RÄTTSLIGT STÖD FÖR MYNDIGHETERS UTVECKLING AV AI-TJÄNSTER FÖR UTLÄMNANDE AV ALLMÄNNA HANDLINGAR

ADVOKATFIRMAN KAHN PEDERSEN

Sammanfattning

1. Denna rapport har upprättats inom ramen för projektet ”Rätt till insyn”, i vilket Lidingö stad, Kungsbacka kommun, Skaraborgs kommunalförbund, Linköpings universitet och Advokatfirman Kahn Pedersen gemensamt – och med finansiellt stöd av Vinnova – utreder förutsättningarna att förenkla och automatisera processen för begäran om allmän handling med hjälp av AI.

2. I rapporten analyserar vi två generella och centrala dataskyddsrättsliga frågor för myndigheters utveckling och användning av AI:

   - Om och i vilken mån kan AI utvecklas och användas **utan behandling** av personuppgifter? Denna fråga utreds i avsnitt 2.
   - Med **vilken rättslig grund** kan myndigheter behandla personuppgifter vid utveckling och användning av AI-tjänster? Denna fråga utreds i avsnitt 3.

3. I rapporten utreder vi även de rättsliga förutsättningar för den specifika AI-tjänst som utvecklas i projektet ”Rätt till insyn”. Vi bedömer att det finns rättsligt stöd för behandling av personuppgifter för den s.k. maskeringstjänsten och, troligen, även för den mer omfattande ”helhetstjänsten”. Se avsnitt 4.

4. I rapporten presenterar vi bl.a. följande slutsatser och observationer:

   - **Det finns typiskt sett rättsligt stöd för myndigheter att utveckla AI i syfte att främja myndigheternas verksamhet och uppdrag**. Vi menar att sådan verksamhetsutveckling typiskt sett ryms inom den rättsliga grunden ”uppgift av allmänt intresse” och många gånger även utgör ett sådant ”viktigt allmänt intresse” som krävs för behandling av känsliga personuppgifter. Det förutsätter dock en bedömning av (i) det allmänna intresset som ska främjas, (ii) det rättsliga stödet i nationell rätt eller unionsrätten, (iii) att behandlingen är nödvändig, (iv) att behandlingen är proportionerlig i förhållande till integritetsriskerna, samt (v) särskilt stöd för behandling av känsliga personuppgifter.

   - **Myndigheterna behöver tydligare rättsligt stöd för utveckling av AI-tjänster**. För att främja en rättssäker och effektiv användning av AI-tekniken och underlätta utvecklingen av innovativa AI-tjänster inom den offentliga sektorn kan det krävas förtydliganden i lag som underlättar myndigheternas bedömningar och sätter gränserna för hur personuppgifter får behandlas. Det kan exempelvis handla om förtydliganden av eller kompletterande ändamålsbestämmelser, översyn av sökbegränsningar i registerförfattningar eller en justering av bestämmelsen i 3 kap. 3 § 3 dataskyddslagen. Se avsnitt 5.1.

   - **AI-tjänster kan medföra att offentlighetsprincipen utökas väsentligt**. En sådan utveckling ökar allmänhetens insyn men medför samtidigt större integritetsrisker i och med att personuppgifter i myndigheternas diarier och databaser kan komma att bli tillgängliga på ett sätt som inte går att förutse och spridas i större omfattning än tidigare. Detta kan i sin tur leda till komplicerade och svåra sekretessbedömningar vid utlämnande av allmänna handlingar. Se avsnitt 5.2.

kahn pedersen

Innehåll

1. UPPDRAG, AVGRÄNSNINGAR OCH DEFINITIONER

1.1 Rapportens innehåll

Denna rapport är en del av det Vinnova-finansierade projektet ”Rätt till insyn”[1] och avser rättsliga frågeställningar kopplade till AI inom den offentliga sektorn med särskilt fokus på utlämnande av allmänna handlingar i kommunal verksamhet.

I rapporten analyserar vi två centrala dataskyddsrättliga frågor för utveckling av AI-tjänster inom den offentliga sektorn, dels under vilka förutsättningar utvecklingen av AI-tjänster kan göras utan personuppgifter (avsnitt 2), dels vilket rättsligt stöd som finns för behandling av personuppgifter i samband med utvecklingen av AI-tjänster inom den offentliga sektorn (avsnitt 2.7). Därutöver tittar vi särskilt på de AI-tjänster som Lidingö stad tillsammans med Skaraborgs kommunalförbund, Linköpings universitet, Kungsbacka kommun och Atea utvecklar inom projektet ”Rätt till insyn”, inklusive planerna på en mer omfattande tjänst för automatisering av utlämnande av allmänna handlingar, den s.k. helhetstjänsten (avsnitt 4). Avslutningsvis redovisar vi några övergripande observationer och pekar därvid på behovet av åtgärder från lagstiftarens sida (avsnitt 5).

