

Databeskyttelsesrådets udtalelser (artikel 64)



Udtalelse 28/2024 om visse databeskyttelsesaspekter vedrørende behandling af personoplysninger i forbindelse med AI-modeller

Vedtaget den 17. december 2024

Translations proofread by EDPB Members.

This language version has not yet been proofread.

Resumé

AI-teknologier skaber mange muligheder og fordele på tværs af en lang række sektorer og sociale aktiviteter.

Ved at beskytte den grundlæggende ret til databeskyttelse støtter GDPR disse muligheder og fremmer andre grundlæggende rettigheder i EU, herunder retten til at tænke frit, retten til ytrings- og informationsfrihed, retten til uddannelse eller friheden til at oprette og drive egen virksomhed. På denne måde er GDPR en retlig ramme, der tilskynder til ansvarlig innovation.

I denne forbindelse og under hensyntagen til de databeskyttelsesspørgsmål, som disse teknologier rejser, anmodede den irske tilsynsmyndighed Databeskyttelsesrådet om at afgive en udtalelse om almengyldige spørgsmål i henhold til artikel 64, stk. 2, i GDPR. Anmodningen vedrører behandling af personoplysninger i forbindelse med udviklingsfasen og idriftsættelsesfasen for modeller for kunstig intelligens ("AI"). I anmodningen blev der stillet følgende spørgsmål: 1) hvornår og hvordan kan en AI-model betragtes som "anonym", 2) hvordan kan dataansvarlige påvise, at en legitim interesse er relevant som retsgrundlag i udviklingsfasen og 3) idriftsættelsesfasen, og 4) hvilke konsekvenser har ulovlig behandling af personoplysninger i udviklingsfasen af en AI-model for den efterfølgende behandling eller anvendelse af AI-modellen.

Med hensyn til spørgsmål 1 nævnes det i udtalelsen, at påstande om en AI-models anonymitet bør vurderes af de kompetente tilsynsmyndigheder fra sag til sag, eftersom Databeskyttelsesrådet mener, at AI-modeller, der er trænet med personoplysninger, ikke i alle tilfælde kan betragtes som anonyme. For at en AI-model kan betragtes som anonym, bør 1) sandsynligheden for direkte (herunder probabilistisk) udtræk af personoplysninger vedrørende personer, hvis personoplysninger er blev anvendt til at udvikle modellen, og 2) sandsynligheden for forsætlig eller uforsætlig indhentning af sådanne personoplysninger gennem forespørgsler være meget ubetydelig, under hensyntagen til "*alle de midler, som med rimelighed kan forventes at blive anvendt*" af den dataansvarlige eller en anden person.

For at foretage en vurdering bør tilsynsmyndighederne gennemgå den dokumentation, som den dataansvarlige har indleveret for at vise, at modellen er anonym. I den forbindelse indeholder udtalelsen en ikkebindende og ikkeudtømmende liste over metoder, der kan anvendes af dataansvarlige i deres påvisning af anonymitet, og dermed tages i betragtning af tilsynsmyndighederne, når de vurderer en dataansvarligs påstand om anonymitet. Dette omfatter bl.a. de tilgange, som dataansvarlige anvender i udviklingsfasen for at forhindre eller begrænse indsamlingen af personoplysninger, der anvendes til træning, for at reducere deres identificerbarhed, for at forhindre udtræk af dem eller for at sikre, at den seneste teknik for at afværge trusler og angreb anvendes.

Med hensyn til spørgsmål 2 og 3 indeholder udtalelsen generelle betragtninger, som tilsynsmyndighederne skal tage hensyn til, når de vurderer, om dataansvarlige kan påberåbe sig legitim interesse som et relevant retsgrundlag for behandling, der foretages i forbindelse med udvikling og idriftsættelse af AI-modeller.

I udtalelsen mindes der om, at der ikke er noget hierarkisk forhold mellem de retsgrundlag, der er fastsat i GDPR, og at det er op til dataansvarlige at udpege det relevante retsgrundlag for deres behandlingsaktiviteter. Udtalelsen henviser derefter til den tretrinshenvisning, der bør foretages ved vurderingen af anvendelsen af legitim interesse som retsgrundlag. Testen skal 1) kortlægge den

dataansvarliges eller en tredjeparts legitime interesse, 2) analysere nødvendigheden af behandlingen med henblik på den eller de legitime interesser, der forfølges (også kaldet "nødvendighedstesten"), og 3) vurdere, om den legitime interesse/de legitime interesser tilsidesætter de registreredes interesser eller grundlæggende rettigheder og frihedsrettigheder (også kaldet "afvejningstesten").

Med hensyn til det første trin mindes der i udtalelsen om, at en interesse kan betragtes som legitim, hvis følgende tre kumulative kriterier er opfyldt: interessen 1) er lovlig, 2) er klart og præcist formuleret og 3) er reel og aktuel (dvs. ikke spekulativ). En sådan interesse kan i forbindelse med udviklingen af en AI-model f.eks. omfatte udvikling af en samtalerobottjeneste til brugerne, og i forbindelse med idriftsættelsen kan det omfatte forbedring af trusselsopsporing i et informationssystem.

Med hensyn til det andet trin mindes der i udtalelsen om, at nødvendighedsvurderingen indebærer, at der tages hensyn til: 1) om behandlingsaktiviteten vil gøre det muligt at forfølge den legitime interesse, og 2) om der ikke findes nogen mindre indgribende måde at forfølge denne interesse på. Ved vurderingen af, om betingelsen om nødvendighed er opfyldt, bør tilsynsmyndighederne være særligt opmærksomme på mængden af behandlede personoplysninger, og om det er forholdsmæssigt afpasset i forhold til at forfølge den legitime interesse, der gør sig gældende, også set i lyset af princippet om dataminimering.

Med hensyn til det tredje trin mindes der i udtalelsen om, at afvejningstesten skal udføres under hensyntagen til de specifikke omstændigheder i hver enkelt sag. Desuden gives der en oversigt over de elementer, som tilsynsmyndighederne kan tage hensyn til, når de vurderer, om den registreredes interesser, grundlæggende rettigheder og frihedsrettigheder går forud for den dataansvarliges eller en tredjeparts interesser.

Som en del af det tredje trin fremhæver udtalelsen specifikke risici for de grundlæggende rettigheder, der kan opstå enten i udviklings- eller idriftsættelsesfasen af AI-modeller. Den præciserer også, at den behandling af personoplysninger, der finder sted i udviklings- og idriftsættelsesfasen af AI-modeller, kan påvirke de registrerede på forskellige måder, som kan være positive eller negative. For at vurdere en sådan indvirkning kan tilsynsmyndighederne tage hensyn til arten af de data, der behandles af modellerne, konteksten for behandlingen og de mulige yderligere konsekvenser af behandlingen.

Udtalelsen fremhæver desuden den rolle, som de registreredes rimelige forventninger spiller i afvejningstesten. Dette kan være vigtigt på grund af kompleksiteten af de teknologier, der anvendes i AI-modeller, og det faktum, at det kan være vanskeligt for de registrerede at forstå de mange forskellige mulige anvendelser af dem samt de forskellige behandlingsaktiviteter, det omfatter. I denne henseende kan både de oplysninger, der gives til de registrerede, og baggrunden for behandlingen være blandt de elementer, der skal tages i betragtning for at vurdere, om de registrerede med rimelighed kan forvente, at deres personoplysninger behandles. I forhold til konteksten kan det f.eks. være: om personoplysningerne var offentligt tilgængelige eller ej, arten af forholdet mellem den registrerede og den dataansvarlige (og om der er en forbindelse mellem de to), tjenestens art, den kontekst, hvori personoplysningerne blev indsamlet, den kilde, hvorfra oplysningerne blev indsamlet (dvs. det websted eller den tjeneste, hvor personoplysningerne blev indsamlet, og de privatlivsindstillinger, de tilbyder), de potentielle yderligere anvendelser af modellen, og om de registrerede rent faktisk er klar over, at deres personoplysninger overhovedet er online.

I udtalelsen mindes der også om, at når de registreredes interesser, rettigheder og frihedsrettigheder synes at gå forud for den dataansvarliges eller en tredjeparts legitime interesse/legitime interesser, kan den dataansvarlige overveje at indføre afbødende foranstaltninger for at begrænse behandlingens indvirkning på de pågældende registrerede. Afbødende foranstaltninger bør ikke forveksles med de foranstaltninger, som den dataansvarlige under alle omstændigheder er retligt forpligtet til at træffe

for at sikre overholdelse af GDPR. Desuden bør foranstaltningerne afpasses sagens omstændigheder og AI-modellens karakteristika, herunder dens tilsigtede anvendelse. I den henseende indeholder udtalelsen en ikke-udtømmende liste over eksempler på afhjælpende foranstaltninger i forbindelse med udviklingsfasen (også med hensyn til webskrabning) og idriftsættelsesfasen. Afbødende foranstaltninger kan videreudvikles hurtigt og bør tilpasses sagens omstændigheder. Det er derfor fortsat op til tilsynsmyndighederne at vurdere, om de afbødende foranstaltninger, der er truffet i hvert enkelt tilfælde, er hensigtsmæssige.

Med hensyn til spørgsmål 4 mindes der i udtalelsen generelt om, at tilsynsmyndighederne har skønsbeføjelser til at vurdere de mulige overtrædelser og vælge passende, nødvendige og forholdsmæssige foranstaltninger under hensyntagen til omstændighederne i hver enkelt sag. I udtalelsen behandles tre scenarier.

I scenario 1 opbevares personoplysninger i AI-modellen (hvilket betyder, at modellen ikke kan betragtes som anonym, som beskrevet i det første spørgsmål) og behandles efterfølgende af den samme dataansvarlige (f.eks. i forbindelse med idriftsættelsen af modellen). I udtalelsen angives, at det i hvert enkelt tilfælde – og afhængigt af konteksten – bør vurderes, hvorvidt udviklingsfasen og idriftsættelsesfasen har separate formål (og dermed udgør separate behandlingsaktiviteter), og i hvilket omfang manglen på retsgrundlag for den indledende behandlingsaktivitet påvirker lovligheden af den efterfølgende behandling.

I scenario 2 opbevares personoplysninger i modellen og behandles af en anden dataansvarlig i forbindelse med idriftsættelsen af modellen. I den henseende anføres det i udtalelsen, at tilsynsmyndighederne bør tage hensyn til, om den dataansvarlige, der anvender modellen, gennemførte en passende vurdering som led i sine ansvarlighedsforpligtelser til at påvise overholdelse af artikel 5, stk. 1, litra a), og artikel 6 i GDPR, for at sikre at AI-modellen ikke blev udviklet ved ulovlig behandling af personoplysninger. Denne vurdering bør f.eks. tage hensyn til kilden til personoplysningerne og til, om behandlingen i udviklingsfasen var genstand for en konstateret overtrædelse, især hvis den blev fastslået af en tilsynsmyndighed eller en domstol. Den bør desuden være mere eller mindre detaljeret afhængigt af de risici, der er forbundet med behandlingen i idriftsættelsesfasen.

I scenario 3 behandler en dataansvarlig personoplysninger ulovligt for at udvikle AI-modellen og sikrer derefter, at de anonymiseres, inden den samme eller en anden dataansvarlig indleder en anden behandling af personoplysninger i forbindelse med idriftsættelsen. I denne henseende anføres det i udtalelsen, at Databeskyttelsesrådet ikke mener, at GDPR finder anvendelse, hvis det kan påvises, at den efterfølgende anvendelse af AI-modellen ikke indebærer behandling af personoplysninger. At den oprindelige behandling var ulovlig, bør ikke påvirke den efterfølgende anvendelse af modellen. Databeskyttelsesrådet tilføjer, at når dataansvarlige efterfølgende behandler personoplysninger, der er indsamlet i idriftsættelsesfasen, efter at modellen er blevet anonymiseret, bør GDPR finde anvendelse i forbindelse med denne behandling. Hvad angår GDPR mener Databeskyttelsesrådet i disse tilfælde, at lovligheden af den behandling, der er udført i idriftsættelsesfasen, ikke påvirkes af, at den oprindelige behandling var ulovlig.

Indholdsfortegnelse

1	Introduktion	6
1.1	Resumé af de faktiske omstændigheder	6
1.2	Antagelighed af anmodningen om en udtalelse i henhold til artikel 64, stk. 2, i GDPR	8
2	Anvendelsesområde og centrale begreber	9
2.1	Udtalelsens omfang	9
2.2	Centrale begreber	11
2.3	AI-modeller inden for rammerne af udtalelsen	12
3	Anmodningens berettigelse	13
3.1	AI-modellers egenskaber i forhold til definitionen af personoplysninger	13
3.2	De omstændigheder, hvorunder AI-modeller kan betragtes som anonyme, og dokumentationen herfor	14
3.2.1	Generelle overvejelser vedrørende anonymisering i den foreliggende sammenhæng	15
3.2.2	Elementer til evaluering af den resterende sandsynlighed for identifikation	17
3.3	Legitim interesse som passende retsgrundlag for behandling af personoplysninger i forbindelse med udvikling og idriftsættelse af AI-modeller	20
3.3.1	Generelle bemærkninger	20
3.3.2	Overvejelser om de tre trin i vurderingen af den legitime interesse i forbindelse med udviklingen og idriftsættelsen af AI-modeller	22
3.4	Mulig indvirkning af ulovlig behandling i forbindelse med udviklingen af en AI-model på lovligheden af den efterfølgende behandling eller anvendelse af AI-modellen	32
3.4.1	Scenarie 1. En dataansvarlig behandler ulovligt personoplysninger for at udvikle modellen, personoplysningerne opbevares i modellen og behandles efterfølgende af den samme dataansvarlige (f.eks. i forbindelse med idriftsættelsen af modellen).	34
3.4.2	Scenarie 2. En dataansvarlig behandler personoplysninger ulovligt for at udvikle modellen, personoplysningerne opbevares i modellen og behandles af en anden dataansvarlig i forbindelse med idriftsættelsen af modellen	35
3.4.3	Scenarie 3. En dataansvarlig behandler personoplysninger ulovligt for at udvikle modellen og sikrer derefter, at modellen anonymiseres, inden den samme eller en anden dataansvarlig indleder en anden behandling af personoplysninger i forbindelse med idriftsættelsen	36
4	Afsluttende bemærkninger	37

Det Europæiske Databeskyttelsesråd har —

under henvisning til artikel 63 og artikel 64, stk. 2, i Europa-Parlamentets og Rådets forordning 2016/679/EU af 27. april 2016 om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger og om fri udveksling af sådanne oplysninger og om ophævelse af direktiv 95/46/EF (herefter "GDPR"),

under henvisning til EØS-aftalen, særlig bilag XI og protokol 37, som ændret ved Det Blandede EØS-Udvalgs afgørelse nr. 154/2018 af 6. juli 2018,¹

under henvisning til artikel 10 og artikel 22 i forretningsordenen,

som tager følgende i betragtning:

1) Det Europæiske Databeskyttelsesråd (herefter "**Databeskyttelsesrådet**") har som sin vigtigste opgave at sikre en konsekvent anvendelse af GDPR i hele Det Europæiske Økonomiske Samarbejdsområde ("**EØS**"). I henhold til artikel 64, stk. 2, i GDPR kan en ("**tilsynsmyndighed**"), formanden for Databeskyttelsesrådet eller Kommissionen kræve, at ethvert almengyldigt spørgsmål eller ethvert spørgsmål, der har virkninger i mere end én EØS-medlemsstat, drøftes af Databeskyttelsesrådet med henblik på en udtalelse. Formålet med denne udtalelse er at undersøge et almengyldigt spørgsmål eller et spørgsmål, der har virkninger i mere end én EØS-medlemsstat.

2) Databeskyttelsesrådets udtalelse vedtages i overensstemmelse med artikel 64, stk. 3, i GDPR sammenholdt med artikel 10, stk. 2, i Databeskyttelsesrådets forretningsorden inden for otte uger regnet fra det tidspunkt, hvor formanden og de kompetente tilsynsmyndigheder har besluttet, at sagsakterne er komplette. Efter formandens afgørelse kan denne frist forlænges med yderligere seks uger under hensyntagen til spørgsmålets kompleksitet –

HAR VEDTAGET FØLGENDE UDTALELSE:

1 Introduktion

1.1 Resumé af de faktiske omstændigheder

1. Den 4. september 2024 anmodede den irske tilsynsmyndighed "herefter kaldet **den anmodende tilsynsmyndighed**") Databeskyttelsesrådet om at afgive en udtalelse i henhold til artikel 64, stk. 2, i GDPR vedrørende AI-modeller og behandling af personoplysninger (herefter kaldt anmodningen "**anmodningen**").
2. Formanden for Databeskyttelsesrådet og den irske tilsynsmyndighed anså sagsakterne for at være komplette den 13. september 2024. Den følgende arbejdsdag, den 16. september 2024, blev sagsakterne udsendt af Databeskyttelsesrådets sekretariat. I betragtning af sagens kompleksitet besluttede formanden for Databeskyttelsesrådet at forlænge den lovbestemte frist i overensstemmelse med artikel 64, stk. 3, i GDPR og artikel 10, stk. 4, i forretningsordenen for Databeskyttelsesrådet.

