt att uteslutas från behandling ("opt-out") redan från början, dvs. att ge de registrerade en skönsmässig rätt att motsätta sig behandlingen innan den äger rum, för att stärka enskilda personers kontroll över sina uppgifter, vilket går utöver villkoren i artikel 21 i dataskyddsförordningen⁷⁷.
- c. Att göra det möjligt för de registrerade att utöva sin rätt till radering även när de särskilda skäl som anges i artikel 17.1 i dataskyddsförordningen inte är tillämpliga⁷⁸.
- d. Att tillåta registrerade att lämna in klagomål om att modellen "återflödar" eller memorerar personuppgifter och de omständigheter och medel genom vilka klagomålen kan återges, så att personuppgiftsansvariga kan reproducera och bedöma relevanta avinläringstekniker för att ta itu med problemen.

103. **Insynsrelaterade åtgärder:** I vissa fall kan riskreducerande åtgärder omfatta åtgärder som ger större insyn i utvecklingen av AI-modellen. Förutom att bidra till efterlevnaden av skyldigheterna i dataskyddsförordningen kan vissa åtgärder också bidra till att komma till rätta med informationsasymmetrin och göra det möjligt för de registrerade att få en bättre förståelse för den behandling som äger rum i utvecklingsfasen:

- a. Offentliggörande av offentlig och lätt tillgänglig kommunikation som går utöver den information som krävs enligt artikel 13 eller 14 i dataskyddsförordningen, till exempel genom att tillhandahålla ytterligare uppgifter om insamlingskriterierna och alla dataset som används, med beaktande av särskilt skydd för barn och sårbara personer.
- b. Alternativa former för att informera de registrerade, till exempel mediekampanjer med olika medier för att informera de registrerade, informationskampanjer via e-post, användning av grafisk visualisering, svar på vanliga frågor, öppenhetsmärkning och modellkort vars systematisering skulle kunna strukturera presentationen av information om AI-modeller, samt årliga rapporter om öppenhet och insyn på frivillig basis.

⁷⁷ Se föregående fotnot.

⁷⁸ Se föregående fotnot.

104. **Specifika riskreducerande åtgärder i samband med webbinsamling:** Med tanke på att webbinsamling, som nämnts ovan, medför särskilda risker⁷⁹ kan särskilda begränsande åtgärder identifieras i detta sammanhang. När en tillsynsmyndighet granskar personuppgiftsansvariga som utför webbinsamling kan myndigheten i tillämpliga fall beakta dessa särskilda åtgärder utöver de begränsande åtgärder som nämnts ovan.
105. Även om de inte är nödvändiga enligt det andra steget i bedömningen av berättigat intresse kan särskilda åtgärder visa sig användbara för att minska riskerna vid webbinsamling. Dessa kan exempelvis omfatta följande **tekniska åtgärder**:
- a. Undvika datainnehåll som kommer från publikationer innehållande personuppgifter som medför risker för särskilda personer eller grupper av personer (till exempel personer som kan vara föremål för fördomar eller utsättas för övergrepp eller fysisk skada om informationen offentliggörs).
 - b. Säkerställa att vissa kategorier av uppgifter inte samlas in eller att vissa källor utesluts vid datainsamlingen. Detta kan till exempel omfatta vissa webbplatser som är särskilt inkräktande på grund av deras känsliga innehåll.
 - c. Undvika insamling från webbplatser (eller delar av webbplatser) som tydligt motsätter sig webbinsamling och vidareutnyttjande av deras innehåll i syfte att bygga upp databaser för AI-träning (till exempel genom att respektera robots.txt- eller ai.txt-filer eller någon annan erkänd mekanism för att uttrycka att automatiska genomsökningar eller webbinsamling inte ska tillämpas på deras innehåll).
 - d. Införa andra relevanta begränsningar av insamlingen, eventuellt även kriterier baserade på tidsperioder.
106. I samband med webbinsamling kan exempel på specifika åtgärder **som underlättar utövandet av enskildas rättigheter och öppenhet** omfatta upprättande av en undantagslista som förvaltas av den personuppgiftsansvarige och som gör det möjligt för registrerade att motsätta sig insamling av deras uppgifter från vissa webbplatser eller onlineplattformar genom att tillhandahålla information som identifierar dem på dessa webbplatser, även innan datainsamlingen sker⁸⁰.
107. **Särskilda överväganden avseende riskreducerande åtgärder i driftsättningsfasen:** Även om vissa av de åtgärder som nämns ovan också kan vara relevanta för driftsättningsfasen (beroende på omständigheterna), tillhandahåller EDPB nedan en icke uttömmande förteckning över ytterligare stödåtgärder som kan genomföras och som bör bedömas av tillsynsmyndigheterna i varje enskilt fall.
- a. **Tekniska åtgärder** kan till exempel vidtas för att förhindra lagring, återflöde eller generering av personuppgifter, särskilt i samband med generativa AI-modeller (till exempel utdatafilter), och/eller för att minska risken för olagligt vidareutnyttjande av AI-modeller för allmänna ändamål (till exempel digital vattenmärkning av AI-genererat innehåll).
 - b. **Åtgärder som underlättar eller påskyndar utövandet av enskilda personers rättigheter** under driftsättningsfasen, utöver vad som krävs enligt lag, särskilt men utan att vara begränsat