Målgruppen är i första hand jurister inom offentlig förvaltning. Vår förhoppning är att rapporten ska kunna fungera som vägledning och utgångspunkt för andra projekt för utveckling av AI för myndigheter.

1.2 Avgränsningar

Vi behandlar huvudsakligen rättsfrågor som är kopplade till utvecklingsfasen av AI-tjänster. Andra rättsliga bedömningar kan behöva göras vid användning av AI-tjänster. Många gånger är användningen av färdigutvecklade AI-tjänster närmare knuten till en behandling som redan utförs av myndigheten medan utvecklingen av en ny AI-tjänst utgör ett särskilt ändamål skilt från myndighetens normala verksamhet.

Vi har i denna rapport fokuserat på tillämpningen av dataskyddsförordningen vid utvecklingen av AI-tjänster inom den offentliga sektorn. Vi bedömer således vissa rättsliga frågor som uppkommer vid tillämpningen av dataskyddsförordningen men även i viss utsträckning frågor som har samband med offentlighets- och sekretesslagen, 2009:400 (OSL), förvaltningslagen 2017:900 (FL) och AI-förordningen[2]. Notera att rapporten inte gör anspråk på att analysera samtliga dataskyddsfrågor kopplade till utvecklingen av AI-tjänster. Exempelvis har vi inte utrett skyldigheten att informera de registrerade eller de registrerades rättigheter inklusive rätten att inte bli föremål för automatiserat beslutsfattande i artikel 22 dataskyddsförordningen.

[1] För mera information om Vinnova-projektet, se https://www.insyn.info/.

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

1.3 Några begrepp som vi använder

Begreppsförvirring präglar stora delar av AI-utvecklingen. I denna rapport använder vi begreppet AI med en bred innebörd och preciserar vad vi menar när det finns behov av preciseringar. I våra exempel försöker vi dock i möjligaste mån utgå från de senaste årens rätts- och teknikutveckling rörande AI.

Vi använder således begreppet AI-tjänst för tjänster som innehåller delar som skulle kunna definieras som AI, inklusive generativ AI, maskininlärning och symbolisk AI (det som i AI-förordningen benämns ”AI-system”)[3] men som även kan innefatta mer traditionella komponenter konstruerade med traditionell systemutveckling.

Med AI-modell menar vi en mjukvarubaserad arkitektur som med parametrar/vikter eller på annat sätt kan tränas för en viss uppgift – vanligen någon typ av kategorisering eller förutsägelse baserad på vissa indata – med hjälp av träningsdata. Detta innefattar bl.a. AI-modeller för allmänna ändamål så som dessa definieras i AI-förordningen, exempelvis generativa modeller, men kan även omfatta mer specialiserade modeller. Exempel på sådana specialiserade modeller som används av vissa myndigheter är riskbaserade kontroller av utbetalningar från välfärdssystemen.[4]

Av särskilt intresse är stora språkmodeller (Large language models, LLM) som utvecklas och tränats av flera aktörer, exempelvis Google, Meta och OpenAI. Dessa språkmodeller kan användas för allmänna ändamål, exempelvis för att granska en text utifrån givna instruktioner (”promptning”). En LLM kan användas som den är, och då enbart kontrolleras genom promptning, eller anpassas (”fine-tuning”) till vissa specifika uppgifter genom exempel. Leverantörer av LLM erbjuder regelmässigt åtkomst till dessa som tjänst, vilket innebär att indata till modellen överförs till leverantören. Det är också vanligt att LLMs erbjuds i nedladdningsbar form för att kunna tas i drift i exempelvis en myndighets egen it-miljö eller hos en tredje part. I dessa fall har myndigheten större kontroll över vem som får tillgång till indata till AI-systemet.

[3] Se bl.a. definitionen av AI-system i AI-förordningen (artikel 3.1) och OECD

[4] Se exv. Delegationen för korrekta utbetalningar från välfärdssystemen, *Digitalisering och AI för korrekta utbetalningar från välfärdssystemen, Rapport 5*, s 46.