¹ Henvisninger til "medlemsstater" i denne udtalelse skal forstås som henvisninger til "EØS-medlemsstater". Henvisninger til "Unionen" i denne udtalelse skal forstås som henvisninger til "EØS".

3. Anmodningen omhandler visse elementer af træningen, opdateringen, udviklingen og anvendelsen af AI-modeller, hvor personoplysninger udgør en del af det relevante datasæt. Den irske tilsynsmyndighed fremhæver, at anmodningen vedrører centrale spørgsmål, der har stor indvirkning på registrerede og dataansvarlige i EØS, og at der på nuværende tidspunkt ikke er nogen harmoniseret holdning blandt de nationale tilsynsmyndigheder². Den terminologi, der anvendes i forbindelse med denne udtalelse, findes i afsnit 2.2 og 2.3 nedenfor.

4. Den irske tilsynsmyndighed stillede følgende spørgsmål:

Spørgsmål 1: Anses den endelige AI-model, som er blevet trænet ved hjælp af personoplysninger, i alle tilfælde for ikke at opfylde definitionen af personoplysninger (som fastsat i artikel 4, stk. 1, i GDPR)?

Såfremt spørgsmål 1 besvares bekræftende:

- i. På hvilket stadium af de behandlingsaktiviteter, der fører til en AI-model, behandles personoplysninger ikke længere?
 - a) Hvordan kan det påvises, at AI-modellen ikke behandler personoplysninger?
- ii. Er der nogen faktorer, som ville medføre, at anvendelsen af den endelige AI-model ikke længere kan betragtes som anonym?
 - a) Hvis ja, hvordan kan de foranstaltninger, der er truffet for at afbøde, forebygge eller beskytte mod disse faktorer (for at sikre, at AI-modellen ikke behandler personoplysninger), påvises?

Såfremt spørgsmål 1 besvares benægtende:

- i. Under hvilke omstændigheder kan dette ske?
 - a) I bekræftende fald, hvordan kan de tiltag, der træffes for at sikre, at AI-modellen ikke behandler personoplysninger, påvises?

Spørgsmål 2: Hvis en dataansvarlig baserer sig på legitime interesser som retsgrundlag for behandling af personoplysninger med henblik på at fremstille, opdatere og/eller udvikle en AI-model, hvordan bør den pågældende dataansvarlige så påvise, at legitime interesser er hensigtsmæssige som retsgrundlag, både i forbindelse med behandling af tredjepartsdata og førstepartsdata?

- i. Hvilke overvejelser bør den dataansvarlige tage hensyn til for at sikre, at de registreredes interesser, hvis personoplysninger behandles, er passende afstemt mod den dataansvarliges interesser i forbindelse med:
 - a) Tredjepartsdata
 - b) Førstepartsdata

Spørgsmål 3: Hvordan bør en dataansvarlig påvise, at legitime interesser er hensigtsmæssige som retsgrundlag, hvis en dataansvarlig efter træning baserer sig på legitime interesser som retsgrundlag for behandling af personoplysninger, der finder sted inden for rammerne af en AI-model, eller et AI-system, som en AI-model er en del af?

Spørgsmål 4: Hvis det konstateres, at en AI-model er blevet skabt, opdateret eller udviklet ved hjælp af ulovligt behandlede personoplysninger, hvilken indvirkning har dette i givet fald på lovligheden af

² Anmodning, s. 1.

den fortsatte eller efterfølgende behandling eller anvendelse af AI-modellen, enten alene eller som en del af et AI-system, hvor:

- i. AI-modellen, enten alene eller som en del af et AI-system, behandler personoplysninger?
- ii. Hverken AI-modellen, alene eller som en del af et AI-system, behandler personoplysninger?

1.2 Antagelighed af anmodningen om en udtalelse i henhold til artikel 64, stk. 2, i GDPR

5. Artikel 64, stk. 2, i GDPR fastsætter navnlig, at en tilsynsmyndighed kan kræve, at ethvert almengyldigt spørgsmål eller ethvert spørgsmål, der har virkninger i mere end én medlemsstat, drøftes af Databeskyttelsesrådet med henblik på en udtalelse.
6. Den anmodende tilsynsmyndighed stillede spørgsmål til Databeskyttelsesrådet vedrørende databeskyttelsesaspekter i forbindelse med AI-modeller. Den præciserede i anmodningen, at selv om mange organisationer nu anvender AI-modeller, herunder store sprogmodeller ("**LLM-modeller**"), giver aktiviteter, træning og anvendelse i forbindelse hermed anledning til "*en række vidtrækkende databeskyttelsesproblemer*"³, som "*påvirker registrerede i hele EU/EØS*"⁴.
7. Anmodningen rejser i det væsentlige spørgsmål om i) anvendelsen af begrebet personoplysninger, ii) princippet om lovlighed, navnlig med hensyn til retsgrundlaget for legitim interesse, i forbindelse med AI-modeller samt om iii) konsekvenserne af ulovlig behandling af personoplysninger i udviklingsfasen af AI-modeller for den efterfølgende behandling eller anvendelse af modellen.
8. Databeskyttelsesrådet mener, at anmodningen vedrører et "*almengyldigt spørgsmål*" i henhold til artikel 64, stk. 2, i GDPR. Spørgsmålet vedrører navnlig fortolkningen og anvendelsen af artikel 4, stk. 1, artikel 5, stk. 1, litra a), og artikel 6 i GDPR i forbindelse med behandling af personoplysninger ved udvikling og idriftsættelse af AI-modeller. Som fremhævet af den anmodende tilsynsmyndighed rejser anvendelsen af disse bestemmelser på AI-modeller systemiske, teoretiske og nye spørgsmål⁵. Den hurtige udvikling og idriftsættelse af AI-modeller i flere og flere organisationer rejser særlige spørgsmål, og som det påpeges i anmodningen, "*vil Databeskyttelsesrådet have stor gavn af at nå frem til en fælles holdning til de spørgsmål, der rejses i denne anmodning, da sådanne spørgsmål er centrale for Databeskyttelsesrådets planlagte arbejde på kort og mellemlang sigt*"⁶. Desuden skaber AI-teknologier mange muligheder og fordele på tværs af et bredt spektrum af sektorer og sociale aktiviteter. Derudover er GDPR en juridisk ramme, der tilskynder til ansvarlig innovation. Det følger heraf, at der er en almen interesse i at foretage denne vurdering i form af en udtalelse fra Databeskyttelsesrådet for at sikre en konsekvent anvendelse af visse bestemmelser i GDPR i forbindelse med AI-modeller.
9. Den alternative betingelse i artikel 64, stk. 2, i GDPR omhandler spørgsmål, "*der har virkninger i mere end én medlemsstat*". Databeskyttelsesrådet minder om, at udtrykket "virkninger" skal fortolkes i bred

³ Anmodning, s. 1.

⁴ Ibid.

⁵ Anmodningen, s. 1.

⁶ Anmodning, s. 1. Som nævnt i Databeskyttelsesrådets arbejdsprogram for 2024-2025, der blev vedtaget den 8. oktober 2024, og som er tilgængeligt på https://www.edpb.europa.eu/system/files/2024-10/edpb_work_programme_2024-2025_en.pdf, har Databeskyttelsesrådet til hensigt bl.a. at udstede retningslinjer for anonymisering, pseudonymisering og dataskrabning i forbindelse med generativ kunstig intelligens.

forstand og derfor ikke blot er begrænset til retsvirkninger⁷. Da flere og flere AI-modeller trænes og anvendes af et stigende antal organisationer i EØS, påvirker de et stort antal registrerede i hele EØS, hvoraf nogle allerede har udtrykt bekymring til de kompetente tilsynsmyndighed⁸. Databeskyttelsesrådet mener derfor, at det spørgsmål, som den anmodende tilsynsmyndighed har rejst, også opfylder denne betingelse.

10. Anmodningen indeholder et skriftligt ræsonnement vedrørende baggrunden og begrundelserne for at stille spørgsmålene til Databeskyttelsesrådet, herunder de relevante juridiske rammer. Databeskyttelsesrådet mener derfor, at anmodningen er begrundet i overensstemmelse med artikel 10, stk. 3, i Databeskyttelsesrådets forretningsorden.
11. I henhold til artikel 64, stk. 3, i GDPR⁹ afgiver Databeskyttelsesrådet ikke en udtalelse, hvis det allerede har afgivet en udtalelse om samme spørgsmål. Databeskyttelsesrådet har ikke afgivet udtalelse om samme emne og har endnu ikke besvaret de spørgsmål, der er opstået som følge af anmodningen.
12. Af disse grunde er Databeskyttelsesrådet af den opfattelse, at anmodningen kan antages til behandling, og at de spørgsmål, der opstår i forbindelse med den, bør analyseres i denne udtalelse (herefter kaldet "**udtalelsen**"), som er vedtaget i henhold til artikel 64, stk. 2, i GDPR.

2 Anvendelsesområde og centrale begreber

2.1 Udtalelsens omfang

13. Databeskyttelsesrådet er enig med den anmodende tilsynsmyndighed i, at udviklingen og idriftsættelsen af AI-modeller ud fra et databeskyttelsesperspektiv rejser grundlæggende databeskyttelsesspørgsmål. Spørgsmålene vedrører navnlig: i) hvornår og hvordan kan en AI-model betragtes som "anonym" (spørgsmål 1 i anmodningen), ii) hvordan kan dataansvarlige påvise, at legitim interesse er et passende retsgrundlag i udviklingsfasen (spørgsmål 2 i anmodningen) og idriftsættelsesfasen (spørgsmål 3 i anmodningen), og iii) hvorvidt har ulovlig behandling af personoplysninger i udviklingsfasen konsekvenser for lovligheden af den efterfølgende behandling eller anvendelse af AI-modellen (spørgsmål 4 i anmodningen).
14. Databeskyttelsesrådet minder om, at tilsynsmyndighederne er ansvarlige for at overvåge anvendelsen af GDPR og bør bidrage til en konsekvent anvendelse af den i hele Unionen¹⁰. Det ligger derfor inden for tilsynsmyndighedernes kompetence at undersøge specifikke AI-modeller og i den forbindelse foretage vurderinger fra sag til sag.
15. Denne udtalelse udgør en ramme for kompetente tilsynsmyndigheder til at vurdere specifikke tilfælde, hvor (nogle af) de spørgsmål, der rejses i anmodningen, vil opstå. Det er ikke meningen, at denne udtalelse skal være udtømmende, men den skal snarere give nogle generelle overvejelser om fortolkningen af de relevante bestemmelser, som de kompetente tilsynsmyndigheder bør tage størst muligt hensyn til, når de udøver deres undersøgelsesbeføjelser. Selv om denne udtalelse er rettet til kompetente tilsynsmyndigheder og vedrører deres aktiviteter og beføjelser, berører den ikke dataansvarliges og databehandlers forpligtelser i henhold til GDPR. I henhold til

⁷ Databeskyttelsesrådet, internt dokument 3/2019 om intern vejledning om artikel 64, stk. 2, i GDPR, vedtaget den 8. oktober 2019, afsnit 15, tilgængeligt på https://www.edpb.europa.eu/system/files/2022-07/internaledpb_document_201903_art64.2_en.pdf.

⁸ Anmodning, s. 1-2.

⁹ Artikel 64, stk. 3, i GDPR og artikel 10, stk. 4, i Databeskyttelsesrådets forretningsorden.

¹⁰ Artikel 51, stk. 1, i GDPR og artikel 51, stk. 2, i GDPR.

ansvarlighedsprincippet, der er fastsat i artikel 5, stk. 2, i GDPR, er dataansvarlige navnlig ansvarlige for og skal kunne påvise overholdelse af alle principper vedrørende behandling af personoplysninger.

16. I nogle tilfælde gives der eksempler i udtalelsen, men i betragtning af bredden af spørgsmålene i anmodningen samt de forskellige typer af AI-modeller, der er omfattet heraf, er det ikke muligt at behandle alle potentielle scenarier i denne udtalelse. De teknologier, der er forbundet med AI-modeller, er i hastig udvikling, og Databeskyttelsesrådets overvejelser i denne udtalelse bør derfor fortolkes i lyset heraf.
17. **Denne udtalelse analyserer ikke nedenstående bestemmelser, som dog stadig kan spille en vigtig rolle ved vurderingen af de databeskyttelseskrav, der gælder for AI-modeller:**

- **Behandling af særlige kategorier af oplysninger:** Databeskyttelsesrådet minder om forbuddet i artikel 9, stk. 1, i GDPR vedrørende behandling af særlige kategorier af oplysninger og de begrænsede undtagelser i artikel 9, stk. 2, i GDPR¹¹. I denne henseende præciserede Den Europæiske Unions Domstol (herefter kaldet "**EU-Domstolen**") yderligere, at "*såfremt en helhed af oplysninger, der både omfatter følsomme oplysninger og ikke-følsomme oplysninger [...] er indsamlet i blokke, uden at oplysningerne kan adskilles fra hinanden under indsamlingen, skal behandlingen af denne helhed af oplysninger anses for at være forbudt som omhandlet i databeskyttelsesforordningens artikel 9, stk. 1, eftersom den i hvert fald indebærer én følsom oplysning, og ingen af undtagelserne i forordningens artikel 9, stk. 2, finder anvendelse*"¹². Domstolen understregede også, at ¹³ "*med henblik på anvendelsen af undtagelsen i databeskyttelsesforordningens artikel 9, stk. 2, litra e), skal det efterprøves, om den registrerede udtrykkeligt og ved en klar bekræftelse har haft til hensigt at offentliggøre de omhandlede personoplysninger*". Disse overvejelser bør tages i betragtning, når behandling af personoplysninger i forbindelse med AI-modeller involverer særlige kategorier af data.
- **Automatiseret beslutningstagning, herunder profilering:** De behandlingsaktiviteter, der udføres i forbindelse med AI-modeller, kan falde ind under anvendelsesområdet for artikel 22 i GDPR, som pålægger dataansvarlige yderligere forpligtelser og giver de registrerede yderligere garantier. Databeskyttelsesrådet minder i denne forbindelse om sine retningslinjer for automatiseret individuel beslutningstagning og profilering i medfør af forordning 2016/679¹⁴.
- **Forenelighed af formål:** Artikel 6, stk. 4, i GDPR fastsætter for visse retsgrundlag kriterier, som en dataansvarlig skal tage hensyn til for at fastslå, om behandling til et andet formål er forenelig med det formål, hvortil personoplysningerne oprindeligt blev indsamlet. Denne bestemmelse kan være relevant i forbindelse med udviklingen og idriftsættelsen af AI-modeller, og dens anvendelighed bør vurderes af tilsynsmyndigheder.

¹¹ Se også Databeskyttelsesrådets rapport om det arbejde, der er udført af ChatGPT-taskeforcen, der blev vedtaget den 23. maj 2024, punkt 18: "Med hensyn til behandlingen af særlige kategorier af personoplysninger skal en af undtagelserne i artikel 9, stk. 2, desuden finde anvendelse, for at behandlingen er lovlig. *I princippet kan en af disse undtagelser være artikel 9, stk. 2, litra e), i GDPR. Alene det faktum at personoplysninger er offentligt tilgængelige, indebærer imidlertid ikke, at "sådanne oplysninger tydeligvis er offentliggjort af den registrerede" [...]*".

¹² EU-Domstolens dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 89.

¹³ EU-Domstolens dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 77.

¹⁴ Artikel 29-Gruppens retningslinjer for automatiseret individuel beslutningstagning og profilering i medfør af forordning 2016/679, senest revideret og vedtaget den 6. februar 2018 og godkendt af Databeskyttelsesrådet den 25. maj 2018. Se også EU-Domstolens dom af 7. december 2023, sag C-634/21, *SCHUFA Holding m.fl.* (ECLI:EU:C:2023:957).

- **Konsekvensanalyser vedrørende databeskyttelse** (artikel 35 i GDPR): Konsekvensanalyser vedrørende databeskyttelse er et vigtigt element med hensyn til ansvarlighedsprincippet, hvis behandlingen i forbindelse med AI-modeller må antages at kunne medføre en høj risiko for fysiske personers rettigheder og friheder¹⁵.
- **Princippet om databeskyttelse gennem design** (artikel 25, stk. 1, i GDPR): Databeskyttelse gennem design er en væsentlig sikkerhedsforanstaltning, der skal vurderes af tilsynsmyndighederne i forbindelse med udviklingen og idriftsættelsen af en AI-model.