⁷⁹ Dessa metoder kan också ge upphov till ytterligare frågor som inte behandlas i detta yttrande, se till exempel Pagallo U., Ciani Sciolla J., *Anatomy of web data scraping: ethics, standards, and the troubles of the law*. European Journal of Privacy Law & Technologies, (2023) 2, s. 1–19, tillgänglig på: <https://doi.org/10.57230/EJPLT232PS>.

⁸⁰ Den personuppgiftsansvarige får då inte längre behandla personuppgifterna såvida denne inte kan påvisa tvingande berättigade skäl för behandlingen som väger tyngre än den registrerades intressen, rättigheter och friheter eller om det sker för fastställande, utövande eller försvar av rättsliga anspråk.

till utövandet av rätten till radering av personuppgifter från modellers utdata eller deduplicering samt tekniker för att ta bort eller radera personuppgifter efter träning av modellen.

108. När tillsynsmyndigheterna utreder användningen av en viss AI-modell bör de överväga om den personuppgiftsansvarige har offentliggjort det avvägningstest som den utfört, med tanke på att detta kan öka insynen och korrektheten. Såsom nämns i EDPB:s riktlinjer om berättigat intresse kan andra åtgärder övervägas för att förse de registrerade med information från avvägningstestet innan någon insamling av personuppgifter sker⁸¹. EDPB upprepar också⁸² att en faktor som bör beaktas är huruvida den personuppgiftsansvarige har gjort dataskyddsombudet delaktigt, i tillämpliga fall.

3.4 Om den möjliga inverkan som olaglig behandling vid utvecklingen av en AI-modell kan ha på lagligheten av den efterföljande behandlingen eller användningen av AI-modellen

109. Detta avsnitt i yttrandet behandlar fråga 4 i begäran. I denna fråga efterfrågas ett förtydligande av den möjliga inverkan som en olaglig behandling i utvecklingsfasen kan ha på den efterföljande behandlingen (dvs. AI-modellens driftsättningsfas) eller på modellens funktion. Frågan avser både den situation där en sådan AI-modell behandlar personuppgifter som lagras i modellen (fråga 4 (i) i begäran) och den situation där ingen behandling av personuppgifter längre äger rum vid användning av AI-modellen (dvs. modellen är anonym) (fråga 4 (ii) i begäran).
110. Innan vissa specifika scenarier behandlas vill EDPB framföra följande allmänna överväganden.
111. För det första kommer de förtydliganden som lämnas i detta avsnitt att inriktas på behandling av personuppgifter i utvecklingsfaser där principen om laglighet enligt artikel 5.1 a och artikel 6 i dataskyddsförordningen inte efterlevs (nedan kallad **olaglighet**)⁸³. På samma sätt kommer EDPB:s överväganden att inriktas på hur den olagliga behandlingen under utvecklingsfasen påverkar lagligheten (dvs. efterlevnaden av artikel 5.1 a och artikel 6 i dataskyddsförordningen) i den efterföljande behandlingen eller driften av modellen. EDPB påpekar dock att den behandling som utförs under utvecklingsfasen också kan leda till överträdelser av andra bestämmelser i dataskyddsförordningen, såsom bristande öppenhet gentemot registrerade eller inbyggt dataskydd och/eller dataskydd som standard, vilka är aspekter som inte analyseras i detta yttrande.
112. För det andra spelar principen om ansvarsskyldighet, som kräver att personuppgiftsansvariga är ansvariga för och kan bevisa efterlevnad av bland annat artikel 5.1 och artikel 6 i dataskyddsförordningen⁸⁴, en nyckelroll när det gäller att hantera denna fråga. Detta gäller även för behovet av att bedöma vilken organisation som är personuppgiftsansvarig för den aktuella behandlingsaktiviteten och om situationer med gemensamt personuppgiftsansvar uppstår (eftersom personuppgiftsansvariga kan vara oupplösligt sammankopplade)⁸⁵. Med tanke på betydelsen av de faktiska omständigheterna i varje enskilt fall, inbegripet den roll som innehas av varje part som deltar