2. HUR KAN AI UTVECKLAS OCH ANVÄNDAS UTAN PERSONUPPGIFTER?

2.1 I vilka steg kan AI-användning medföra personuppgiftsbehandling?

För AI-modeller som baseras på inställning av parametrar/vikter är det viktigt att skilja på den initiala träningen av modellen och den efterföljande användningen (även kallad inferensfasen). Specifikt för stora generativa språkmodeller kan man skilja på grundträning, finjustering, promptutforning, vilka utgör olika steg i *utvecklingen* av modellen, respektive generering, som utgör *användning* av modellen.

Under grundträningsfasen (i tekniska sammanhang ofta benämnd ”pre-training”) behandlas mycket stora datamängder för att lära upp modellen om hur språk i allmänhet fungerar, specifikt hur ord och delar av ord vanligtvis följer varandra. Den kunskap som finns i datamängderna representeras i någon form av den upptränade modellen. Träningsmängden i detta steg är typiskt sett inte avgränsad till endast material av en viss typ (såsom allmänna handlingar från myndigheter) eller ens ett specifikt språk, utan ju större datamängder desto bättre. De aktuella datamängderna innehåller typiskt sett personuppgifter, exempelvis i form av tidningsartiklar. Grundträning av en stor språkmodell kräver mycket stora investeringar i datorkraft och är utanför de flesta myndigheters förmåga.

Under finjusteringsfasen (”finetuning”) tränas modellen vidare på mer specifika datamängder och utifrån vissa mål. Det kan exempelvis vara fråga om uppsättningar av frågor med önskvärda svar. Denna träningsdatamängd är typiskt sett utvald med större omsorg utifrån det önskade beteendet hos systemet, bl.a. vilken sakkunskap man önskar att systemet ska lära sig av. Detta kan kräva att datat innehåller (riktiga eller fingerade) personuppgifter om det önskade beteendet är att behandla personuppgifter på något visst sätt. Finetuning är typiskt sett möjligt för en myndighet att genomföra om man själv tagit modellen i drift, men det förekommer även att finetuning erbjuds som tjänst på samma sätt som språkmodellen som sådan.

Promptutformningen instruerar språkmodellen hur den ska generera innehåll i det specifika fallet. Detta sker genom utformning av instruktioner för en uppgift med naturligt språk, exempelvis ”Summera följande dokument i högst tio punkter”. Den specifika promptutformningen är, tillsammans med egenskaperna hos och kunskapen i den specifika modellen, avgörande för AI-systemets beteende. En myndighet som vill använda generativa språkmodeller i sin egen tjänst behöver utforma en eller flera lämpliga prompter för systemets önskade beteende. Prompter kan innehålla (typiskt sett fingerade) personuppgifter om detta gör instruktionerna tydligare.

Generering använder den finjusterade modellen, den utformade prompten och specifik indata (såsom innehållet i ett dokument) för att generera någon form av utdata. Den specifika indatan kommer i många fall innehålla personuppgifter.

2.2 Informationssamlingar innehåller vanligtvis personuppgifter

Utveckling av AI-modeller är, som beskrivits ovan, regelmässigt beroende av tillgång till data för träning. För stora språkmodeller utgörs träningsdata av löpande text. I många fall är det svårt att ta fram träningsdata som inte innehåller personuppgifter, vilket medför behov av att analysera hur sådan AI-träning är förenlig med dataskyddslagstiftningen. Om dataskyddsförordningen är tillämplig medför det begränsningar i vilken data som får användas för det avsedda ändamålet och hur data med personuppgifter ska hanteras (vi återkommer till den frågan i avsnitt 3 nedan).

Det är således viktigt att granska vad som utgör en personuppgift enligt dataskyddslagstiftningen och vilka metoder som kan användas för att ”rensa” data från personuppgifter eller att minska integritetsriskerna. Ett problem i sammanhanget är att begreppet personuppgift har fått en relativt vag definition och är avsett att ha en vidsträckt innebörd. Att avidentifiera eller anonymisera data som innehåller personuppgifter anses i regel svårt att åstadkomma i praktiken.

2.3 Vad är en personuppgift?

En personuppgift är enligt dataskyddsförordningen all information som avser en identifierad eller identifierbar fysisk person.[5] Definitionen kan framstå som enkel men innehåller många komplicerade frågor som har analyserats och hanteras i ett stort antal vägledningar och domar från bl.a. EU-domstolen.[6] I huvudsak är det två kriterier som är av särskild betydelse, dels vad som utgör en identifierad, och framför allt, en identifierbar person, dels vilken information som ”avser” en sådan person.