2.2 Centrale begreber

18. Som en indledende bemærkning ønsker Databeskyttelsesrådet at præcisere den terminologi og de begreber, der anvendes i denne udtalelse, og kun med henblik på denne udtalelse:
- **"Førstepartsdata"** henviser til personoplysninger, som den dataansvarlige har indsamlet fra de registrerede.
 - **"Tredjeparts-data"** henviser til personoplysninger, som dataansvarlige ikke har indhentet fra de registrerede, men indsamlet eller modtaget fra en tredjepart, f.eks. fra en dataformidler eller indsamlet via webskrabning.
 - **"Webskrabning"** er en almindeligt anvendt teknik til indsamling af oplysninger fra offentligt tilgængelige onlinekilder. Oplysninger, der er hentet fra f.eks. tjenester som nyhedsformidling, sociale medier, forumdiskussioner og personlige hjemmesider, kan indeholde personoplysninger.
 - Anmodningen vedrører **"AI-modellers livscyklus"** samt forskellige faser vedrørende bl.a. "design", "udvikling", "træning", "opdatering", "finjustering", "idriftsættelse" eller "efterfølgende træning" af AI-modeller. Databeskyttelsesrådet anerkender, at sådanne faser afhængigt af omstændighederne kan finde sted i forbindelse med udvikling og idriftsættelse af AI-modeller og kan omfatte behandling af personoplysninger til forskellige behandlingsformål. Med henblik på denne udtalelse finder Databeskyttelsesrådet det ikke desto mindre vigtigt at strømline kategoriseringen af de faser, der sandsynligvis vil finde sted. Med henblik på denne udtalelse henviser Databeskyttelsesrådet derfor til **"udviklingsfasen"** og til **"idriftsættelsesfasen"**. Udviklingen af en AI-model omfatter alle stadier før en eventuel idriftsættelse af AI-modellen og omfatter bl.a. kodeudvikling, indsamling og forbehandling af de personoplysninger, der skal anvendes til at træne modellen, samt selve træningen af modellen. Idriftsættelsen af en AI-model omfatter alle faser i forbindelse med brugen af en AI-model og kan omfatte alle aktiviteter, der gennemføres efter udviklingsfasen. Databeskyttelsesrådet er opmærksomt på de mange forskellige brugsscenarier og deres potentielle konsekvenser for behandlingen af personoplysninger, derfor bør tilsynsmyndighederne overveje, om bemærkningerne i denne udtalelse er relevante for den behandling, de vurderer.
 - Databeskyttelsesrådet fremhæver også, at begrebet **"træning"**, når det er nødvendigt, henviser til den del af udviklingsfasen, hvor AI-modeller lærer af data for at udføre de tiltænkte opgave (som forklaret i næste afsnit i denne udtalelse).

¹⁵ Artikel 29-Gruppen: Retningslinjer for konsekvensanalyse vedrørende databeskyttelse (DPIA) og bestemmelse af, om behandlingen "sandsynligvis indebærer en høj risiko" i henhold til forordning 2016/679, revideret og vedtaget den 4. oktober 2017 og godkendt af Databeskyttelsesrådet den 25. maj 2018.

- Begrebet og anvendelsesområdet for **AI-modeller**, som det forstås af Databeskyttelsesrådet med henblik på denne udtalelse, er yderligere præciseret i det følgende særlige afsnit.

2.3 AI-modeller inden for rammerne af udtalelsen

19. I EU's forordning om kunstig intelligens ("**AI-forordningen**")¹⁶ defineres et "AI-system" som "*et maskinbaseret system, som er udformet med henblik på at fungere med en varierende grad af autonomi, og som efter idriftsættelsen kan udvise en tilpasningsevne, og som til eksplicitte eller implicitte mål af det input, det modtager, udleder, hvordan det kan generere output såsom forudsigelser, indhold, anbefalinger eller beslutninger, som kan påvirke fysiske eller virtuelle miljøer*"¹⁷. I betragtning 12 i retsaken om kunstig intelligens gives en mere uddybende forklaring af begrebet "AI-system". I overensstemmelse hermed er en central egenskab ved AI-systemer deres evne til at udlede. De teknikker, der muliggør inferens ved opbygningen af et AI-system, omfatter maskinlæringstilgange og logik- og videnbaserede tilgange.
20. "AI-modeller" er på den anden side kun indirekte defineret i AI-forordningen: "*Selv om AI-modeller er vigtige komponenter i AI-systemer, udgør de ikke i sig selv AI-systemer. AI-modeller kræver, at der tilføjes yderligere komponenter, f.eks. en brugergrænseflade, for at blive til AI-systemer. AI-modeller er typisk integreret i og udgør en del af AI-systemer*"¹⁸.
21. Databeskyttelsesrådet forstår, at den definition af en AI-model, der foreslås i anmodningen, er snævrere end definitionen i AI-forordningen, da den henviser til "AI-model" som "*omfattende det produkt, der er resultatet af de træningsmekanismer, der anvendes på et sæt træningsdata i forbindelse med kunstig intelligens, maskinlæring, dyb læring eller i andre relaterede behandlingssammenhænge*", og præciserer endvidere, at "*udtrykket finder anvendelse på AI-modeller, der er beregnet til at gennemgå yderligere træning, finjustering og/eller udvikling, samt AI-modeller, der ikke er.*"¹⁹
22. Databeskyttelsesrådet vedtog derfor denne udtalelse ud fra den antagelse, at et AI-system baserer sig på en AI-model for at opnå det ønskede mål ved at integrere modellen i en større ramme (f.eks. kan et AI-system til kundeservice bruge en AI-model, der er trænet med historiske samtaledata, til at give svar på brugernes spørgsmål).
23. Desuden er AI-modeller (eller herefter kaldet "**modellerne**"), der er relevante for denne udtalelse, dem, der er udviklet gennem en træningsproces. En sådan træningsproces er en del af udviklingsfasen, hvor modellerne lærer af data for at udføre den specifikke opgave. Træningsprocessen kræver derfor et datasæt, som modellen skal bruge til at identificere og "lære" mønstre. I disse tilfælde vil modellen bruge forskellige teknikker til at opbygge en repræsentation af den viden, der er hentet fra træningsdatasættet. Dette er navnlig tilfældet med maskinlæring.
24. I praksis er enhver AI-model en algoritme, hvis funktion bestemmes af et sæt elementer. Dybe læringsmodeller udgøres f.eks. ofte af neurala netværk med flere lag, der er forbundet med hinanden gennem samarbejdende neuroner (noder) med en eller flere vægtede "indgange". Disse vægte ændres gennem læring, hvilket sker gennem en kobling fra output fra en neuron til input fra en anden neuron.

¹⁶Europa-Parlamentets og Rådets forordning (EU) 2024/1689 af 13. juni 2024 om harmoniserede regler for kunstig intelligens og om ændring af forordning (EF) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 og (EU) 2019/2144 samt direktiv 2014/90/EU, (EU) 2016/797 og (EU) 2020/1828 (AI-forordningen).

¹⁷ Se artikel 3, stk. 1, i AI-forordningen.

¹⁸ Betragtning 97 i AI-forordningen.

¹⁹ Anmodningen, s. 3.

En enkel dyb læringsmodel har følgende kendetegn: i) type og størrelse af hvert lag, ii) den vægt, som tildeles hvert output (nogle gange kaldet "parametre"), iii) aktiveringsfunktionerne²⁰ mellem lagene og eventuelt iv) andre processer, der kan finde sted mellem lagene. Når man f.eks. træner en enkel dyb læringsmodel til billedklassifikation, vil input ("**billedpixels**") blive forbundet med output, og vægtene kan justeres, så der produceres det rigtige output det meste af tiden.

25. Andre eksempler på dybe læringsmodeller er LLM-modeller og generativ kunstig intelligens, som anvendes til f.eks. at skabe nye data og indhold, der ligner menneskeskabt indhold.
26. **På grundlag af ovenstående betragtninger og i overensstemmelse med anmodningen omfatter denne udtalelses anvendelsesområde kun den undergruppe af AI-modeller, der er resultatet af træningen af sådanne modeller med personoplysninger.**

3 Anmodningens berettigelse

3.1 AI-modellers egenskaber i forhold til definitionen af personoplysninger

27. I artikel 4, stk. 1, i GDPR defineres personoplysninger som "*enhver form for information om en identificeret eller identificerbar fysisk person*" (dvs. den registrerede). Endvidere fastsættes det i betragtning 26 i GDPR, at databeskyttelsesprincipperne ikke bør gælde for anonyme oplysninger, dvs. oplysninger, der ikke vedrører en identificeret eller identificerbar fysisk person, under hensyntagen til "*alle de midler, som med rimelighed kan tænkes bragt i anvendelse*" af den dataansvarlige eller en anden person. Dette omfatter: i) oplysninger, der aldrig har været knyttet til en identificeret eller identificerbar person, og ii) personoplysninger, der er gjort anonyme på en sådan måde, at den registrerede ikke længere kan identificeres.
28. Spørgsmål 1²¹ i anmodningen kan således besvares ved at analysere, om en AI-model, der er resultatet af træning, som indebærer behandling af personoplysninger, i alle tilfælde bør betragtes som anonym. Med udgangspunktet i spørgsmålets formulering vil Databeskyttelsesrådet i dette afsnit henvise til processen med at "træne" en AI-model.
29. Først og fremmest ønsker Databeskyttelsesrådet at fremsætte følgende generelle betragtninger. AI-modeller, uanset om de er trænet med personoplysninger eller ej, er normalt designet til at komme med forudsigelser eller drage konklusioner, dvs. de er designet til at udlede. Desuden er AI-modeller, der er trænet med personoplysninger, ofte designet til at drage slutninger om enkeltpersoner, der er forskellige fra dem, hvis personoplysninger blev anvendt til at træne AI-modellen. Nogle AI-modeller er dog specifikt designet til at levere personoplysninger om personer, hvis personoplysninger blev brugt til at træne modellen, eller på en eller anden måde at gøre sådanne oplysninger tilgængelige. I disse tilfælde vil sådanne AI-modeller i sagens natur (og typisk nødvendigvis) omfatte oplysninger om en identificeret eller identificerbar fysisk person, og de vil derfor involvere behandling af personoplysninger. Derfor kan disse typer AI-modeller ikke betragtes som anonyme. Dette ville f.eks. være tilfældet med i) en generativ model, der finjusteres ved hjælp af en persons stemmeoptagelse til at efterligne vedkommendes stemme, eller ii) enhver model, der er udformet til at svare med personoplysninger, som den trænes med, når der anmodes om oplysninger om en bestemt person.

²⁰ Det vi sige funktioner, der på baggrund af input og vægte beregner en neurons output, som siden overføres til det næste lag i det neurale netværk.

²¹ "Anses den endelige AI-model, som er blevet trænet ved hjælp af personoplysninger, i alle tilfælde for ikke at opfylde definitionen af personoplysninger (som fastsat i artikel 4, stk. 1, i GDPR)?"

30. På baggrund af ovenstående overvejelser fokuserer Databeskyttelsesrådet i besvarelsen af spørgsmål 1 i anmodningen på situationen med AI-modeller, der ikke er designet til at levere personoplysninger relateret til træningsdataene.
31. Databeskyttelsesrådet mener, at selv når en AI-model ikke bevidst er designet til at producere oplysninger om en identificeret eller identificerbar fysisk person fra træningsdataene, kan oplysninger fra træningsdatasættet, herunder personoplysninger, stadig være "indlemmet" i modellens parametre, dvs. repræsenteret gennem matematiske objekter. De kan afvige fra de oprindelige træningsdatapunkter, men kan stadig opbevare den oprindelige information om disse data, som i sidste ende kan udtrækkes eller på anden måde indhentes direkte eller indirekte fra modellen. Når information om identificerede eller identificerbare personer, hvis personoplysninger blev anvendt til at træne modellen, kan indhentes fra en AI-model med midler, som med rimelig sandsynlighed vil blive anvendt, kan det konkluderes, at en sådan model ikke er anonym.
32. I den forbindelse hedder det i anmodningen, at "*Eksisterende forskningspublikationer fremhæver visse potentielle sårbarheder i AI-modeller, som kan resultere i behandling af personoplysninger,²² samt den form for behandling af personoplysninger, der finder sted, når modeller anvendes til brug med andre data, enten via applikationsprogrammeringsgrænseflader ("API'er") eller "prompt"-grænseflader²³.*
33. På samme måde er forskning i udtræk af træningsoplysninger særlig dynamisk²⁴. Den viser, at det i nogle tilfælde er muligt at anvende midler, der med rimelig sandsynlighed vil udtrække personoplysninger fra visse AI-modeller, eller blot utilsigtet indhente personoplysninger gennem interaktioner med en AI-model (f.eks. som en del af et AI-system). Vedvarende forskning på dette område vil bidrage til en yderligere vurdering af de resterende risici for gengivelse²⁵ og udtræk af personoplysninger i alle givne tilfælde.
34. **På baggrund af ovenstående betragtninger mener Databeskyttelsesrådet, at AI-modeller, der er trænet med personoplysninger, ikke i alle tilfælde kan betragtes som anonyme. I stedet bør det vurderes fra sag til sag, om en AI-model er anonym, baseret på specifikke kriterier.**

3.2 De omstændigheder, hvorunder AI-modeller kan betragtes som anonyme, og dokumentationen herfor

²² F.eks. Membership Inference Attacks ([OWASP](#)) og Model Inversion Attacks ([OWASP](#) og [Veale et al](#), 2018).

²³ Anmodning, s. 1-2.

²⁴ Se i denne henseende f.eks.: i) Veale Michael, Binns Reuben and Edwards Lilian 2018 " *Algorithms that remember: model inversion attacks and data protection law*" Phil. Phil. Trans. R. Soc. A 376: 20180083, tilgængelig på; (ii) Brown H., Lee K., Mireshghallah F., Shokri R., and Tramèr F., *What Does it Mean for a Language Model to Preserve Privacy?*, 2022, ACM Digital Library, FAccT '22, June 20, 2022, Seoul, Republic of Korea, tilgængelig 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*, januar 2024, National Institute of Standards and Technology, tilgængelig på; (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 Jun 2021, tilgængelig 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, tilgængelig på; (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 Apr 2020, tilgængelig på <https://arxiv.org/pdf/1911.07135>.

²⁵ I et AI-system baseret på generativ AI indebærer "regurgiterende" en tilbagestrømning, hvor outputdata direkte relaterer til træningsdata.

35. Med hensyn til spørgsmål 1 i anmodningen²⁶ anmodes Databeskyttelsesrådet om at præcisere, under hvilke omstændigheder en AI-model, som er blevet trænet ved hjælp af personoplysninger, kan betragtes som anonym. Med hensyn til spørgsmål 1(i)(a) i anmodningen²⁷ anmodes Databeskyttelsesrådet om at præcisere, hvilke beviser og/eller hvilken dokumentation tilsynsmyndighederne skal tage i betragtning, når de vurderer, om en AI-model er anonym.

3.2.1 Generelle overvejelser vedrørende anonymisering i den foreliggende sammenhæng

36. Brugen af udtrykket "*enhver form for information*" i definitionen af "*personoplysninger*" i artikel 4, stk. 1, i GDPR afspejler målet om at give dette begreb et bredt anvendelsesområde, der omfatter alle former for oplysninger, forudsat at de "*vedrører*" den registrerede, som er identificeret eller kan identificeres direkte eller indirekte.
37. Oplysninger kan vedrøre en fysisk person, selv når de teknisk set er struktureret eller kodet (f.eks. i et udelukkende maskinlæsbart format, uanset om det er proprietært eller åbent) på en måde, der ikke gør forbindelsen til den fysiske person umiddelbart synlig. I sådanne tilfælde kan softwareapplikationer anvendes til let at identificere, genkende og udtrække specifikke data. Dette gælder især for AI-modeller, hvor parametrene repræsenterer statistiske forhold mellem træningsoplysningerne, og hvor det kan være muligt at udtrække nøjagtige eller fejlagtige (på grund af inferens) personoplysninger, enten direkte fra forbindelsen mellem de oplysninger, der indgår i modellen, eller ved at stille spørgsmål til modellen.
38. Da AI-modeller normalt ikke indeholder data, der kan isoleres eller forbindes direkte, men snarere parametre, der repræsenterer probalitiske relationer mellem de data, der indgår i modellen, kan det være muligt at udtrække²⁸ information fra modellen, f.eks. medlemsinferens, i realistiske scenarier.

²⁶ "Under hvilke omstændigheder kan dette ske?"

²⁷ "Hvordan kan de tiltag, der er taget for at sikre, at AI-modellen ikke behandler personoplysninger, i bekræftende fald dokumenteres?"