⁸¹ EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 68.

⁸² EDPB:s riktlinjer 1/2024 om behandling av personuppgifter på grundval av artikel 6.1 f i den allmänna dataskyddsförordningen, version 1.0, antagna den 8 oktober 2024, punkt 12.

⁸³ EU-domstolens dom av den 4 maj 2023, mål C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), punkterna 55–57.

⁸⁴ EU-domstolens dom i mål C-60/22 av den 4 maj 2023, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), punkt 53.

⁸⁵ EDPB:s riktlinjer 07/2020 angående begreppen personuppgiftsansvarig och personuppgiftsbiträde i dataskyddsförordningen, version 2.1, antagna den 7 juli 2021, punkt 55.

i behandlingen, bör EDPB:s överväganden betraktas som allmänna observationer som bör bedömas från fall till fall av tillsynsmyndigheterna.

113. För det tredje framhåller EDPB att tillsynsmyndigheterna, i enlighet med artikel 51.1 i dataskyddsförordningen, är *"ansvariga för att övervaka tillämpningen av denna förordning [den allmänna dataskyddsförordningen], i syfte att skydda fysiska personers grundläggande rättigheter och friheter i samband med behandling samt att underlätta det fria flödet av sådana uppgifter inom unionen [...]"*. Det ligger därför inom tillsynsmyndigheternas behörighet att bedöma behandlingens laglighet och att utöva sina befogenheter i enlighet med dataskyddsförordningen och de nationella ramar som de omfattas av⁸⁶. I dylika fall har tillsynsmyndigheterna befogenhet att göra en skönsmässig bedömning av möjliga överträdelser och välja lämpliga, nödvändiga och proportionerliga åtgärder bland dem som nämns i artikel 58 i dataskyddsförordningen, med beaktande av omständigheterna i varje enskilt fall⁸⁷.
114. **Om en överträdelse har konstaterats kan tillsynsmyndigheterna vidta korrigerande åtgärder, exempelvis att förelägga personuppgiftsansvariga att med beaktande av omständigheterna i varje enskilt fall vidta åtgärder för att avhjälpa den ursprungliga behandlingens olaglighet.** Dessa kan till exempel omfatta utfärdande av böter, införande av en tillfällig begränsning av behandlingen, radering av en del av det dataset som behandlades olagligt eller, om detta inte är möjligt och med beaktande av de aktuella omständigheterna samt åtgärdens proportionalitet, föreläggande om radering av hela det dataset som använts för att utveckla AI-modellen och/eller själva AI-modellen. Vid bedömningen av den planerade åtgärdens proportionalitet får tillsynsmyndigheterna beakta åtgärder som kan vidtas av den personuppgiftsansvarige för att avhjälpa den ursprungliga behandlingens olaglighet (t.ex. att träna om modellen).
115. EDPB framhåller också att registrerade kan begära radering av sina personuppgifter om dessa har behandlats olagligt, med förbehåll för de villkor som anges i artikel 17 i dataskyddsförordningen, och att tillsynsmyndigheterna på eget initiativ kan beordra radering av personuppgifterna⁸⁸.
116. Vid bedömningen av om en åtgärd är lämplig, nödvändig och proportionerlig kan tillsynsmyndigheterna bland annat beakta de risker som de registrerade utsätts för, hur allvarlig överträdelsen är, om åtgärden är tekniskt och ekonomiskt genomförbar samt hur många personuppgifter som berörs.
117. Slutligen påminner EDPB om att de åtgärder som tillsynsmyndigheterna vidtar enligt dataskyddsförordningen inte påverkar de åtgärder som vidtas av behöriga myndigheter enligt AI-

⁸⁶ Särskilda nationella bestämmelser kan behöva beaktas. Se till exempel artikel 2-decies i den italienska dataskyddslagen (lagstiftningsdekret 196/2003), där det fastställs att uppgifter som behandlas i strid med dataskyddsbestämmelserna inte får användas. Detta påverkar inte tillämpningen av andra nationella rättsliga ramar, till exempel straffrätten.