En person anses *identifierad eller identifierbar* när denne kan särskiljas från andra personer i en grupp. Identifiering kan förstås ske direkt genom mer eller mindre unika identifierare såsom namn, personnummer eller indirekt genom att identifierare kombineras med andra upplysningar såsom adress eller arbetsplats. Identifiering kan även ske indirekt genom en kombination av upplysningar t.ex. information i en profil som innehåller information om en unik användares beteende, ekonomi, intressen utan att personen identifieras med namn eller annan unik identifierare. Det är således inte nödvändigt att kunna identifiera någon med namn eller personnummer för att informationen ska utgöra en personuppgift. En person kan även urskiljas från en grupp, och därigenom identifieras, genom identifiering av mönster i en större datamängd, vilket är särskilt intressant men tanke på att mönsterigenkänning är en av AI:s främsta förmågor.

Det som komplicerar definitionen är även att man måste väga in inte bara den information och de hjälpmedel som den personuppgiftsansvarige har utan även information och hjälpmedel som någon annan har tillgång till och som ”… rimligen kan

[5] Se fullständig definitionen i artikel 4.1 dataskyddsförordningen.

[6] Se exempelvis EU-domstolens dom den 19 oktober 2016 i mål C-582/14 ("Breyer") och EU-domstolens dom den 7 mars 2024 i mål C-604/22 ("IAB Europe").

komma att användas för att direkt eller indirekt identifiera den fysiska personen".[7] En begränsning finns i ordet "rimligen". Bedömningen av vad som utgör rimliga hjälpmedel för identifiering ska utgå från "samtliga objektiva faktorer, såsom kostnader och tidsåtgång för identifiering, med beaktande av såväl tillgänglig teknik vid tidpunkten för behandlingen som den tekniska utvecklingen".[8]

Enligt det andra kriteriet är det endast sådan information som "*avser*" en identifierad eller identifierbar person som kan utgöra en personuppgift. Det handlar om information som kan kopplas till en person genom en identifierare, men som utan sådan koppling inte utgör personuppgifter. Ett enkelt exempel är att namnet Stockholm i sig inte kan utgöra en personuppgift om alla som bor i Stockholm men i kombination med annan information t.ex. "Martin med personnumret XXX bor i Stockholm" är Stockholm en personuppgift om Martin.

Information kan avse en person på tre sätt; genom informationens innehåll, syfte eller verkan.[9] En patients läkarjournal avser till innehållet personuppgifter om patienten. Ett IP-nummer kan utgöra en personuppgift om syftet med behandlingen är att identifiera vem som har använt IP-numret. GPS-uppgifter om ett företags fordon kan till sin verkan (hur det används) få betydelse för förarnas rättigheter och friheter och därför utgöra personuppgifter om dessa. Däremot torde inte uppgifter om t.ex. fordonens motorstyrka vara personuppgifter om förarna, eftersom sådan information i regel varken till innehåll, syfte eller verkan avser förarna.

Sammanfattningsvis kan vi konstatera att personuppgiftsbegreppet är omfattande och vagt definierat i rättspraxis. Det beror bl.a. på att det kan handla om personuppgifter trots att ingen person är identifierad med namn, personnummer e.d. och trots att det inte finns någon avsikt från den som behandlar uppgifterna att identifiera någon. Användningen av personuppgifter inom AI har accentuerat dessa komplicerade gränsdragningsfrågor. Det är enligt vår bedömning inte uteslutet att AI-utvecklingen kommer att driva fram ny rättspraxis som klargör, och möjligen begränsar, omfattningen av personuppgiftsbegreppet.

2.4 Är träningsdata personuppgifter?

För att träna en AI-modell behöver träningsdata ofta processas på olika sätt. En vanligt förekommande metod är att tokenisera indata, vilket innebär översättande av enskilda ord eller delar av ord till "tokens", normalt uttryckta som heltal.

En sekvens av sådana heltal kan framstå som en datamängd utan personuppgifter, eftersom det från informationen som sådan inte går att utläsa någon information som gör det möjligt att identifiera en person. Om tokeniseringen är reversibel, dvs. om det går att återskapa den ursprungliga informationen genom att exempelvis använda

[7] Skäl 26 dataskyddsförordningen.

[8] Skäl 26 dataskyddsförordningen och EU-domstolens dom den 19 oktober 2016 i mål C-582/14 ("Breyer") p. 46 där EU-domstolen konstaterar att om de åtgärder som krävdes för identifiering av en person utgör inte rimliga hjälpmedel "om identifiering av den aktuella personen var förbjuden i lag eller omöjlig att genomföra i praktiken, exempelvis på grund av att den skulle kräva orimliga resurser i form av tid, kostnader och arbetskraft, med den följden att risken för identifiering i praktiken var försumbar."