²⁸ (i) Carlini N., Chien S., Nasr M., Song S., Terzis A., Tramer F., *Membership Inference Attacks From First Principles*, arXiv: 2112.03570, tilgængelig 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*. Env. Adv.10, eadj9260(2024). DOI:10.1126/sciadv.adj9260 tilgængelig på <https://www.science.org/doi/10.1126/sciadv.adj9260>,

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

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

(v) Hu H., *Membership Inference Attacks and Defenses on Machine Learning Models Literature*, tilgængelig 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., og Lee K., *Scalable Extraction of Training Data from (Production) Language Models*, arXiv: 2311.17035 28. november 2023, tilgængelig 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. marts 2017, tilgængelig 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, tilgængelig 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, tilgængelig 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, tilgængelig på <https://arxiv.org/abs/2409.19798>,

For at en tilsynsmyndighed kan være enig med den dataansvarlige i, at en given AI-model kan betragtes som anonym, bør den derfor som minimum kontrollere, om den har modtaget tilstrækkelig dokumentation for, at der med rimelige midler: (i) ikke kan udtrækkes personoplysninger, der er relateret til træningsdataene,²⁹ ud af modellen, og (ii) ikke kan produceres output, når der stilles spørgsmål til modellen, som vedrører de registrerede, hvis personoplysninger blev brugt til at træne modellen.

39. Tilsynsmyndigheder bør tage tre elementer i betragtning ved vurderingen af, om disse betingelser er opfyldt.
40. For det første bør tilsynsmyndighederne tage hensyn til de elementer, der er afdækket i Artikel 29-Gruppens seneste udtalelser og/eller Databeskyttelsesrådets retningslinjer på området. Med hensyn til anonymisering på datoen for denne udtalelse bør tilsynsmyndighederne tage hensyn til elementerne i Artikel 29-Gruppens udtalelse 05/2014 om anonymiseringsteknikker ("**Artikel 29-Gruppens udtalelse 05/2014**"), hvori det hedder, at hvis det ikke er muligt at udpege, forbinde og udlede oplysninger fra det angiveligt anonyme datasæt, kan dataene anses for at være anonyme³⁰. Det hedder også, at "*når et forslag ikke opfylder et af kriterierne, skal der foretages en grundig evaluering af identifikationsrisikoen*"³¹. **I betragtning af ovennævnte sandsynlighed for udledning og inferens mener Databeskyttelsesrådet, at AI-modeller med stor sandsynlighed vil kræve en så grundig evaluering af risikoen for identifikation.**
41. For det andet bør denne vurdering foretages under hensyntagen til "*alle de midler, som med rimelighed kan forventes at blive anvendt*" af den dataansvarlige eller en anden person til at identificere enkeltpersoner³², og fastlæggelsen af disse midler bør baseres på objektive faktorer som forklaret i betragtning 26 i GDPR. Disse kan omfatte:
- a. egenskaberne ved selve træningsdataene, AI-modellen og træningsprocessen³³
 - b. den kontekst, hvori AI-modellen frigives og/eller behandles³⁴
 - c. de øvrige oplysninger, der ville muliggøre identifikationen, og som kan være tilgængelige for den givne person
 - d. de omkostninger og den tid, som personen skal bruge på at indhente sådanne yderligere oplysninger (hvis de ikke allerede er tilgængelige for dem)³⁵, og
 - e. den teknologi, der er til rådighed på behandlingstidspunktet, samt den teknologiske udvikling³⁶.

(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. nov 2023, tilgængelig på: <https://arxiv.org/abs/2308.15727>.

²⁹ Udtræk omfatter navnlig det tilfælde, hvor personoplysninger udledes af selve AI-modellen, med begrænset eller ingen brug af søgegrænsefladerne.

³⁰ Artikel 29-Gruppens udtalelse 05/2014, side 24.

³¹ Artikel 29-Gruppens udtalelse 05/2014, side 24.

³² EU-Domstolens dom af 19. oktober 2016, sag C-582/14, *Breyer mod Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), præmis 43.

³³ Dette omfatter karakteristika såsom den unikke karakter af dataene i træningssættet, graden af nøjagtighed af oplysningerne, aggregering, randomisering og navnlig, hvordan disse påvirker sårbarheden med hensyn til identifikation.

³⁴ Dette omfatter kontekstuelle elementer såsom begrænsning af adgangen til visse personer og retlige garantier.

³⁵ EU-Domstolens dom af 7. marts 2024, sag C-479/22 P, *OC mod Europa-Kommissionen* (ECLI:EU:C:2024:215), præmis 50.

³⁶ EU-Domstolens dom af 7. marts 2024, sag C-479/22 P, *OC mod Europa-Kommissionen* (ECLI:EU:C:2024:215), præmis 50.

42. For det tredje bør tilsynsmyndigheder overveje, om dataansvarlige har vurderet risikoen for, om der er mulighed for identifikation via den dataansvarlige og forskellige typer af "*andre personer*", herunder utilsigtede tredjeparter, der tilgår AI-modellen, også under hensyntagen til, om de med rimelighed kan anses for at være i stand til at få adgang til eller behandle de pågældende oplysninger.
43. **Kort sagt mener Databeskyttelsesrådet, at for at en AI-model kan betragtes som anonym ved hjælp af rimelige midler, bør både i) sandsynligheden for direkte (herunder probabilistisk) at kunne udtrække personoplysninger vedrørende personer, hvis personoplysninger blev anvendt til at træne modellen, og ii) sandsynligheden for forsætligt eller uforsætligt at indhente sådanne personoplysninger fra forespørgsler være ubetydelig³⁷ for alle registrerede. Tilsynsmyndighederne bør automatisk tage hensyn til, at AI-modeller sandsynligvis vil kræve en grundig evaluering af sandsynligheden for identifikation for at nå frem til en konklusion om deres mulige anonyme karakter. Denne sandsynlighed bør vurderes under hensyntagen til "*alle de midler, der med rimelighed kan forventes at blive anvendt*" af den dataansvarlige eller en anden person, og bør også tage hensyn til utilsigtet (gen)anvendelse eller offentliggørelse af modellen.**

3.2.2 Elementer til evaluering af den resterende sandsynlighed for identifikation

44. Selv om der kan træffes foranstaltninger både i udviklings- og idriftsættelsesfasen for at mindske sandsynligheden for at indhente personoplysninger fra en AI-model, bør evalueringen af anonymiteten af en AI-model også omfatte direkte adgang til modellen.
45. Derudover bør tilsynsmyndighederne fra sag til sag vurdere, om de foranstaltninger, som den dataansvarlige har implementeret for at sikre og bevise, at en AI-model er anonym, er hensigtsmæssige og effektive.
46. Konklusionen efter tilsynsmyndighedens vurdering kan navnlig variere, alt efter om der er tale om en offentligt tilgængelig AI-model, som er tilgængelig for et ukendt antal personer med et ukendt antal metoder, hvormed man kan søge at udtrække personoplysninger, eller en intern AI-model, der kun er tilgængelig for medarbejdere. Selv om tilsynsmyndighederne i begge tilfælde bør kontrollere, at de dataansvarlige har opfyldt deres ansvarlighedsforpligtelse i henhold til artikel 5, stk. 2, og artikel 24 i GDPR, kan "*midler, som med rimelighed kan forventes at blive anvendt*" af andre personer, have en indvirkning på omfanget og arten af de mulige scenarier, der skal tages i betragtning. Derfor kan tilsynsmyndighederne, afhængigt af konteksten for udvikling og idriftsættelse af modellen, overveje forskellige niveauer af testning og modstandsdygtighed over for angreb.
47. I den forbindelse opstiller Databeskyttelsesrådet nedenfor en ikke-præskriptiv og ikkeudtømmende liste over mulige elementer, som tilsynsmyndighederne kan tage i betragtning ved vurderingen af en dataansvarligs påstand om anonymitet. Andre tilgange kan være mulige, hvis de giver et tilsvarende beskyttelsesniveau, navnlig under hensyntagen til, hvordan teknikken udvikles.
48. Tilstedeværelsen eller fraværet af nedenstående elementer er ikke et afgørende kriterium for at vurdere en AI-models anonymitet.

³⁷ EU-Domstolens dom af 19. oktober 2016, sag C-582/14, *Breyer mod Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), præmis 46, og EU-Domstolens dom af 7. marts 2024, sag C-479/22 P, *OC mod Europa-Kommissionen* (ECLI:EU:C:2024:215), præmis 51.

3.2.2.1 Udformning af AI-modeller

49. Med hensyn til udformningen af AI-modeller bør tilsynsmyndighederne evaluere de tilgange, som de dataansvarlige anvender i udviklingsfasen. Der bør i denne forbindelse tages hensyn til anvendelsen og effektiviteten af fire vigtige områder (se nedenfor).

Udvælgelse af kilder

50. Det første evalueringsområde omfatter en undersøgelse af udvælgelsen af de kilder, der anvendes til at træne AI-modellen. Dette omfatter en evaluering foretaget af tilsynsmyndighederne af alle tiltag, der er iværksat for at undgå eller begrænse indsamlingen af personoplysninger, herunder bl.a. i) udvælgelseskriteriernes hensigtsmæssighed, ii) de valgte kilders relevans og tilstrækkelighed i forhold til det/de tilsigtede formål, og iii) hvorvidt uhensigtsmæssige kilder er blevet udelukket.

Dataforberedelse og -minimering

51. Det andet evalueringsområde drejer sig om forberedelsen af data til træningsfasen. Tilsynsmyndighederne bør navnlig undersøge: i) hvorvidt anvendelsen af anonyme og/eller personlige oplysninger, der har været genstand for pseudonymisering, er blevet overvejet, og ii) hvis det blev besluttet ikke at anvende sådanne foranstaltninger, begrundelsen for denne afgørelse under hensyntagen til det tilsigtede formål, iii) de dataminimeringsstrategier og -teknikker, der anvendes til at begrænse mængden af personoplysninger, der indgår i træningsprocessen, og iv) eventuelle datafiltreringsprocesser, der er gennemført forud for modeltræning med henblik på at fjerne irrelevante personoplysninger.

Metodiske valg vedrørende træningen

52. Det tredje område af evalueringen vedrører udvælgelsen af robuste metoder inden for udvikling af AI-modeller. Tilsynsmyndigheder bør vurdere metodiske valg, der væsentligt kan reducere eller fjerne identificerbarheden, herunder bl.a.: i) hvorvidt denne metode anvender reguleringsmetoder til at forbedre modelgeneraliseringen og reducere overtilpasning, og især ii) hvorvidt den dataansvarlige har gennemført hensigtsmæssige og effektive teknikker til beskyttelse af privatlivets fred (f.eks. differentiell beskyttelse af privatlivets fred).

Foranstaltninger vedrørende modellens output

53. Det sidste evalueringsområde handler om metoder eller foranstaltninger, der er tilføjet til AI-modellen, og som måske ikke direkte påvirker risikoen for direkte udtræk af personoplysninger fra modellen, hvis nogen får adgang til den, men som kan reducere sandsynligheden for at få adgang til personoplysninger relateret til træningsdata gennem forespørgsler.

3.2.2.2 Analyse af AI-modeller

54. For at tilsynsmyndighederne kan vurdere robustheden af den designede AI-model med hensyn til anonymisering, er det første skridt at sikre, at designet er blevet udviklet som planlagt og er underlagt en effektiv teknisk styring. Tilsynsmyndigheder bør evaluere, om dataansvarlige har foretaget dokumentbaserede revisioner (interne eller eksterne), der omfatter en evaluering af de valgte foranstaltninger og deres indvirkning for at begrænse sandsynligheden for identifikation. Dette kunne omfatte analyse af rapporter om kodegranskning samt en teoretisk analyse, der dokumenterer hensigtsmæssigheden af de foranstaltninger, der er valgt for at mindske sandsynligheden for genidentifikation af den pågældende model.

3.2.2.3 AI-modeltestning og modstand mod angreb

55. Endelig bør tilsynsmyndigheder tage hensyn til omfanget, hyppigheden, kvantiteten og kvaliteten af de test, som den dataansvarlige har foretaget på modellen. Tilsynsmyndigheder bør navnlig tage hensyn til, at vellykkede test, som omfatter almindeligt kendte, nye angrebsmetoder, kun kan være bevis for modstanden mod netop disse angreb. På datoen for denne udtalelse kan dette blandt andet

omfatte struktureret testning i forhold til: i) "attribut- og medlemsinferens", ii) eksfiltrering, iii) gengivelse af træningsdata, iv) modelinversion eller v) rekonstruktionsangreb.

3.2.2.4 Dokumentation

56. I henhold til artikel 5, 24, 25 og 30 i GDPR og, i tilfælde af en sandsynlig høj risiko for registreredes rettigheder og frihedsrettigheder, artikel 35 i GDPR, skal dataansvarlige dokumentere deres behandlingsaktiviteter på tilstrækkelig vis. Dette gælder også for enhver behandling, der omfatter træning af en AI-model, selv om formålet med behandlingen er anonymisering. Tilsynsmyndigheder bør tage sådan dokumentation i betragtning og enhver regelmæssig vurdering af de deraf følgende risici for den behandling, der udføres af dataansvarlige, da disse foranstaltninger er væsentlige for at vise, at personoplysninger ikke behandles.
57. **Databeskyttelsesrådet mener, at tilsynsmyndigheder bør tage hensyn til dokumentationen, når en påstand om anonymitet vedrørende en given AI-model skal evalueres. Databeskyttelsesrådet bemærker, at hvis en tilsynsmyndighed, efter at have vurderet påstanden om anonymitet på baggrund af dokumentationen, ikke kan bekræfte, at der er truffet effektive foranstaltninger for anonymisere AI-modellen, kan tilsynsmyndigheden antage, at den dataansvarlige ikke har opfyldt sine ansvarlighedsforpligtelser i henhold til artikel 5, stk. 2, i GDPR. Der bør derfor også tages hensyn til overholdelse af andre bestemmelser i GDPR.**
58. Ideelt set bør tilsynsmyndighederne kontrollere, om den dataansvarliges dokumentation omfatter:
- alle oplysninger om konsekvensanalyser vedrørende databeskyttelse, herunder eventuelle vurderinger og afgørelser, der fastslår, at en konsekvensanalyse vedrørende databeskyttelse ikke var nødvendig
 - eventuel rådgivning eller feedback fra databeskyttelsesrådgiveren (hvor en databeskyttelsesrådgiver blev – eller burde være blevet – udpeget)
 - oplysninger om de tekniske og organisatoriske foranstaltninger, der er truffet i forbindelse med udformningen af AI-modellen for at mindske sandsynligheden for identifikation, herunder den trusselsmodel og de risikovurderinger, som disse foranstaltninger er baseret på. Dette bør omfatte de specifikke foranstaltninger for hver kilde til træningsdatasæt, herunder relevante kilde-URL'er og beskrivelser af de foranstaltninger, der er truffet (eller allerede truffet af leverandører af datasæt fra tredjepart)
 - de tekniske og organisatoriske foranstaltninger, der er truffet i alle faser af modellens livscyklus, og som enten har bidraget til eller bekræftet, at der ikke findes personoplysninger i modellen
 - dokumentation, der viser AI-modellens teoretiske modstandsdygtighed over for genidentifikationsteknikker, samt de kontroller, der er designet til at begrænse eller vurdere resultatet og virkningen af hovedangreb (gengivelse, medlemsinferensangreb, eksfiltrering osv.). Dette kan især omfatte: i) forholdet mellem mængden af træningsdata og antallet af parametre i modellen, herunder analysen af dens indvirkning på modellen³⁸, ii) målinger af

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

sandsynligheden for genidentifikation baseret på det aktuelle tekniske niveau, iii) rapporter om, hvordan modellen er blevet testet (af hvem, hvornår, hvordan og i hvilket omfang), og iv) resultaterne af testene

- f. dokumentationen, der leveres til den eller de dataansvarlige, der anvender modellen, og/eller til de registrerede, navnlig dokumentationen vedrørende de foranstaltninger, der er truffet for at mindske sandsynligheden for identifikation, og vedrørende de mulige resterende risici.

3.3 Legitim interesse som passende retsgrundlag for behandling af personoplysninger i forbindelse med udvikling og idriftsættelse af AI-modeller

- 59. For at besvare spørgsmål 2 og 3 i anmodningen vil Databeskyttelsesrådet først fremsætte generelle bemærkninger om visse vigtige aspekter, som tilsynsmyndigheder, uanset retsgrundlaget for behandling, bør tage i betragtning, når de vurderer, hvordan dataansvarlige kan påvise overholdelse af GDPR i forbindelse med AI-modeller. Med udgangspunkt i retningslinjerne 1/2024 for behandling af personoplysninger baseret på artikel 6, stk. 1, litra f), i GDPR³⁹ vil Databeskyttelsesrådet herefter se på de tre trin, der kræves i vurderingen af legitim interesse i forbindelse med udvikling og idriftsættelse af AI-modeller.