⁸⁷ Se i detta avseende skäl 129 i dataskyddsförordningen samt EU-domstolens dom av den 26 september 2024 i mål C-768-21, *TR mot Land Hessen* (ECLI:EU:C:2024:785), punkt 37, EU-domstolens dom av den 7 december 2023 i de förenade målen C-26/22 och C-64/22, *SCHUFA Holding (Libération de reliquat de dette)* (ECLI:EU:C:2023:958), punkt 57, och EU-domstolens dom av den 14 mars 2024 i mål C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), punkt 34.

⁸⁸ Se i detta avseende punkt 28 i EDPB:s yttrande 39/2021 om huruvida artikel 58.2 g i den allmänna dataskyddsförordningen skulle kunna fungera som rättslig grund för en tillsynsmyndighet att på eget initiativ beordra radering av personuppgifter i en situation där en sådan begäran inte lämnats in av den registrerade. Se även i detta avseende EU-domstolens dom av den 14 mars 2024 i mål C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), punkt 42.

förordningen och/eller enligt andra tillämpliga rättsliga ramar (t.ex. lagstiftning om civilrättsligt ansvar).

118. I de följande avsnitten kommer EDPB att behandla tre scenarier som omfattas av fråga 4 i begäran, där skillnaderna ligger i huruvida de personuppgifter som behandlas för att utveckla modellen kvarhålls i modellen och/eller huruvida den efterföljande behandlingen utförs av samma eller en annan personuppgiftsansvarig.

3.4.1 Fall 1. En personuppgiftsansvarig behandlar olagligen personuppgifter för att utveckla modellen. Personuppgifterna kvarhålls i modellen och behandlas därefter av samma personuppgiftsansvarige (till exempel i samband med att modellen tas i bruk).

119. Detta scenario avser fråga 4 (i) i begäran och gäller en situation där en personuppgiftsansvarig olagligt behandlar personuppgifter (dvs. underlåter att följa artikel 5.1 a och artikel 6 i dataskyddsförordningen) för att utveckla en AI-modell som sparar information som rör en identifierad eller identifierbar fysisk person och därmed inte är anonym. Personuppgifterna behandlas därefter av samma personuppgiftsansvarig (till exempel i samband med driftsättningen av modellen). När det gäller detta scenario konstaterar EDPB följande:
120. Tillsynsmyndighetens befogenhet att införa korrigerande åtgärder för den ursprungliga behandlingen (såsom förklaras i punkterna 113, 114 och 115 ovan) skulle i princip påverka den efterföljande behandlingen (om tillsynsmyndigheten t.ex. beordrar den personuppgiftsansvarige att radera de personuppgifter som behandlades olagligt skulle dessa korrigerande åtgärder inte göra det möjligt för den personuppgiftsansvarige att därefter behandla de personuppgifter som var föremål för åtgärderna).
121. När det gäller konsekvenserna av olaglig behandling i utvecklingsfasen för efterföljande behandling (till exempel i driftsättningsfasen) påminner EDPB om att det är upp till tillsynsmyndigheterna att göra en analys som tar hänsyn till de särskilda omständigheterna i varje enskilt fall.
122. **Huruvida utvecklings- och driftsättningsfaserna har separata syften (och därmed utgör separata behandlingsaktiviteter) och i vilken utsträckning avsaknaden av rättslig grund för den inledande behandlingsaktiviteten påverkar lagligheten av den efterföljande behandlingen bör bedömas från fall till fall beroende på sammanhanget i det enskilda fallet.**
123. När den efterföljande behandlingen grundar sig på berättigat intresse bör det faktum att den ursprungliga behandlingen var olaglig beaktas vid bedömningen av berättigat intresse (t.ex. med hänsyn till riskerna för de registrerade eller det faktum att de registrerade kanske inte förväntar sig en sådan efterföljande behandling), med särskild hänsyn till den rättsliga grunden i artikel 6.1 f i dataskyddsförordningen. I dessa fall kan den olagliga behandlingen under utvecklingsfasen påverka lagligheten av den efterföljande behandlingen.