[9] Artikel 29-gruppen, Yttrande 4/2007 om begreppet personuppgifter WP136. WP 136, s. 11.

samma tokeniseringsalgoritm baklänges, är det dock fortfarande fråga om personuppgifter. Detta gäller även om den part som utför tokenisering själv saknar förmågan att reversera processen, så länge det är möjligt för någon att direkt eller indirekt identifiera en person med beaktande av hjälpmedel och information som rimligen kan komma att användas för identifiering (se punkt 0 ovan). Andra metoder för att anonymisera eller behandla träningsdata med integritetshöjande teknik kan dock förändra denna bedömning. Se avsnitt 2.6 nedan.

2.5 AI-modeller som "läcker" personuppgifter

En AI-modell, eller närmare bestämt modellens variabler (vikter), kan under vissa förutsättningar avslöja uppgifter som har använts för att träna modellen. Det brukar kallas för att AI-modellen minns eller läcker uppgifter. Det förutsätter i regel att medvetna åtgärder genomförs för att modellen ska avslöja sådana uppgifter. I dagsläget hänvisas ofta till två sådana åtgärder eller attacker, Membership Inference Attack och Model Inversion Attack. Men frågan är föremål för omfattande forskning och flera metoder kan komma att utvecklas.[10]

Den danska dataskyddsmyndigheten, Datatilsynet, anser att en AI-modell som utgångspunkt inte innehåller personuppgifter. AI-modeller är endast resultat av en behandling av personuppgifter på samma sätt som en statistisk rapport som endast innehåller aggregerade data och som är resultatet av en statistisk analys. Datatilsynet påpekar dock att vissa maskininlärningsmodeller kan attackeras på olika sätt i syfte att identifiera personer vars uppgifter har inkluderats i modellens träningsdata. Om personuppgifter på detta sätt kan användas för återidentifiering utgör det enligt Datatilsynet en överträdelse av skyldigheten enligt dataskyddsförordningen att skydda personuppgifter med lämpliga säkerhetsåtgärder. Men denna risk innebär enligt Datatilsynet inte att AI-modellen i sig ska anses innehålla personuppgifter.[11]

I IMY-rapporten ”Federerad maskininlärning mellan två vårdgivare” bedömde IMY att det inte kunde uteslutas att ”vikter” som var resultatet av träning av lokala AI-modeller innehöll personuppgifter och att uppgifterna behandlades när dessa vikter sammanställdes i en federerad lösning mellan två vårdgivare. IMY ansåg att vikterna utgjorde personuppgifter även om det är ”omständligt” att få fram personuppgifter ur vikterna som delades mellan vårdgivarna men påpekade också att denna slutsats inte är giltig för alla former av federerad maskininlärning.[12] Någon djupare analys på vilket sätt som uppgifter i vikterna kan identifiera enskilda gjordes dock inte av IMY.

EDPB betonar att frågan huruvida en AI-modell som har tränats med personuppgifter kan vara ”anonym” (dvs. att personuppgifter inte kan extraheras vid användning av modellen), måste avgöras utifrån en bedömning i det enskilda fallet. Bedömningen ska

[10] En aktuell översikt över forskningsfältet med en taxonomi över olika typer av attacker mot AI-system har publicerats av amerikanska NIST, Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations (NIST AI 100-2e2025, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2025.pdf)

[11] Danska datatilsynet, Offentlige myndigheders brug af kunstig intelligens, Inden I går i gang, Oktober 2023, s. 7.

[12] IMY, Federerad maskininlärning mellan två vårdgivare. Slutrapport om Integritetsskyddsmyndighetens pilotprojekt med regulatorisk testverksamhet om dataskydd, dnr IMY-2023-2602, 2023-03-15, s. 19 f.

enligt EDPB göras med utgångspunkt i definitionen av personuppgifter enligt dataskyddsförordningen. En AI-modell anses därför som ”anonym” endast om både (i) sannolikheten för direkt (inklusive parametrar som representerar sannolikhetsbaserade relationer) extrahering av personuppgifter som har använts vid träning av modellen och (ii) sannolikheten för att erhålla (avsiktligt eller inte) sådana personuppgifter vid förfrågningar till AI-modellen, är obetydlig (”insignificant”), med hänsyn till "alla medel som rimligen kommer att användas av den registeransvarige eller en annan person". Metoder som kan visa på en AI-modells anonymitet kan exempelvis innefatta de åtgärder som vidtagits under utvecklingsfasen för att förhindra eller begränsa insamlingen av personuppgifter som används för träning, minska deras identifierbarhet, förhindra att de extraheras eller tillhandahålla försäkran om toppmodern motståndskraft mot attacker.[13]

Vilka attackmetoder eller liknande åtgärder som ska tas med i bedömningen av om en AI-modell innefattar personuppgifter, innefattar en komplicerad bedömning av personuppgiftsbegreppet enligt dataskyddsförordningen. Som nämnts ovan, är det fråga om vilka hjälpmedel (metoder) som är rimligen kan komma att användas för att direkt eller indirekt identifiera en person med hjälp av AI-modellen.