3.3.1 Generelle bemærkninger

- 60. Databeskyttelsesrådet minder om, at GDPR ikke fastsætter noget hierarki mellem de forskellige retsgrundlag, der er fastsat i artikel 6, stk. 1, i GDPR⁴⁰.
- 61. Artikel 5 i GDPR fastlægger principperne for behandling af personoplysninger. Databeskyttelsesrådet fremhæver dem, der er væsentlige for denne udtalelse, og der som minimum bør overvejes af tilsynsmyndighederne, når de vurderer specifikke AI-modeller, samt de mest relevante krav fra andre bestemmelser i GDPR under hensyntagen til anvendelsesområdet for denne udtalelse.
- 62. **Ansvarlighedsprincippet** (artikel 5, stk. 2, i GDPR) — Dette princip fastsætter, at den dataansvarlige skal være ansvarlig for og være i stand til at påvise overholdelse af GDPR. I denne forbindelse bør roller og ansvarsområder for de parter, der behandler personoplysninger i forbindelse med udviklingen eller idriftsættelsen af en AI-model, vurderes, inden behandlingen finder sted, med henblik på fra starten at fastlægge de forpligtelser, der påhviler de dataansvarlige eller fælles dataansvarlige og eventuelle databehandlere.
- 63. **Lovligheds-, rimeligheds- og gennemsigtighedsprincipper** (artikel 5, stk. 1, litra a), i GDPR) — Ved vurderingen af lovligheden af behandlingen i forbindelse med AI-modeller i lyset af artikel 6, stk. 1, i GDPR finder Databeskyttelsesrådet det hensigtsmæssigt at skelne mellem de forskellige faser⁴¹ af behandlingen af personoplysninger. Rimelighedsprincippet, som er tæt forbundet med princippet om gennemsigtighed, kræver, at personoplysninger ikke behandles med urimelige metoder, ved svig eller på en måde, som er "*urimeligt skadelig, ulovligt diskriminerende, uventet eller vildledende for den registrerede*"⁴². I betragtning af kompleksiteten af de involverede teknologier bør oplysninger om

³⁹ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024.

⁴⁰ Ibidem., stk. 1.

⁴¹ Databeskyttelsesrådets rapport om det arbejde, der er udført af ChatGPT-taskforcen, vedtaget den 23. maj 2024, punkt 14.

⁴² Databeskyttelsesrådets rapport om det arbejde, der er udført af ChatGPT-taskforcen, vedtaget den 23. maj 2024, punkt 23, Databeskyttelsesrådets retningslinjer 4/2019 for databeskyttelse gennem design og

behandling af personoplysninger inden for AI-modeller derfor gives på en tilgængelig, forståelig og brugervenlig måde⁴³. Gennemsigtighed i forbindelse med behandling af personoplysninger omfatter navnlig overholdelse af oplysningsforpligtelserne i artikel 12-14 i GDPR⁴⁴, som i tilfælde af automatiske afgørelser, herunder profilering, også kræver meningsfuld information om den involverede logik samt betydningen og de forventede konsekvenser af behandlingen for den registrerede⁴⁵. I betragtning af at udviklingsfaserne i AI-modeller kan omfatte indsamling af store mængder data fra offentligt tilgængelige kilder (f.eks. via webskrabningsteknikker), er anvendelse af undtagelsen i artikel 14, stk. 5, litra b), i GDPR strengt begrænset til, når kravene i denne bestemmelse er opfyldt fuldt ud⁴⁶.

64. **Principperne om formålsbegrænsning og dataminimering** (artikel 5, stk. 1, litra b) og c), i GDPR) – I overensstemmelse med princippet om dataminimering kræver udviklingen og idriftsættelsen af AI-modeller, at personoplysninger skal være tilstrækkelige, relevante og begrænset til, hvad der er nødvendigt i forhold til det formål, der forfølges. Dette kan omfatte behandling af personoplysninger for at undgå risikoen for potentielle bias og fejl, når dette klart og specifikt er angivet i forbindelse med formålet, og personoplysningerne er nødvendige til dette formål (kan f.eks. ikke opnås effektivt ved behandling af andre oplysninger, herunder syntetiske eller anonymiserede data⁴⁷). Artikel 29-Gruppen har allerede understreget, at "*formålet med indsamlingen skal være klart og specifikt angivet [...]'*"⁴⁸. Når man vurderer, om det forfulgte formål er legitimt, specifikt og udtrykkeligt, og om behandlingen er i overensstemmelse med princippet om dataminimering, skal man først angive den pågældende behandlingsaktivitet. Navnlig kan de forskellige faser i udviklings- eller idriftsættelsesfasen udgøre de samme eller forskellige behandlingsaktiviteter og kan omfatte flere på hinanden følgende dataansvarlige eller fælles dataansvarlige. I nogle tilfælde er det muligt at bestemme det formål, der skal forfølges under idriftsættelsen af AI-modellen, på et tidligt udviklingsstadium. Selv om dette ikke er tilfældet, bør der allerede være en vis kontekst for idriftsættelsen, og derfor bør der tages højde for, hvordan denne kontekst påvirker formålet med udviklingen. Når tilsynsmyndighederne gennemgår formålet med behandlingen på et givet udviklingstrin, bør de forvente en vis detaljeringsgrad fra den eller de dataansvarlige og en redegørelse for, hvordan disse oplysninger bidrager til formålet med behandlingen. Dette kan f.eks. omfatte oplysninger om den type AI-model, der er udviklet, dens forventede funktioner og enhver anden relevant kontekstviden, der allerede er kendt på det tidspunkt. Baggrunden for idriftsættelsen kan f.eks. også omfatte, om en model udvikles med henblik på intern idriftsættelse, om den dataansvarlige har til hensigt at sælge eller distribuere modellen til tredjeparter efter dens udvikling, herunder om modellen primært idriftsættes med henblik på forskningsmæssige eller kommercielle formål.

databeskyttelse gennem standardindstillinger i henhold til artikel 25, version 2.0, vedtaget den 20. oktober 2020, punkt 69 og artikel 29-Gruppens retningslinjer for gennemsigtighed i henhold til forordning 2016/679, revideret og vedtaget den 11. april 2018 og godkendt af Databeskyttelsesrådet den 25. maj 2018, punkt 2.

⁴³ Artikel 29-Gruppens retningslinjer for gennemsigtighed i henhold til forordning 2016/679, revideret og vedtaget den 11. april 2018, godkendt af Databeskyttelsesrådet den 25. maj 2018, punkt 5.

⁴⁴ Se også betragtning 39 i GDPR, som fastsætter, at det "*bør være gennemsigtigt for de pågældende fysiske personer, at personoplysninger, der vedrører dem, indsamles, anvendes, tilgås eller på anden vis behandles, og i hvilket omfang personoplysningerne behandles eller vil blive behandlet [...]'*".

⁴⁵ Artikel 13, stk. 2, litra f), i GDPR og artikel 14, stk. 2, litra g), i GDPR.

⁴⁶ Databeskyttelsesrådets rapport om det arbejde, der er udført af ChatGPT-taskforcen, vedtaget den 23. maj 2024, punkt 27.

⁴⁷ Desuden indeholder artikel 10, stk. 5, i AI-forordningen specifikke regler for behandling af særlige kategorier af personoplysninger i forbindelse med højrisiko-AI-systemer med henblik på påvisning og korrektion af bias.

⁴⁸ Artikel 29-Gruppens udtalelse 03/2013 om formålsbegrænsning (WP203), s. 15-16.

65. **Registreredes rettigheder** (kapitel III i GDPR) — selvom tilsynsmyndighederne skal sikre, at alle registreredes rettigheder respekteres, når AI-modeller udvikles og anvendes af dataansvarlige, minder Databeskyttelsesrådet om, at når en dataansvarlig påberåber sig legitime interesser som retsgrundlag, finder retten til at gøre indsigelse i henhold til artikel 21 i GDPR anvendelse og bør sikres⁴⁹.

3.3.2 Overvejelser om de tre trin i vurderingen af den legitime interesse i forbindelse med udviklingen og idriftsættelsen af AI-modeller

66. For at afgøre, om en given behandling af personoplysninger kan være baseret på artikel 6, stk. 1, litra f), i GDPR, bør tilsynsmyndighederne kontrollere, at de dataansvarlige omhyggeligt har vurderet og dokumenteret, om de tre følgende kumulative betingelser er opfyldt: i) den dataansvarliges eller en tredjeparts forfølgelse af en legitim interesse, ii) behandlingen er nødvendig for at forfølge den legitime interesse, og iii) den legitime interesse går ikke forud for de registreredes interesser eller grundlæggende rettigheder og frihedsrettigheder⁵⁰.

3.3.2.1 Første trin – Den dataansvarliges eller en tredjeparts forfølgelse af en legitim interesse

67. En interesse er den bredere drivkraft eller fordel, som en dataansvarlig eller tredjepart kan få ved at udføre en bestemt behandlingsaktivitet⁵¹. Selv om GDPR og EU-Domstolen anerkender flere interesser som legitime⁵², bør vurderingen af en given interesses legitimitet være resultatet af en analyse fra sag til sag.
68. Som Det Europæiske Databeskyttelsesråd påpegede i sine retningslinjer for legitim interesse⁵³, kan en interesse betragtes som legitim, hvis følgende tre kumulative kriterier er opfyldt:
- a. Interessen er lovlig⁵⁴

⁴⁹ Hvis den registrerede af grunde, der vedrører den pågældendes særlige situation, gør indsigelse mod behandling af personoplysninger om vedkommende, må den dataansvarlige i henhold til artikel 21 i GDPR ikke længere behandle personoplysningerne, medmindre den dataansvarlige påviser vægtige legitime grunde til behandlingen, der går forud for den registreredes interesser, rettigheder og frihedsrettigheder, eller behandlingen er nødvendig for, at retskrav kan fastlægges, gøres gældende eller forsvares. De to aspekter, som tilsynsmyndighederne skal tage hensyn til, er derfor, om den dataansvarlige er i stand til at påvise sådanne vægtige tungerevejende legitime grunde, og om retten til at gøre indsigelse kan udøves.

⁵⁰ EU-Domstolen, dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis. 106; EU-Domstolens dom af 11. december 2019, sag C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), præmis 40. Se også Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 12 og ff. Som nævnt i disse retningslinjer "bør denne vurdering foretages ved behandlingens start under inddragelse af databeskyttelsesrådgiveren (hvis en sådan er udpeget) og bør dokumenteres af den dataansvarlige i overensstemmelse med ansvarlighedsprincippet i artikel 5, stk. 2, i GDPR".

⁵¹ Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 14.

⁵² Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 16.

⁵³ Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 17.

⁵⁴ EU-Domstolen, dom af 4. oktober 2024, sag C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), præmis 49, hvor EU-Domstolen understregede, at en legitim interesse ikke kan være i strid med loven. I den forbindelse understreger Databeskyttelsesrådet, at der i givet fald bør tages hensyn til de lovgivningsmæssige rammer, når man vurderer lovligheden af en given interesse. Se f.eks. Artikel 26, stk. 3, og artikel 28 i Europa-Parlamentets og Rådets forordning (EU) 2022/2065 af 19. oktober 2022 om et indre marked for digitale tjenester og om ændring af direktiv 2000/31/EF (forordningen om digitale tjenester) om forbudt målrettet reklame for mindreårige, artikel 5, stk. 1 og 2, i AI-forordningen om forbudte AI-praksisser

- b. Interessen er klart og præcist formuleret, og
 - c. interessen er reel og ikke hypotetisk.
69. Med forbehold af de to andre trin, der kræves i forbindelse med vurderingen af den legitime interesse, kan følgende eksempler udgøre en legitim interesse i forbindelse med AI-modeller: i) udvikling af en samtalerobot til at hjælpe brugerne, ii) udvikling af et AI-system til at opdage svigagtigt indhold eller svigagtig adfærd og iii) forbedring af trusselsopsporing i et informationssystem.
- 3.3.2.2 Andet trin – Analyse af nødvendigheden af behandlingen for at forfølge den legitime interesse*
70. Det andet trin i vurderingen består i at afgøre, om behandlingen af personoplysninger er nødvendig af hensyn til den eller de legitime interesser, der forfølges⁵⁵ ("nødvendighedstest").
71. Det fremgår af betragtning 39 i GDPR, at *personoplysninger kun bør behandles, hvis formålet med behandlingen ikke med rimelighed kan opfyldes på anden måde*. Ifølge EU-Domstolen og tidligere retningslinjer fra Det Europæiske Databeskyttelsesråd bør betingelsen vedrørende nødvendigheden af behandlingen undersøges i lyset af de registreredes grundlæggende rettigheder og frihedsrettigheder og i sammenhæng med princippet om dataminimering, der er fastsat i artikel 5, stk. 1, litra c), i GDPR⁵⁶.
72. Den metode, som EU-Domstolen henviser til, tager hensyn til behandlingens kontekst samt virkningerne for den dataansvarlige og de registrerede. Vurderingen af nødvendighed indebærer

(manipulerende praksis og ud over den menneskelige bevidsthed) og behandling i strid med intellektuelle ejendomsrettigheder og bestemmelserne i direktiv (EU) 2019/790 om ophavsret og beslægtede rettigheder på det digitale indre marked.

⁵⁵ Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 28-30.

⁵⁶ EU-Domstolen, dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis. 108 og 109, der også henviser til EU-Domstolens dom af 11. december 2019, sag C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), præmis 48; EU-Domstolens dom af 9. november 2010, forenede sager C-92/09 og C-93/09, *Volker und Markus Schecke* (ECLI:EU:C:2010:662), præmis 85 og 86; EU-Domstolens dom af 22. juni 2021, sag C-439/19, *Latvijas Republikas Saeima* (ECLI:EU:C:2021:504), præmis 98, 109, 110 og 113. Se f.eks. også: Databeskyttelsesrådets retningslinjer 3/2019 for behandling af personoplysninger via videoenheder, version 2.0, vedtaget den 29. januar 2020, punkt 24-26 og 73, Databeskyttelsesrådets retningslinjer 2/2019 for behandling af personoplysninger i henhold til artikel 6, stk. 1, litra b), i GDPR i forbindelse med levering af onlinetjenester til registrerede, version 2.0, vedtaget den 8. oktober 2019, punkt 23-25 og Databeskyttelsesrådets udtalelse 11/2024 om brug af ansigtsgenkendelse til at strømline passagerstrømmen i lufthavne, version 1.1, vedtaget den 23. maj 2024, punkt 27.

derfor to elementer: i) hvorvidt behandlingsaktiviteten vil gøre det muligt at forfølge formålet, og⁵⁷ii) om der ikke findes en mindre indgribende måde at forfølge dette formål på⁵⁸.

73. Den tilsigtede mængde personoplysninger, der indgår i AI-modellen, skal f.eks. alt efter omstændighedern vurderes i lyset af mindre indgribende alternativer, der med rimelighed kan være tilgængelige for lige så effektivt at nå formålet med den legitime interesse, der forfølges. Hvis forfølgelsen af formålet også er mulig via en AI-model, der ikke indebærer behandling af personoplysninger, bør behandling af personoplysninger anses for ikke at være nødvendig. Dette er særlig relevant for udviklingen af AI-modeller. Ved vurderingen af, om betingelsen om nødvendighed er opfyldt, bør tilsynsmyndighederne være særligt opmærksomme på mængden af behandlede personoplysninger, og om det er forholdsmæssigt afpasset i forhold til at forfølge den legitime interesse, der gør sig gældende, også set i lyset af princippet om dataminimering.
74. Ved nødvendighedsvurderingen bør der også tages hensyn til den bredere sammenhæng med den påtænkte behandling af personoplysninger. Om der forefindes løsninger, der er mindre indgribende i de registreredes grundlæggende rettigheder og frihedsrettigheder, kan variere afhængigt af, om den dataansvarlige har en direkte relation til de registrerede (førstepartsoplysninger) eller ej (tredjepartsoplysninger). EU-Domstolen har fremlagt nogle overvejelser, der skal tages i betragtning, når man analyserer nødvendigheden af behandling af førstepartsoplysninger med henblik på legitime interesser, der forfølges (om end i forbindelse med videregivelse af sådanne oplysninger til tredjeparter)⁵⁹.
75. Gennemførelse af tekniske garantier for at beskytte personoplysninger kan også bidrage til at opfylde nødvendighedstesten. Dette kunne f.eks. omfatte gennemførelsesforanstaltninger som dem, der er anført i afsnit 3.2.2, på en sådan måde, at anonymisering ikke opnås, men stadig mindsker muligheden for at identificere registrerede. Databeskyttelsesrådet bemærker, at nogle af disse foranstaltninger, når de ikke er nødvendige for at overholde GDPR, kan udgøre yderligere garantier som yderligere analyseret i underafsnittet "afbødende foranstaltninger " i afsnit 3.3.2.3⁶⁰.