3.4.2 Fall 2. En personuppgiftsansvarig behandlar olagligt personuppgifter för att utveckla modellen. Personuppgifterna sparas i modellen och behandlas sedan av en annan personuppgiftsansvarig i samband med driftsättningen av modellen.

124. Detta scenario hänför sig till fråga 4 (i) i begäran. Detta skiljer sig från scenario 1 (i avsnitt 3.4.1 i detta yttrande) i så måtto att personuppgifterna i det senare skedet behandlas av en annan personuppgiftsansvarig inom ramen för driftsättningen av AI-modellen.
125. EDPB framhåller att fastställandet av de roller som tilldelats dessa olika aktörer inom dataskyddsramen är ett viktigt steg för att identifiera vilka skyldigheter som gäller enligt dataskyddsförordningen och

vem som är ansvarig för vad, samt att situationer med gemensamt personuppgiftsansvar också bör beaktas vid bedömningen av varje parts ansvar enligt dataskyddsförordningen. Observationerna som anges nedan bör därför betraktas som allmänna iakttagelser som bör beaktas av tillsynsmyndigheterna när så är relevant. När det gäller scenario 2 konstaterar EDPB följande:

126. För det första bör det erinras om att enligt artikel 5.1 a i dataskyddsförordningen, tolkad mot bakgrund av artikel 5.2 i samma förordning, ska varje personuppgiftsansvarig säkerställa att den behandling som den utför är laglig och dessutom kunna visa detta. Därför bör tillsynsmyndigheterna bedöma lagligheten av den behandling som utförs av i) den personuppgiftsansvarige som ursprungligen utvecklar AI-modellen, och ii) den personuppgiftsansvarige som förvärvar AI-modellen och därefter själv behandlar personuppgifterna.
127. För det andra är de överväganden som gjorts i punkterna 113, 114 och 115 ovan relevanta i detta fall, när det gäller tillsynsmyndigheternas befogenhet att ingripa i samband med den inledande behandlingen. Artikel 17.1 d i dataskyddsförordningen (om radering av olagligt behandlade uppgifter) och artikel 19 i samma förordning (om anmälningsskyldighet avseende rättelse eller radering av personuppgifter och begränsning av behandling) kan, beroende på omständigheterna i det enskilda fallet, också vara relevanta i detta sammanhang, till exempel när det gäller den anmälan som den personuppgiftsansvarige som utvecklar modellen ska göra till den personuppgiftsansvarige som tar modellen i bruk.
128. För det tredje, när det gäller den eventuella inverkan av den ursprungliga behandlingens olaglighet på den efterföljande behandling som utförs av en annan personuppgiftsansvarig, bör en sådan bedömning göras av tillsynsmyndigheterna med beaktande av omständigheterna i varje enskilt fall.
129. **Tillsynsmyndigheterna bör beakta huruvida den personuppgiftsansvarige som driftsätter modellen har gjort en lämplig bedömning, som en del av sina ansvarsskyldigheter⁸⁹, för att visa att artikel 5.1 a och artikel 6 i dataskyddsförordningen efterlevs och därmed säkerställer att AI-modellen inte har utvecklats genom olaglig behandling av personuppgifter.** En sådan utvärdering från tillsynsmyndigheternas sida bör ta hänsyn till huruvida den personuppgiftsansvarige har bedömt vissa icke uttömmande kriterier, såsom källan till uppgifterna och huruvida AI-modellen är resultatet av en överträdelse av dataskyddsförordningen, särskilt om detta fastställts av en tillsynsmyndighet eller en domstol, så att den personuppgiftsansvarige som driftsätter modellen inte kan ignorera att den inledande behandlingen varit olaglig.
130. Den personuppgiftsansvarige bör till exempel överväga om uppgifterna härrör från en personuppgiftsincident eller om behandlingen varit föremål för ett konstaterande av en överträdelse från en tillsynsmyndighet eller en domstol. **Omfattningen av den personuppgiftsansvariges bedömning och detaljnivån som förväntas av tillsynsmyndigheterna kan variera beroende på olika faktorer, däribland typen och graden av risk som behandlingen i AI-modellen ger upphov till när den tas i drift i förhållande till de registrerade vars uppgifter användes för att utveckla modellen.**
131. EDPB noterar att AI-förordningen kräver att leverantörer av sådana AI-system som innebär hög risk upprättar en EU-försäkran om överensstämmelse⁹⁰ och att denna försäkran innehåller en förklaring om att AI-systemet i fråga efterlever EU:s dataskyddslagstiftning⁹¹. EDPB noterar att en sådan försäkran inte utgör något slutgiltigt bevis på att dataskyddsförordningen efterlevs. Tillsynsmyndigheterna kan dock beakta sådana EU-försäkringar när de undersöker AI-modeller.