En ytterligare komplexitet uppstår till följd av att AI-modeller till sin natur måste ge ifrån sig någon utdata som svar på indata, och att den exakta promptningen kan påverka hur detta går till. Det är därför möjligt för AI-modeller att skapa utdata som innehåller information som utgör personuppgifter, trots att dessa inte återfinns i träningsdata (en form av s.k. hallucinationer). Även sådana ”falska” uppgifter som av en tillfällighet kan knytas till (identifiera) en person kan utgöra personuppgifter.

Vägledning i denna fråga kan ytterst endast lämnas av EU-domstolen. Domstolens praxis har än så länge inte gett något entydigt svar, men domstolen tycks gå mot ett mer vidsträckt personuppgiftsbegrepp. Enligt vår uppfattning kommer troligen AI-utvecklingen driva fram nya avgöranden i EU-domstolen, vilka möjligen medger en mer flexibel tolkning.[14]

2.6 Vad kan göras för att undvika att behandla personuppgifter?

Det finns i grunden två sätt att minska risken för att AI-modeller avslöjar personuppgifter. Antingen kan man undvika att använda personuppgifter vid träningen av AI-modellen, t.ex. genom fiktiva personuppgifter (se avsnitt 0 nedan), eller också kan man vidta åtgärder som förhindrar eller i vart fall försvårar att

[13] EDPB, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models Adopted on 17 December 2024, s. 16.

[14] I skrivande stund är ett intressant mål på väg för EU-domstolens avgörande. I målet (C-413/23) väcks frågan om pseudonymiserade uppgifter ska anses vara personuppgifter hos en mottagare som inte har tillgång till information som möjliggör identifiering. Om EU-domstolen finner att så inte är fallet kan det öppna större möjligheter att använda pseudonymiserade personuppgifter vid träning av AI-modeller. Avgörande väntas under 2025.

personuppgifter kan läcka från AI-modellen och få AI-modellen att glömma eventuella personuppgifter.[15]

IMY nämner några faktorer som skulle kunna utredas och tillämpas för att minska risken för att personuppgifter avslöjas och därigenom höja integritetsskyddet vid användning av federerad maskininlärning:[16]

- Utveckling av tekniken i en riktning som påverkar vikternas förmåga att lagra personuppgifter, till exempel genom att begränsa modellens komplexitet och träning för att minska risken för så kallad överträning.
- Utveckling av medel för att motverka användning av attacker eller andra metoder som möjliggör att personuppgifter kan återskapas från modellen.
- Riskreduktion genom säkerhetsåtgärder.
- Användning av metoder för att anonymisera personuppgifter genom utveckling av teknik som exempelvis differential privacy (en teknik som i korthet går ut på att man lägger på ett så kallat ”brus” på data för att minska risken att personuppgifter kan återidentifieras).

Till detta kan även läggas åtgärder som är tillämpliga på specifika it-lösningar. Beroende på tillämpningens krav kan både indata och utdata från modellen kontrolleras så att en utomstående saknar möjlighet att förmå modellen att ge ifrån sig personuppgifter.

Det är även möjligt att använda flera steg i behandlingen, exempelvis ett inledande steg där en separat AI-modell används för att skrubba bort personuppgifter från indata, eller ett avslutande steg där en annan AI-modell detekterar förekomster av personuppgifter (sanna eller hallucinerade) i utdata.

2.7 Fiktiva (fingerade) personuppgifter kan vara en lösning

Fiktiva personuppgifter är påhittade identitetsuppgifter om personer som inte finns. Uppgifterna ska således inte kunna användas för att identifiera "riktiga" personer. Fiktiva personuppgifter används vanligen som en åtgärd för att minska integritetsriskerna t.ex. vid testning av it-system eller AI-modeller.[17] Under vissa förutsättningar kan dock fiktiva uppgifter utgöra personuppgifter enligt dataskyddsförordningen.