⁵⁷ Jf. Domstolens dom af 16. december 2008, sag C-524/06, *Heinz Huber mod Bundesrepublik Deutschland* (ECLI:EU:C:2008:724), præmis 66. Se også i samme sag forslag til afgørelse fra generaladvokat Poiares Maduro i sag C-524/06, *Heinz Huber mod Bundesrepublik Deutschland* (ECLI:EU:C:2008:194), præmis 16, som fastsætter: "Den rigtige test her er en effektivitetstest, og det er op til den nationale domstol at anvende den. Det spørgsmål, den skal stille, er, om der findes andre måder at behandle oplysninger på, som udlændingemyndighederne kan anvende for at håndhæve bestemmelserne om opholdsret. Hvis dette spørgsmål besvares bekræftende, skal den centrale lagring og behandling af oplysninger om unionsborgere anses for ulovlig. Det er ikke nødvendigt, at det alternative system udgør den mest effektive eller hensigtsmæssige løsning. Det er tilstrækkeligt, at systemet kan fungere. Sagt på en anden måde: Selv om det centrale register er mere effektivt, passende eller brugervenligt end alternativerne (såsom decentrale, lokale registre), skal lokale registre med andre ord klart foretrækkes, hvis de anvendes til at give oplysninger om unionsborgeres opholdsstatus."

⁵⁸ Jf. Domstolens dom af 27. september 2017, sag C-73/16, *Peter Puškár* (ECLI:EU:C:2017:725), præmis 113: "Det tilkommer således den forelæggende ret at efterprøve, om udarbejdelsen af den omtvistede liste og opførelsen af de pågældende personers navne på denne er egnet til at opfylde de mål, der forfølges med udarbejdelsen, og om der ikke findes mindre indgribende foranstaltninger til at opnå disse mål"; se også f.eks. forslag til afgørelse fra generaladvokat Rantos i sag C-252/21, *Meta mod Bundeskartellamt*, ECLI:EU:C:2022:704, præmis 61, hvori det anføres: "[...] Det er derfor nødvendigt, at der er en tæt forbindelse mellem behandlingen og den interesse, der forfølges, i mangel af alternativer med en bedre beskyttelse af personoplysninger, idet det ikke er tilstrækkeligt, at behandlingen alene skyldes, at den er nyttig for den dataansvarlige".

⁵⁹ EU-Domstolen, dom af 4. oktober 2024, sag C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), præmis 51-53.

⁶⁰ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 57.

3.3.2.3 Tredje trin - Afvejningstest

76. Det tredje trin i vurderingen af den legitime interesse er "**afvejningen**" (i dette dokument også kaldet "**afvejningstesten**")⁶¹. Dette trin består i at identificere og beskrive de forskellige modstridende rettigheder og interesser, der er på spil⁶², dvs. på den ene side de registreredes interesser, grundlæggende rettigheder og frihedsrettigheder og på den anden side den dataansvarliges eller en tredjeparts interesser. Sagens særlige omstændigheder bør derefter behandles for at påvise, at en legitim interesse er et passende retsgrundlag for de pågældende behandlingsaktiviteter⁶³.

De registreredes interesser, grundlæggende rettigheder og frihedsrettigheder

77. I henhold til artikel 6, stk. 1, litra f), i GDPR skal den dataansvarlige ved vurderingen af de forskellige komponenter i forbindelse med afvejningstesten tage hensyn til de registreredes interesser, grundlæggende rettigheder og frihedsrettigheder. De registreredes interesser er dem, der kan blive berørt af den pågældende behandling. I forbindelse med udviklingsfasen af en AI-model kan disse omfatte, men er ikke begrænset til, interessen for selvbestemmelse og bevarelse af kontrollen over ens egne personlige data (f.eks. de data, der er indsamlet til udvikling af modellen). I forbindelse med idriftsættelsen af en AI-model kan de registreredes interesser omfatte, men er ikke begrænset til, interesse i at bevare kontrollen over ens egne personoplysninger (f.eks. de data, der behandles, når modellen er taget i brug), økonomiske interesser (f.eks. hvis en AI-model anvendes af den registrerede til at generere indtægter eller anvendes af en person i forbindelse med vedkommendes erhverv), interesse på grund af personlige fordele (f.eks. hvis en AI-model anvendes til at forbedre adgangen til visse tjenester) eller socioøkonomiske interesser (f.eks. hvis en AI-model giver adgang til bedre sundhedspleje eller letter udøvelsen af en grundlæggende rettighed såsom adgang til uddannelse)⁶⁴.
78. Jo mere præcist en interesse defineres i lyset af det tilsigtede formål med behandlingen, desto lettere vil det være at få et overblik over de reelle fordele og risici, der skal tages i betragtning i afvejningstesten.
79. I forbindelse med de registreredes grundlæggende rettigheder og frihedsrettigheder kan udviklingen og idriftsættelsen af AI-modeller medføre alvorlige risici for de rettigheder, der er beskyttet af EU's charter om grundlæggende rettigheder ("**EU's charter**"), herunder, men ikke begrænset til, retten til privatlivet og familielivet (artikel 7 i EU's charter) og retten til beskyttelse af personoplysninger (artikel 8 i EU's charter). Disse risici kan opstå i udviklingsfasen, f.eks. hvis personoplysninger er indsamlet ved webskrabning mod de registreredes vilje eller uden deres viden. Disse risici kan også opstå i idriftsættelsesfasen, f.eks. hvis personoplysninger behandles af (eller som en del af) modellen på en måde, der er i strid med de registreredes rettigheder, eller hvis det er muligt at udlede, utilsigtet eller gennem angreb (f.eks. medlemsinferens, udtræk eller modelinversion), hvilke personoplysninger der er indeholdt i læringsdatabasen. Sådanne situationer udgør en risiko for at krænke retten til privatlivets fred hos de registrerede, hvis data kan dukke op i idriftsættelsesfasen af AI-systemet (f.eks. omdømmerisiko, identitetstyveri eller svindel, sikkerhedsrisiko afhængigt af dataenes art).

⁶¹ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 31-60.

⁶² Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 32.

⁶³ Jf. Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, præmis 32, der også henviser til Domstolens dom af 4. juli 2023, sag C-252/21, *Meta mod Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 110.

⁶⁴ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 38.

80. Afhængigt af den pågældende sag kan der også være risici for andre grundlæggende rettigheder. For eksempel kan omfattende og vilkårlig dataindsamling af AI-modeller i udviklingsfasen skabe en følelse af overvågning for de registrerede, især når man tænker på, hvor svært det er at forhindre, at offentlige data bliver skrabt. Dette kan få enkeltpersoner til at udøve selvzensur, og udgør en risiko for at undergrave deres ytringsfrihed (artikel 11 i EU's charter). I idriftsættelsesfasen er der også risici for ytringsfriheden, når AI-modeller bruges til at blokere offentliggørelse af indhold fra de registrerede. Derudover kan en AI-model, der anbefaler upassende indhold til sårbare personer, udgøre en risiko for deres mentale sundhed (artikel 3, stk. 1, i EU's charter). I andre tilfælde kan idriftsættelse af AI-modeller også have negative konsekvenser for den enkeltes ret til at arbejde (artikel 15 i EU's charter), f.eks. når jobansøgninger udvælges på forhånd ved hjælp af en AI-model. På samme måde kan en AI-model udgøre en risiko for retten til ikkeforskelsbehandling (artikel 21 i EU's charter), hvis den diskriminerer enkeltpersoner på baggrund af visse personlige karakteristika (f.eks. nationalitet eller køn). Desuden kan idriftsættelsen af AI-modeller også udgøre en risiko for den enkeltes sikkerhed og tryghed (f.eks. hvor AI-modellen anvendes med ondsindet hensigt) samt for deres fysiske og mentale integritet⁶⁵.
81. Idriftsættelse af AI-modeller kan også have en positiv indvirkning på visse grundlæggende rettigheder, f.eks. kan modellen understøtte en persons ret til mental integritet (artikel 3 i chartret), f.eks. når en AI-model anvendes til at identificere skadeligt indhold på nettet, eller modellen kan lette adgangen til visse væsentlige tjenester eller lette udøvelsen af grundlæggende rettigheder, såsom adgang til information (artikel 11 i EU's charter) eller adgang til uddannelse (artikel 14 i EU's charter).

Behandlingens indvirkning på registrerede

82. Den behandling af personoplysninger, der finder sted under udviklingen og idriftsættelsen af AI-modeller, kan påvirke registrerede på forskellige måder, hvilket kan være positivt eller negativt⁶⁶. Hvis en behandlingsaktivitet f.eks. medfører fordele for den registrerede, kan der tages hensyn til disse i afvejningstesten. Selv om sådanne fordele kan føre til, at en tilsynsmyndighed konkluderer, at den dataansvarliges eller en tredjeparts interesser ikke går forud for den registreredes interesser, grundlæggende rettigheder og frihedsrettigheder, kan det kun konkluderes efter en analyse af den enkelte konkrete sag, hvor der tages hensyn til alle relevante faktorer.
83. Behandlingens indvirkning på de registrerede kan påvirkes af i) karakteren af de data, der behandles af modellerne, ii) konteksten for behandlingen og iii) de yderligere konsekvenser, som behandlingen kan have⁶⁷.
84. Med hensyn til **karakteren af de data, der behandles**, bør det erindres, at – bortset fra særlige kategorier af personoplysninger og data vedrørende straffedomme og lovovertrædelser, der nyder yderligere beskyttelse i henhold til artikel 9 og 10 i GDPR – kan behandlingen af visse andre kategorier af personoplysninger have betydelige konsekvenser for de registrerede. I denne forbindelse bør behandlingen af visse typer personoplysninger, der afslører meget private oplysninger (f.eks. finansielle data eller lokaliseringsdata) i forbindelse med udviklingen og idriftsættelsen af en AI-model, anses for muligvis at have en alvorlig indvirkning på de registrerede. I idriftsættelsesfasen kan konsekvenserne af en sådan behandling for den registrerede f.eks. være af økonomisk art (f.eks.

⁶⁵ Retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 46.

⁶⁶ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 39.

⁶⁷ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 32.

diskrimination i forbindelse med ansættelse) og/eller skade den pågældendes omdømme (f.eks. ærekrænkelse).

85. I forbindelse med **behandlingen** er det først nødvendigt at udpege de elementer, der kan skabe risici for de registrerede (f.eks. den måde, hvorpå modellen blev udviklet, den måde, hvorpå modellen kan anvendes, og/eller hvorvidt de sikkerhedsforanstaltninger, der anvendes til at beskytte personoplysningerne, er hensigtsmæssige). Modellens art og de tilsigtede anvendelser i praksis spiller en vigtig rolle, når man skal afdække sådanne potentielle årsager.
86. Det er også nødvendigt at vurdere, hvor alvorlige disse risici er for den registrerede. Det kan bl.a. overvejes, hvordan personoplysningerne behandles (f.eks. hvis de kombineres med andre datasæt), hvad omfanget af behandlingen og mængden af behandlede personoplysninger er⁶⁸ (f.eks. den samlede mængde data, mængden af data pr. registreret, antallet af berørte registrerede)⁶⁹, den registreredes status (f.eks. børn eller andre sårbare registrerede) og deres forhold til den dataansvarlige (f.eks. hvis den registrerede er en kunde). For eksempel kan brugen af webskrabning i udviklingsfasen – i mangel af tilstrækkelige sikkerhedsforanstaltninger – føre til betydelige konsekvenser for enkeltpersoner på grund af den store mængde data, der indsamles, det store antal registrerede og den vilkårlige indsamling af personoplysninger.
87. De **yderligere konsekvenser**, som behandlingen kan have, bør også tages i betragtning, når man vurderer behandlingens indvirkning på den registrerede. De bør vurderes af tilsynsmyndighederne fra sag til sag under hensyntagen til de specifikke faktiske omstændigheder.
88. Sådanne konsekvenser kan omfatte (men er ikke begrænset til) risici for overtrædelse af den registreredes grundlæggende rettigheder, som beskrevet i det foregående underafsnit⁷⁰. Risiciene kan være af varierende sandsynlighed og alvor og opstå som følge af behandling af personoplysninger, der kan føre til fysisk, materiel eller immateriel skade, navnlig hvis behandlingen kan give anledning til forskelsbehandling⁷¹.
89. Hvis idriftsættelsen af en AI-model indebærer behandling af personoplysninger for både i) registrerede, hvis personoplysninger indgår i det datasæt, der anvendes i udviklingsfasen, og ii) registrerede, hvis personoplysninger behandles i idriftsættelsesfasen, bør tilsynsmyndigheder sondre og overveje de risici, der påvirker hver af disse kategorier af registreredes interesser, rettigheder og frihedsrettigheder, når de kontrollerer den afvejningstest, der er udført af den dataansvarlige.
90. **Endelig bør analysen af de mulige yderligere konsekvenser af behandlingen også tage højde for sandsynligheden for, at disse yderligere konsekvenser opstår.** Vurderingen af en sådan sandsynlighed bør foretages under hensyntagen til de tekniske og organisatoriske foranstaltninger, der er truffet, og sagens særlige omstændigheder. Tilsynsmyndigheder kan f.eks. overveje, om der er gennemført foranstaltninger for at undgå et potentielt misbrug af AI-modellen. For AI-modeller, der kan anvendes til forskellige formål, såsom generativ AI, kan dette omfatte kontroller, der så vidt muligt begrænser deres brug til skadelig praksis, for eksempel: fremstilling af deepfakes, chatbots, der bruges til

⁶⁸ Se Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 43.

⁶⁹ EU-Domstolen, dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 116.

⁷⁰ Se underafsnittet "De registreredes interesser, grundlæggende rettigheder og frihedsrettigheder" ovenfor.

⁷¹ Se afsnit 2.3 i Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024. Se også betragtning 75 i GDPR for yderligere eksempler.

misinformation, phishing og andre former for svindel og manipulerende AI/AI-agenter (især hvis de er menneskelignende eller giver vildledende oplysninger).

Rimelige forventninger hos de registrerede

91. Baseret på betragtning 47 i GDPR, "[I] alle tilfælde kræver tilstedeværelsen af en legitim interesse en nøje vurdering, herunder af, om en registreret på tidspunktet for og i forbindelse med indsamling af personoplysninger med rimelighed kan forvente, at behandling med dette formål kan finde sted. Den registreredes interesser og grundlæggende rettigheder kan navnlig gå forud for den dataansvarliges interesser, hvis personoplysninger behandles under omstændigheder, hvor den registrerede ikke med rimelighed kan forvente viderebehandling"⁷².
92. Rimelige forventninger spiller en central rolle i afvejningstesten, ikke mindst på grund af kompleksiteten af den teknologi, der anvendes i AI-modeller, og det faktum, at det kan være vanskeligt for de registrerede at forstå de mange forskellige potentielle anvendelser af en AI-model og den involverede databehandling⁷³. I denne henseende kan de oplysninger, der gives til den registrerede tages i betragtning for at vurdere, om den registrerede med rimelighed kan forvente behandling af deres personoplysninger. Selv om udeladelsen af oplysninger kan bidrage til, at den registrerede ikke forventer en bestemt behandling, er alene opfyldelsen af gennemsigtighedskravene i GDPR imidlertid ikke i sig selv tilstrækkelig til at antage, at den registrerede med rimelighed kan forvente en vis behandling⁷⁴. Og blot fordi den dataansvarliges politik for beskyttelse af privatlivets fred indeholder oplysninger om udviklingsfasen af en AI-model, betyder det ikke nødvendigvis, at den registrerede med rimelighed kan forvente, at det sker. Dette bør analyseres af tilsynsmyndighederne på baggrund af de specifikke omstændigheder i sagen og under hensyntagen til alle relevante faktorer.
93. Ved vurderingen af de registreredes rimelige forventninger i forbindelse med behandling, der finder sted i udviklingsfasen, er det vigtigt at henvise til de elementer, der er nævnt i Databeskyttelsesrådets retningslinjer for legitim interesse⁷⁵. Inden for emneområdet for denne udtalelse er det desuden vigtigt at overveje den bredere kontekst for behandlingen. Dette kan omfatte, men er ikke begrænset til, om personoplysningerne var offentligt tilgængelige eller ej, arten af forholdet mellem den registrerede og den dataansvarlige (og om der findes en forbindelse mellem de to), tjenestens art, den sammenhæng, hvori personoplysningerne blev indsamlet, den kilde, hvorfra oplysningerne blev indsamlet (f.eks. det websted eller den tjeneste, hvor personoplysningerne blev indsamlet, og de privatlivsindstillinger, de tilbyder), de potentielle yderligere anvendelser af modellen, og om de registrerede faktisk er klar over, at deres personoplysninger overhovedet er online.