⁸⁹ Artikel 5.2 och artikel 24 i den allmänna dataskyddsförordningen.

⁹⁰ Artiklarna 16 g och 47 i AI-förordningen.

⁹¹ Bilaga V, punkt 5 i AI-förordningen.

132. De överväganden som gjordes i punkt 123 ovan är även relevanta i detta fall. När tillsynsmyndigheterna kontrollerar om och hur den personuppgiftsansvarige har bedömt det berättigade intressets lämplighet som rättslig grund för den behandling som utförs, bör olagligheten i den inledande behandlingen beaktas som en del av bedömningen av berättigat intresse. Det potentiella risker som kan uppstå för de registrerade vars personuppgifter behandlades olagligt för att utveckla modellen bör till exempel bedömas. Under avvägningstestet måste olika aspekter beaktas. De kan vara av både teknisk natur (t.ex. förekomsten av filter eller åtkomstbegränsningar som införts under utvecklingen av modellen och som den efterföljande personuppgiftsansvarige inte kan kringgå eller påverka och som kan förhindra åtkomst till eller utlämnande av personuppgifter) och av rättslig natur (t.ex. typ av olaglighet vid den inledande behandlingen och hur allvarlig den var).

3.4.3 Fall 3. En personuppgiftsansvarig behandlar olagligen personuppgifter för att utveckla modellen och ser sedan till att modellen anonymiseras innan samma eller en annan personuppgiftsansvarig påbörjar en annan behandling av personuppgifter i samband med driftsättningen.

133. Detta scenario avser fråga 4 (ii) i begäran och avser ett fall där en personuppgiftsansvarig olagligen behandlar personuppgifter för att utveckla AI-modellen, men gör det på ett sätt som säkerställer att personuppgifterna anonymiseras, innan samma eller en annan personuppgiftsansvarig påbörjar en annan behandling av personuppgifter i samband med driftsättningen. För det första påminner EDPB om att tillsynsmyndigheterna är behöriga och har befogenhet att ingripa när det gäller behandling i samband med anonymisering av modellen samt när det gäller den behandling som utförs under utvecklingsfasen. Således kan tillsynsmyndigheterna, beroende på de särskilda omständigheterna i fallet, införa korrigerande åtgärder för denna inledande behandling (enligt vad som förklaras i punkterna 113, 114 och 115 ovan).
134. Om det kan visas att den efterföljande driften av AI-modellen inte medför behandling av personuppgifter anser EDPB att dataskyddsförordningen inte är tillämplig⁹². Det faktum att den inledande behandlingen var olaglig bör därför inte påverka den efterföljande användningen av modellen i detta fall. EDPB betonar dock att enbart ett påstående om modellens anonymitet inte är tillräckligt för att undanta den från tillämpningen av dataskyddsförordningen, och noterar att tillsynsmyndigheterna bör bedöma detta genom att i varje enskilt fall ta hänsyn till de överväganden som EDPB har lämnat för att behandla fråga 1 i begäran.
135. **Om de personuppgiftsansvariga därefter behandlar personuppgifter som samlats in under driftsättningsfasen, efter det att modellen har anonymiserats, är dataskyddsförordningen tillämplig på dessa behandlingsaktiviteter. Enligt dataskyddsförordningen bör lagligheten i den behandling som utförs i driftsättningsfasen i dessa fall inte påverkas av att den inledande behandlingen var olaglig.**

⁹² Skäl 26 i den allmänna dataskyddsförordningen.

4 Avslutande anmärkningar

136. Detta yttrande riktas till samtliga tillsynsmyndigheter och kommer att offentliggöras i enlighet med artikel 64.5 b i dataskyddsförordningen.

För Europeiska dataskyddsstyrelsen

Ordförande

Anu Talus