Vid användningen av fiktiva personuppgifter för statistiska ändamål kan det krävas att ”riktiga” personuppgifter behandlas för att kunna skapa fiktiva personuppgifter som

[15] Se bl.a. Saskia Keskpaik, Machine unlearning i EDPS TechSonar 2025 report, s. 19.

[16] IMY, Federerad maskininlärning mellan två vårdgivare. Slutrapport om Integritetsskyddsmyndighetens pilotprojekt med regulatorisk testverksamhet om dataskydd, dnr IMY-2023-2602, 2023-03-15, s. 20.

[17] EDPB nämner fiktiva personuppgifter som en lämplig åtgärd vid träning av AI-modeller när det faktiska innehållet i data inte är relevant. EDPB, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, p. 101.

ska spegla de riktiga personuppgifterna på ett sätt som gör dem användbara för statistik, men även för träning och testning av AI-modeller. Det finns i sådana fall en risk att det går att identifiera personer från den ursprungliga uppgiftsmängden om denna mängd innehåller ett mindre antal personer eller om det finns någon eller ett fåtal individer i uppgiftsmängden som särskiljer sig från de övriga.

Ett sådant behandlingssteg, där riktiga personuppgifter i ett material ersätts av fiktiva, exempelvis genom att byta ut namn eller andra identifierande uppgifter, utgör i sig personuppgiftsbehandling, även om resultatet inte utgör personuppgifter. Det bör dock vanligen ses som en del av en större personuppgiftsbehandling (exempelvis träning av en AI-modell), och en åtgärd som typiskt sett är integritetshöjande. Om den större personuppgiftsbehandlingen har en rättslig grund och i övrigt uppfyller dataskyddsförordningens krav bör behandlingen för ”fiktivisering” omfattas av samma rättsliga grund.

I andra sammanhang när fiktiva personuppgifter används för att testa ett it-system eller för att träna AI-modeller eller liknande finns det en risk för identifiering av riktiga personer om de fiktiva uppgifterna är tillräckligt unika, så att de pekar ut en individ med t.ex. ett säreget namn. I sådana fall kan de utgöra personuppgifter enligt dataskyddsförordningen även om det inte har varit avsikten att identifiera någon riktig person, dvs. genereringen av de fiktiva uppgifterna råkade ta fram uppgifter som identifierar en person.

För att det ska handla om en personuppgift enligt dataskyddsförordningen krävs, som nämnts i avsnitt 2.3 ovan, dels att en fysisk person är identifierad eller att denne kan identifieras med kompletterande information, dels att uppgiften i fråga "avser" denna person.

Att uppgiften ska avse en fysisk person är ett objektivt kriterium och det krävs därför inte att det ska finnas en avsikt hos den som behandlar personuppgifterna om att knyta uppgiften till en viss person. Det kan vara tillräckligt att uppgiften på grund av sitt innehåll eller sin verkan är knuten till en fysisk person. Finns det däremot en avsikt att behandla uppgifter om identifierade personer är det i regel tillräckligt för att det ska utgöra personuppgifter. [18]

En förutsättning för att det ska handla om en personuppgift enligt dataskyddsförordningen är dock att uppgiften kan knytas till (avse) en identifierad eller identifierbar fysisk person. Detta medför att exempelvis ett fiktivt namn kan vara en personuppgift enligt dataskyddsförordningen om namnet är så unikt att en fysisk person identifieras med enbart namnet. Det finns troligen ett relativt stort antal personer i Sverige med unika namn, vilka på detta sätt kan identifieras.

Rekommendationen är således att säkerställa att processen för att generera fiktiva personuppgifter sker på ett sådant sätt att endast vanligt förekommande namn (eller

[18] EU-domstolen har slagit fast att en upplysning avser en identifierad eller identifierbar fysisk person när den på grund av sitt innehåll, syfte eller verkan är knuten till en identifierbar person, se EU-domstolens avgörande den 4 maj 2023 i mål C-487/21 "Österreichische Datenschutzbehörde", p. 24

andra uppgifter som genereras) används och att ovanliga namn undviks. Det gäller även om fiktiva personuppgifter tas fram av en extern leverantör som tillhandahåller sådana tjänster.

3. VILKET RÄTTSLIGT STÖD FINNS FÖR ATT BEHANDLA PERSONUPPGIFTER VID UTVECKLING AV AI I OFFENTLIG SEKTOR?

3.1 Behandling av personuppgifter förutsätter rättsligt stöd

Många gånger är det inte möjligt att undvika att använda data som innehåller personuppgifter vid utveckling eller användning av en AI-tjänst. I sådana fall är ett första steg att bedöma om myndigheter har rättsligt stöd för att behandla de aktuella personuppgifterna. Det är en central fråga för att överhuvudtaget kunna gå vidare i ett projekt om AI-utveckling och innefattar ofta svåra juridiska bedömningar.