⁷² Se også EU-Domstolens dom af 4. juli 2023, sag C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 112; EU-Domstolens dom af 11. december 2019, sag C-708/18, *Asociația de Proprietari bloc M5A-Scara A* (ECLI:EU:C:2019:1064), præmis 58; EU-Domstolens dom af 4. oktober 2024, sag C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), præmis 55.

⁷³ For eksempel fandt EU-Domstolen i dom af 4. juli 2023, sag C-252/21, *Meta mod Bundeskartellamt* (ECLI:EU:C:2023:537), præmis 123, at "produktforbedring" i princippet ikke kan udelukkes som en legitim interesse, men den fandt også, at det er "tvivlsomt, om [...] formålet om forbedring af varen med hensyn til den i hovedsagen omhandlede behandling af oplysninger - henset til behandlingens omfang og dens betydelige indvirkning på brugeren samt til den omstændighed, at brugeren ikke med rimelighed kan forvente, at Meta Platforms Ireland behandler disse oplysninger - kan gå forud for en sådan brugers interesser og grundlæggende rettigheder, navnlig når brugeren er et barn".

⁷⁴ Retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 53.

⁷⁵ Retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 50-54.

94. I modellens udviklingsfase kan de registreredes rimelige forventninger variere afhængigt af, om de data, der behandles for at udvikle modellen, offentliggøres af de registrerede eller ej. Desuden kan de rimelige forventninger også variere afhængigt af, om de direkte har givet dataene til den dataansvarlige (f.eks. i forbindelse med deres brug af tjenesten), eller om den dataansvarlige har fået dem fra en anden kilde (f.eks. via en tredjepart eller skrabning). I begge tilfælde bør de skridt, der tages for at informere de registrerede om behandlingsaktiviteterne, tages i betragtning ved vurderingen af de rimelige forventninger.
95. I driftsættelsesfasen af AI-modellen er det lige så vigtigt at tage hensyn til de registreredes rimelige forventninger i forbindelse med modellens specifikke kapaciteter. For eksempel kan det for AI-modeller, der kan tilpasse sig efter de input, den får, være relevant at overveje, om de registrerede var klar over, at de havde givet personoplysninger, så AI-modellen kunne tilpasse sine svar til deres behov, og så de kunne få skræddersyede tjenester. Desuden kan det også være relevant at overveje, om denne behandlingsaktivitet kun vil påvirke den tjeneste, der leveres til de registrerede (f.eks. personalisering af indhold for en specifik bruger), eller om den vil blive anvendt til at ændre den tjeneste, der leveres til alle kunder (f.eks. for at forbedre modellen generelt). Ligesom i udviklingsfasen kan det også være særligt relevant at overveje, om der er en direkte forbindelse mellem de registrerede og den dataansvarlige. En sådan direkte forbindelse kan f.eks. gøre det muligt for den dataansvarlige nemt at informere de registrerede om behandlingsaktiviteten og modellen, hvilket så kan påvirke de rimelige forventninger hos disse registrerede.

Afbødende foranstaltninger

96. Når de registreredes interesser, rettigheder og frihedsrettigheder synes at gå forud for de legitime interesser, der forfølges af den dataansvarlige eller en tredjepart, kan den dataansvarlige overveje at indføre afbødende foranstaltninger for at begrænse behandlingens indvirkning på de registrerede. Afbødende foranstaltninger er sikkerhedsforanstaltninger, der bør skræddersyes til sagens omstændigheder, og som afhænger af forskellige faktorer, herunder den tilsigtede brug af AI-modellen. Disse afbødende foranstaltninger vil have til formål at sikre, at den dataansvarliges eller tredjemands interesser ikke tilsidesættes, således at den dataansvarlige kan påberåbe sig dette retsgrundlag.
97. Som det fremgår af Databeskyttelsesrådets retningslinjer for legitim interesse, bør afhjælpende foranstaltninger ikke forveksles med de foranstaltninger, som den dataansvarlige alligevel er juridisk forpligtet til at træffe for at sikre overholdelse af GDPR, uanset om behandlingen er baseret på artikel 6, stk. 1, litra f), i GDPR⁷⁶. Dette er særlig vigtigt for foranstaltninger, der f.eks. kræver overholdelse af principperne i GDPR, såsom princippet om dataminimering.
98. Nedenstående liste over foranstaltninger er ikke udtømmende og ikke præskriptiv, og gennemførelsen af foranstaltningerne bør vurderes fra sag til sag. Selv om visse af nedenstående foranstaltninger, afhængigt af omstændighederne, kan være nødvendige for at opfylde specifikke forpligtelser i henhold til GDPR, kan der, når dette ikke er tilfældet, tages hensyn til dem som yderligere garantier. Visse af de foranstaltninger, der er nævnt nedenfor, omfatter desuden områder, der er genstand for en hastig og ny udvikling, og som tilsynsmyndighederne bør tage hensyn til, når de behandler en specifik sag.
99. **I forbindelse med udviklingsfasen for AI-modeller** kan der træffes flere foranstaltninger for at afbøde de risici, der er forbundet med behandlingen af både førsteparts- og tredjepartsdata (herunder for at afbøde risici i forbindelse med webskrabning). På baggrund af ovenstående giver

⁷⁶ Retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 57.

Databeskyttelsesrådet nogle eksempler på foranstaltninger, der kan implementeres for at afbøde de risici, der er konstateret i afvejningstesten, og som bør overvejes af tilsynsmyndighederne, når de vurderer specifikke AI-modeller i hvert enkelt tilfælde.

100. Tekniske foranstaltninger:

- a. Foranstaltningerne i afsnit 3.2.2 kan afbøde de pågældende risici, hvor disse foranstaltninger ikke resulterer i anonymisering af modellen og ikke er nødvendige for at overholde andre GDPR-forpligtelser eller nødvendighedstesten (andet trin i vurderingen af den legitime interesse).

101. Derudover kan andre relevante foranstaltninger omfatte:

- b. Pseudonymiseringsforanstaltninger: Dette kan f.eks. omfatte foranstaltninger til at forhindre enhver kombination af data baseret på individuelle identifikatorer. Disse foranstaltninger er muligvis ikke hensigtsmæssige, hvis tilsynsmyndigheden finder, at den dataansvarlige påviste et rimeligt behov for at indsamle forskellige data om en bestemt person med henblik på udviklingen af det pågældende AI-system eller den pågældende AI-model.
- c. Foranstaltninger til at maskere personoplysninger eller erstatte dem med falske personoplysninger i træningssættet (f.eks. udskiftning af navne og e-mailadresser med falske navne og falske e-mailadresser). Denne foranstaltning kan være særlig hensigtsmæssig, når selve indholdet i oplysningerne ikke er relevant for den samlede behandling (f.eks. i forbindelse med LLM-træning).

102. Foranstaltninger, der gør det lettere for enkeltpersoner at udøve deres rettigheder:

- a. at overholde en rimelig tidsperiode mellem indsamlingen af et træningsdatasæt og brugen heraf. Denne yderligere garanti kan gøre det muligt for registrerede at udøve deres rettigheder i denne periode, idet den rimelige frist vurderes i forhold til omstændighederne i den enkelte sag.
- b. at give en ubetinget "fravalgsmulighed" fra begyndelsen, f.eks. ved at give registrerede ret til at gøre indsigelse, inden behandlingen finder sted, med henblik på at give personer en øget kontrol over deres data, som rækker ud over betingelserne i artikel 21 i⁷⁷GDPR.
- c. at give de registrerede mulighed for at udøve deres ret til sletning, selv når de specifikke grunde, der er anført i artikel 17, stk. 1, i GDPR, ikke finder anvendelse⁷⁸.
- d. at give de registrerede mulighed for at anmelde påstande om, at personoplysninger gengives eller memoreres, og de omstændigheder og midler, gennem hvilke påstandene kan reproducere, hvilket gør det muligt for de dataansvarlige at reproducere og vurdere relevante aflæringsteknikker for at løse problemet.

103. Gennemsigtighedsforanstaltninger: I nogle tilfælde kan afbødende foranstaltninger omfatte foranstaltninger, der sikrer større gennemsigtighed med hensyn til udviklingen af AI-modellen. Visse foranstaltninger, ud over overholdelse af GDPR-forpligtelserne, kan bidrage til at overvinde informationsasymmetri og give de registrerede mulighed for at få en bedre forståelse af den behandling, der er involveret i udviklingsfasen:

⁷⁷ Ibid.

⁷⁸ Ibid.

- a. Offentliggørelse af offentlig og let tilgængelig kommunikation, der rækker ud over de oplysninger, der kræves i henhold til artikel 13 eller 14 i GDPR, f.eks. ved at give yderligere oplysninger om indsamlingskriterierne og alle anvendte datasæt under hensyntagen til den særlige beskyttelse af børn og sårbare personer.
 - b. Alternative former for information til de registrerede, f.eks.: mediekampagner med forskellige medier for at informere de registrerede, informationskampagne via e-mail, brug af grafisk visualisering, ofte stillede spørgsmål, gennemsigtighedsmærker og systematiserede modelkort, der kan bidrage til at strukturere præsentationen af oplysninger om AI-modeller, og årlige gennemsigtighedsrapporter på frivillig basis.
104. **Specifikke afbødende foranstaltninger i forbindelse med webskrabning:** I betragtning af at webskrabning, som nævnt ovenfor, medfører specifikke risici⁷⁹, kan der udpeges specifikke afbødende foranstaltninger i denne sammenhæng. Hvor det er relevant, kan tilsynsmyndighederne tage højde for disse i tillæg til de ovenfor nævnte afbødende foranstaltninger, når de undersøger dataansvarlige, der udfører webskrabning.
105. Specifikke foranstaltninger, der ikke er nødvendige i henhold til andet trin i vurderingen af legitim interesse, kan vise sig at være nyttige for at mindske risikoen i forbindelse med webskrabning. Disse kan omfatte **tekniske foranstaltninger**, såsom:
- a. Undgå dataindhold fra publikationer, der kan omfatte personoplysninger, der indebærer risici for bestemte personer eller grupper af personer (f.eks. personer, der kan blive udsat for misbrug, fordomme eller endda fysisk skade, hvis oplysningerne blev offentliggjort).
 - b. Sikre, at visse datakategorier ikke indsamles, eller at visse kilder udelukkes fra dataindsamling; dette kan f.eks. omfatte visse websteder, som er særligt indgribende på grund af deres emneområdes følsomhed.
 - c. Undgå indsamling fra websteder (eller dele af websteder), der klart modsætter sig webskrabning og genbrug af deres indhold med henblik på at opbygge databaser til AI-træning (f.eks. ved at respektere robots.txt- eller ai.txt-filer eller enhver anden anerkendt mekanisme til at udtrykke, at automatisk crawling eller skrabning ikke skal anvendes på deres indhold).
 - d. Indføre andre relevante begrænsninger for indsamling, eventuelt herunder kriterier baseret på tidsperioder.
106. I forbindelse med webskrabning kan eksempler på specifikke foranstaltninger, der **letter udøvelsen af enkeltpersoners rettigheder og gennemsigtighed**, omfatte: oprettelse af en fravalgsliste, der forvaltes af den dataansvarlige, og som gør det muligt for de registrerede at gøre indsigelse mod indsamlingen af deres data på bestemte websteder eller onlineplatforme ved at give oplysninger, der identificerer dem på disse websteder, herunder inden dataindsamlingen finder sted⁸⁰.
107. **Specifikke overvejelser vedrørende afbødende foranstaltninger i idriftsættelsesfasen:** Nogle af ovennævnte foranstaltninger kan også være relevante for idriftsættelsesfasen, men afhængigt af omstændighederne fastlægger Databeskyttelsesrådet nedenfor en ikke-udtømmende liste over

⁷⁹ Denne praksis kan også rejse yderligere spørgsmål, der ikke er omfattet af denne udtalelse, jf. f.eks. Pagallo U., Ciani Sciolla J., *Anatomy of web data scraping: etik, standarder and the troubles of the law*. European Journal of Privacy Law & Technologies, (2023) 2 s. 1 - 19, tilgængelig på: <https://doi.org/10.57230/EJPLT232PS>.

⁸⁰ Medmindre den dataansvarlige påviser vægtige legitime grunde til behandlingen, der går forud for registreredes interesser, rettigheder og frihedsrettigheder, eller behandlingen er nødvendig for, at retskrav kan fastlægges, gøres gældende eller forsvares.

yderligere støtteforanstaltninger, der kan gennemføres, og som bør vurderes af tilsynsmyndighederne i hvert enkelt tilfælde.

- a. **Tekniske foranstaltninger** kan f.eks. iværksættes for at forhindre opbevaring, gengivelse eller generering af personoplysninger, navnlig i forbindelse med generative AI-modeller (f.eks. outputfiltre), og/eller for at mindske risikoen for ulovlig videreanvendelse gennem AI-modeller til generelle formål (f.eks. digital vandmærkning af AI-genererede produkter).
 - b. **Foranstaltninger, der letter eller fremskynder enkeltpersoners udøvelse af deres rettigheder** i idriftsættelsesfasen, ud over hvad der kræves i henhold til loven, især med hensyn til, men ikke begrænset til, udøvelsen af retten til sletning af personoplysninger fra modellens outputdata eller deduplikering og teknikker for at fjerne eller slette personoplysninger efter træning af modellen.
108. Ved undersøgelse af idriftsættelsen af en specifik AI-model bør tilsynsmyndigheder tage hensyn til, om den dataansvarlige har offentliggjort den afvejningstest, den har udført, da der derved kan opnås mere gennemsigtighed og rimelighed. Som nævnt i Databeskyttelsesrådets retningslinjer for legitim interesse kan andre foranstaltninger overvejes for at give registrerede oplysninger fra afvejningstesten, inden indsamlingen af personoplysninger finder sted ⁸¹. Databeskyttelsesrådet gentager også ⁸², at et aspekt, der skal inddrages, er, om den dataansvarlige har inddraget databeskyttelsesrådgiveren, hvor det er relevant.

3.4 Mulig indvirkning af ulovlig behandling i forbindelse med udviklingen af en AI-model på lovligheden af den efterfølgende behandling eller anvendelse af AI-modellen

109. Dette afsnit i udtalelsen omhandler spørgsmål 4 i anmodningen. Dette spørgsmål søger en afklaring af den mulige indvirkning en ulovlig behandling i udviklingsfasen kan have på den efterfølgende behandling (f.eks. i idriftsættelsesfasen af AI-modellen) eller på anvendelsen af modellen. Spørgsmålet har til formål at behandle både den situation, hvor en sådan AI-model behandler personoplysninger, der opbevares i modellen (spørgsmål 4, (i), i anmodningen), og den situation, hvor ingen behandling af personoplysninger længere er involveret i idriftsættelsen af AI-modellen (dvs. modellen er anonym) (spørgsmål 4, (ii), i anmodningen).
110. Inden der tages stilling til visse specifikke scenarier, fremsætter Databeskyttelsesrådet følgende generelle betragtninger.
111. For det første vil afklaringerne i dette afsnit fokusere på behandlingen af personoplysninger i den udviklingsfase, der gennemføres, samtidig med at princippet om lovlighed, jf. artikel 5, stk. 1, litra a), i GDPR og artikel 6 i GDPR specifikt ikke følges (i det følgende benævnt "**ulovlighed**")⁸³. På samme måde vil Databeskyttelsesrådets overvejelser fokusere på, hvordan den ulovlige behandling i udviklingsfasen indvirker på lovligheden (dvs. overholdelse af artikel 5, stk. 1, litra a), i GDPR og artikel 6 i GDPR) af den efterfølgende behandling eller anvendelse af modellen. Databeskyttelsesrådet påpeger imidlertid, at den behandling, der foretages i udviklingsfasen, også kan føre til overtrædelser af andre bestemmelser

⁸¹ Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 68.

⁸² Databeskyttelsesrådets retningslinjer 1/2024 for behandling af personoplysninger på grundlag af artikel 6, stk. 1, litra f), i GDPR, version 1.0, vedtaget den 8. oktober 2024, punkt 12.

⁸³ EU-Domstolen, dom af 4. maj 2023, sag C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), præmis 55-57.

i GDPR såsom manglende gennemsigtighed over for registrerede, eller databeskyttelse gennem design og/eller standardindstillinger, som ikke analyseres i denne udtalelse.

112. For det andet vil ansvarlighedsprincippet, som kræver, at dataansvarlige er ansvarlige for og påviser overholdelse af *bl.a.* artikel 5, stk. 1 og artikel 6 i GDPR, spille en central rolle i forbindelse⁸⁴ med behandlingen af dette spørgsmål. Dette gælder også for behovet for at vurdere, hvilken organisation der er dataansvarlig for den pågældende behandlingsaktivitet, og om der opstår situationer med fælles dataansvar (da de kan være uløseligt forbundne)⁸⁵. I betragtning af betydningen af de faktiske omstændigheder i hver enkelt sag, herunder med hensyn til den rolle, som hver part, der er involveret i behandlingen, spiller, bør Databeskyttelsesrådets overvejelser forstås som generelle bemærkninger, der bør vurderes fra sag til sag af tilsynsmyndighederne.
113. For det tredje fremhæver Databeskyttelsesrådet, at tilsynsmyndighederne i henhold til GDPR's artikel 51, stk. 1, er "*ansvarlige for at overvåge anvendelsen af [GDPR] for at beskytte fysiske personers grundlæggende rettigheder og frihedsrettigheder i forbindelse med behandling og lette den frie udveksling af personoplysninger i Unionen*". Det ligger derfor inden for tilsynsmyndighedernes kompetence at vurdere lovligheden af behandlingen og at udøve deres beføjelser i henhold til GDPR i overensstemmelse med deres nationale rammer⁸⁶. I sådanne tilfælde har tilsynsmyndighederne skønsmæssige beføjelser til at vurdere den eller de mulige overtrædelser og vælge passende, nødvendige og forholdsmæssige foranstaltninger blandt dem, der er nævnt i artikel 58 i GDPR, under hensyntagen til omstændighederne i hver enkelt sag⁸⁷.
114. **Når det konstateres, at der foreligger en overtrædelse, kan tilsynsmyndigheder træffe korrigerende foranstaltninger, såsom at pålægge de dataansvarlige under hensyntagen til omstændighederne i hver enkelt sag at træffe foranstaltninger for at afhjælpe ulovligheden ved den oprindelige behandling.** Det kan f.eks. omfatte udstedelse af en bøde, pålæggelse af en midlertidig begrænsning af behandlingen, sletning af en del af det datasæt, der blev behandlet ulovligt, eller, hvis dette ikke er muligt, afhængigt af de foreliggende faktiske omstændigheder, under hensyntagen til foranstaltningens proportionalitet, påbud om sletning af hele det datasæt, der blev brugt til at udvikle AI-modellen og/eller selve AI-modellen. Ved vurderingen af proportionaliteten af den påtænkte foranstaltning kan tilsynsmyndighederne tage hensyn til foranstaltninger, som den dataansvarlige kan træffe for at afhjælpe ulovligheden af den oprindelige behandling (f.eks. træne modellen om).

⁸⁴ EU-Domstolen, dom af 4. maj 2023, sag C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), præmis 53.

⁸⁵ Databeskyttelsesrådets retningslinjer 07/2020 for begreberne dataansvarlig og databehandler i GDPR, udgave 2.1, vedtaget den 7. juli 2021, punkt 55.

⁸⁶ Det kan være nødvendigt at tage hensyn til specifikke nationale bestemmelser. Se f.eks. artikel 2-decies i den italienske lov om databeskyttelse (lovdekret 196/2003), som fastlægger, at oplysninger, der behandles i strid med databeskyttelsesreglerne, ikke kan anvendes. Dette berører ikke andre nationale juridiske rammer, som f.eks. straffeloven.

⁸⁷ Se i denne forbindelse betragtning 129 i GDPR samt EU-Domstolens dom af 26. september 2024, sag C-768-21, *TR mod Land Hessen* (ECLI:EU:C:2024:785), præmis 37; EU-Domstolens dom af 7. december 2023, forenede sager C-26/22 og C-64/22, *SCHUFA Holding (Libération de reliquat de dette)* (ECLI:EU:C:2023:958), præmis 57; og EU-Domstolens dom af 14. marts 2024, sag C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), præmis 34.

115. Databeskyttelsesrådet fremhæver også, at når personoplysninger behandles ulovligt, kan registrerede anmode om sletning af deres personoplysninger på de betingelser, der er fastsat i artikel 17 i GDPR, og at tilsynsmyndigheder kan beordre sletning af personoplysninger ⁸⁸*på eget initiativ*.
116. Ved vurderingen af, om en foranstaltning er hensigtsmæssig, nødvendig og forholdsmæssig, kan tilsynsmyndighederne bl.a. tage hensyn til de risici, der er forbundet hermed for de registrerede, overtrædelsens grovhed, foranstaltningens tekniske og finansielle gennemførlighed samt mængden af personoplysninger.
117. Endelig minder Databeskyttelsesrådet om, at de foranstaltninger, der træffes af tilsynsmyndigheder i henhold til GDPR, ikke berører dem, der træffes af kompetente myndigheder i henhold til AI-forordningen og/eller i henhold til andre gældende retlige rammer (f.eks. lovgivning om civilretligt ansvar).
118. I de efterfølgende afsnit vil Databeskyttelsesrådet behandle tre scenarier, der er omfattet af spørgsmål 4 i anmodningen, hvor forskellene består i, om de personoplysninger, der behandles for at udvikle modellen, opbevares i modellen, og/eller om den efterfølgende behandling foretages af den samme eller en anden dataansvarlig.

3.4.1 Scenarie 1. En dataansvarlig behandler ulovligt personoplysninger for at udvikle modellen, personoplysningerne opbevares i modellen og behandles efterfølgende af den samme dataansvarlige (f.eks. i forbindelse med idriftsættelsen af modellen).

119. Dette scenarie vedrører spørgsmål 4(i) i anmodningen, i en situation, hvor en dataansvarlig ulovligt behandler personoplysninger (dvs. ikke overholder artikel 5, stk. 1, litra a), i GDPR og artikel 6 i GDPR) for at udvikle en AI-model, og AI-modellen indeholder oplysninger om en identificeret eller identificerbar fysisk person og er derfor ikke anonym. Personoplysninger behandles derefter efterfølgende af den samme dataansvarlige (f.eks. i forbindelse med idriftsættelsen af modellen). Med hensyn til dette scenarie fremsætter Det Europæiske Databeskyttelsesråd følgende betragtninger.
120. Tilsynsmyndighedens beføjelse til at pålægge korrigerende foranstaltninger for den oprindelige behandling (som forklaret i punkt 113, 114 og 115 ovenfor) vil i princippet have en indvirkning på den efterfølgende behandling (hvis tilsynsmyndigheden f.eks. pålægger den dataansvarlige at slette de personoplysninger, der blev behandlet ulovligt, vil sådanne korrigerende foranstaltninger ikke give sidstnævnte mulighed for efterfølgende at behandle de personoplysninger, der var omfattet af foranstaltningerne).
121. Med hensyn til indvirkningen af den ulovlige behandling i udviklingsfasen på den efterfølgende behandling (f.eks. i idriftsættelsesfasen) minder Databeskyttelsesrådet om, at det er op til tilsynsmyndighederne at foretage en analyse fra sag til sag, der tager hensyn til de specifikke omstændigheder i hver enkelt sag.
122. **Hvorvidt udviklings- og idriftsættelsesfaserne omfatter særskilte formål (og dermed udgør særskilte behandlingsaktiviteter), og i hvilket omfang manglen på retsgrundlag for den oprindelige behandlingsaktivitet påvirker lovligheden af den efterfølgende behandling, bør vurderes fra sag til sag afhængigt af sagens kontekst.**

⁸⁸ Se Databeskyttelsesrådets udtalelse 39/2021 om, hvorvidt artikel 58, stk. 2, litra g), i GDPR kan tjene som retsgrundlag for en tilsynsmyndighed til uden videre at beordre sletning af personoplysninger i en situation, hvor en sådan anmodning ikke er indgivet af den registrerede, punkt 28. Se også i denne forbindelse EU-Domstolens dom af 14. marts 2024, sag C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), præmis 42.

123. For så vidt angår f.eks. retsgrundlaget i artikel 6, stk. 1, litra f), i GDPR, bør der, når den efterfølgende behandling er baseret på legitim interesse, tages hensyn til den omstændighed, at den indledende behandling var ulovlig, i vurderingen af den legitime interesse (f.eks. med hensyn til risiciene for de registrerede eller den omstændighed, at de registrerede måske ikke forventer en sådan efterfølgende behandling). I disse tilfælde kan den ulovlige behandling i udviklingsfasen påvirke lovligheden af den efterfølgende behandling.

3.4.2 Scenarie 2. En dataansvarlig behandler personoplysninger ulovligt for at udvikle modellen, personoplysningerne opbevares i modellen og behandles af en anden dataansvarlig i forbindelse med idriftsættelsen af modellen

124. Dette scenarie vedrører spørgsmål 4,(i), i anmodningen. Det adskiller sig fra scenarie 1 (i afsnit 3.4.1 i denne udtalelse), da personoplysninger efterfølgende behandles af en anden dataansvarlig i forbindelse med idriftsættelsen af AI-modellen.
125. Databeskyttelsesrådet minder om, at det er vigtigt at fastlægge de roller, der er tildelt disse forskellige aktører i henhold til databeskyttelsesrammen, for at kortlægge, hvilke forpligtelser i henhold til GDPR der finder anvendelse, og hvem der er ansvarlig for disse forpligtelser, og at situationer med fælles dataansvar også bør tages i betragtning ved vurderingen af hver parts ansvar i henhold til GDPR. Derfor bør nedenstående bemærkninger betragtes som generelle faktorer, som tilsynsmyndighederne bør tage hensyn til, hvor det er relevant. Med hensyn til scenarie 2 fremsætter Databeskyttelsesrådet følgende betragtninger.
126. For det første skal det erindres, at i henhold til artikel 5, stk. 1, litra a), i GDPR, set i sammenhæng med artikel 5, stk. 2, i GDPR, skal hver dataansvarlig sikre lovligheden af den behandling, vedkommende selv udfører, og være i stand til at påvise den. Derfor bør tilsynsmyndighederne vurdere lovligheden af den behandling, der udføres af i) den dataansvarlige, der oprindeligt udviklede AI-modellen, og ii) den dataansvarlige, der erhvervede AI-modellen og selv behandler personoplysningerne.
127. For det andet er overvejelserne i afsnit 113, 114 og 115 ovenfor relevante i denne sag med hensyn til tilsynsmyndighedernes beføjelser til at gribe ind i forbindelse med den indledende behandling. Artikel 17, stk. 1, litra d), i GDPR (sletning af ulovligt behandlede oplysninger) og artikel 19 i GDPR (underretningspligt vedrørende berigtigelse eller sletning af personoplysninger eller begrænsning af behandling) kan, afhængigt af sagens omstændigheder, også være relevante i denne sammenhæng, f.eks. i forbindelse med den underretning, som den dataansvarlige, der udvikler modellen, skal foretage over for den dataansvarlige, der idriftsætter modellen.
128. For det tredje bør tilsynsmyndigheder i hvert enkelt tilfælde foretage en vurdering af de konsekvenser en ulovlig indledende behandling kan have for en efterfølgende behandling, der foretages af en anden dataansvarlig.
129. **Tilsynsmyndigheder bør tage hensyn til, om den dataansvarlige, der anvender modellen, har foretaget en passende vurdering som led i sine ansvarlighedsforpligtelser⁸⁹ for at påvise overholdelse af artikel 5, stk. 1, litra a), og artikel 6 i GDPR for at sikre, at AI-modellen ikke blev udviklet ved ulovlig behandling af personoplysninger.** En sådan evaluering foretaget af tilsynsmyndighederne bør tage hensyn til, om den dataansvarlige har vurderet visse ikkeudtømmende kriterier såsom kilden til oplysningerne, og om AI-modellen er resultatet af en overtrædelse af GDPR, navnlig hvis den blev konstateret af en tilsynsmyndighed eller en domstol, således at den

⁸⁹ Artikel 5, stk. 2, og artikel 24 i GDPR.

dataansvarlige, der idriftsætter modellen, ikke kunne se bort fra, at den oprindelige behandling var ulovlig.

130. Den dataansvarlige bør f.eks. overveje, om oplysningerne stammer fra et brud på persondatasikkerheden, eller om behandlingen var genstand for en tilsynsmyndigheds eller en domstols konstaterede overtrædelse. **Omfanget af den dataansvarliges vurdering og den detaljeringsgrad, der forventes af tilsynsmyndighederne, kan variere afhængigt af forskellige faktorer, herunder typen og graden af risici, som behandlingen i AI-modellen medfører, når den idriftsættes, for de registrerede, hvis data blev brugt til at udvikle modellen.**
131. Databeskyttelsesrådet bemærker, at AI-forordningen kræver, at udbydere af højrisiko-AI-systemer udarbejder en EU-overensstemmelseserklæring⁹⁰, og at en sådan erklæring indeholder en erklæring om, at det relevante AI-system overholder EU's databeskyttelseslovgivning⁹¹. Databeskyttelsesrådet bemærker, at en sådan egenerklæring måske ikke udgør en endelig konstatering af overensstemmelse i henhold til GDPR. Tilsynsmyndigheder kan ikke desto mindre tage hensyn til det, når de undersøger en specifik AI-model.
132. De samme overvejelser, som blev gjort under punkt 123 ovenfor, er også relevante i dette tilfælde. Når tilsynsmyndigheder kontrollerer, om og hvordan den dataansvarlige har vurderet, om legitim interesse er et passende retsgrundlag for den behandling, denne foretager, bør der tages højde for ulovligheden af den oprindelige behandling som en del af vurderingen af legitim interesse, f.eks. ved at vurdere de potentielle risici, der kan opstå for de registrerede, hvis personoplysninger er blevet behandlet ulovligt for at udvikle modellen. Forskellige aspekter, enten af teknisk art (f.eks. eksistensen af filtre eller adgangsbegrænsninger, der er indført under udviklingen af modellen, som den efterfølgende dataansvarlige ikke kan omgå eller påvirke, og som kan forhindre adgang til eller videregivelse af personoplysninger) eller af retlig karakter (f.eks. hvilken form for ulovlighed og alvoren heraf i forbindelse med den oprindelige behandling), skal tages behørigt i betragtning i forbindelse med afvejningstesten.

3.4.3 Scenarie 3. En dataansvarlig behandler personoplysninger ulovligt for at udvikle modellen og sikrer derefter, at modellen anonymiseres, inden den samme eller en anden dataansvarlig indleder en anden behandling af personoplysninger i forbindelse med idriftsættelsen

133. Dette scenarie vedrører spørgsmål 4, (ii), i anmodningen og henviser til en sag, hvor en dataansvarlig ulovligt behandler personoplysninger for at udvikle AI-modellen, men gør det på en måde, der sikrer, at personoplysninger anonymiseres, inden den samme eller en anden dataansvarlig indleder en anden behandling af personoplysninger i forbindelse med idriftsættelsen. For det første minder Databeskyttelsesrådet om, at tilsynsmyndighederne er kompetente og har beføjelse til at gribe ind med hensyn til behandlingen i forbindelse med anonymiseringen af modellen samt den behandling, der foretages i udviklingsfasen. Tilsynsmyndigheder kan således, afhængigt af sagens specifikke omstændigheder, pålægge korrigerende foranstaltninger i forbindelse med denne oprindelige behandling (som forklaret i punkt 113, 114 og 115 ovenfor)
134. Hvis det kan påvises, at den efterfølgende anvendelse af AI-modellen ikke indebærer behandling af personoplysninger, mener Databeskyttelsesrådet, at GDPR ikke vil finde anvendelse⁹². Ulovligheden af den oprindelige behandling bør derfor ikke påvirke den efterfølgende anvendelse af modellen.

⁹⁰ Artikel 16, litra g), og artikel 47 i AI-forordningen.

⁹¹ Bilag V, punkt 5 i AI-forordningen.

⁹² Betragtning 26 i GDPR.

Databeskyttelsesrådet understreger imidlertid, at et simpelt udsagn om, at modellen er anonym, ikke er tilstrækkelig til at fritage den fra anvendelsen af GDPR, og bemærker, at tilsynsmyndigheder bør vurdere den i hvert enkelt tilfælde under hensyntagen til de overvejelser, som Databeskyttelsesrådet har fremsat for at besvare spørgsmål 1 i anmodningen.

135. **Når de dataansvarlige efterfølgende behandler personoplysninger, der er indsamlet i løbet af idriftsættelsesfasen, efter at modellen er blevet anonymiseret, vil GDPR finde anvendelse i forbindelse med disse behandlingsaktiviteter. I disse tilfælde bør lovligheden af den behandling, der foretages i idriftsættelsesfasen, for så vidt angår GDPR ikke påvirkes af den oprindelige behandlings ulovlighed.**

4 Afsluttende bemærkninger

136. Denne udtalelse er rettet til alle tilsynsmyndigheder og offentliggøres i henhold til GDPR's artikel 64 stk. 5, litra b).

På vegne af Det Europæiske Databeskyttelsesråd

Formanden

(Anu Talus)