Varje behandling av personuppgifter är en begränsning av rätten till skydd av personuppgifter och respekten för privatlivet enligt EU-stadgan. Sådana begränsningar får endast göras med beaktan av legalitets- och proportionalitetsprincipen. Det sker huvudsakligen genom att det krävs en giltig rättslig grund för att en behandling av personuppgifter ska vara tillåten och principerna om uppgifts- och lagringsminimering.

3.2 Den rättsliga grunden för myndigheters behandling

All behandling av personuppgifter måste ha stöd i någon av de rättsliga grunder som anges i artikel 6 dataskyddsförordningen. För myndigheter som behöver behandla personuppgifter för att utföra sina arbetsuppgifter är det normalt den rättsliga grunden i artikel 6.1 e som är aktuell, dvs. behandlingen är ”nödvändig för att utföra en uppgift av allmänt intresse eller som ett led i den personuppgiftsansvariges myndighetsutövning”.[19]

En förutsättning för denna rättsliga grund är, enligt artikel 6.3, att grunden för uppgiften eller myndighetsutövningen är fastställd i enlighet med unionsrätten eller medlemsstaternas nationella rätt.

Varje behandling av personuppgifter som sker under en AI-tjänsts livscykel behöver ha en rättslig grund enligt dataskyddsförordningen. Den rättsliga grund som används för utvecklingen av en AI-tjänst är inte nödvändigtvis tillämplig när AI-tjänsten används i myndighetens verksamhet. Även om vi nedan fokuserar på utveckling av AI-tjänst är det viktigt att göra en bedömning av AI-tjänstens hela livscykel (utveckling, utbildning, testning, tillämpning, efterinlärning) för att säkerställa att all behandling av personuppgifter uppfyller kraven enligt dataskyddsförordningen.

[19] Då myndighetsutövning utgör en uppgift av allmänt intresse kommer vi inte särskilt behandla myndighetsutövning. I princip gäller samma krav som när myndigheter agerar utanför ramen för deras myndighetsutövning med den skillnaden att kraven på rättsligt stöd är ännu starkare. Myndigheternas behandling kan även ha stöd i den rättsliga grunden "nödvändig för att fullgöra en rättslig förpliktelse" i artikel 6.1 c samt i vissa fall även en intresseavvägning enligt artikel 6.1 f.

Utöver kravet på att behandlingen ska vara nödvändig för en uppgift av allmänt intresse krävs enligt artikel 6.3 dataskyddsförordningen att den rättsliga grunden som fastställs i unionsrätten eller medlemsstaternas nationella rätt ska uppfylla ett mål av allmänt intresse och vara proportionerlig mot det legitima mål som eftersträvas. Kraven på nödvändighet, proportionalitet och stöd i nationell rätt eller unionsrätten är högre ju integritetskänsligare behandling det är fråga om. Särskilda krav i detta avseende ställs också enligt artikel 9 dataskyddsförordningen för behandling av känsliga personuppgifter.

Nödvändighets- och proportionalitetsbedömningen handlar i praktiken om en riskbedömning och intresseavvägning mellan riskerna för de registrerade i förhållande till syftet med behandlingen. Denna risk- och intressebedömning kan till viss del göras av lagstiftaren genom att ge myndigheterna ett tydligt lagstöd för en viss behandling. I avsaknaden av sådant tydligt lagstöd faller ett större ansvar på myndigheten att göra bedömningen om det ligger i myndighetens uppgift att ägna sig åt verksamhet som kräver att personuppgifter behandlas och i sådana fall väga in hur behandlingen ska göras på ett så integritetsvänligt sätt att fördelarna för uppgiften av allmänt intresse väger över integritetsriskerna.

Bedömningen kan vara komplicerad. En grundläggande förutsättning för en korrekt bedömning är dock en noggrann beskrivning av den planerade behandlingen av personuppgifter dels av intresset och behovet som motiverar personuppgiftsbehandlingen, dels av riskerna för de enskilda vars personuppgifter kommer att behandlas.

Sammanfattningsvis är det således ett antal krav som måste vara uppfyllda för att en myndighet ska kunna förlita sig på den rättsliga grunden ”uppgift av allmänt intresse” vid behandling av personuppgifter.

- Det måste för det första finnas en uppgift som myndigheten har i uppdrag att utföra och som avser *ett allmänt intresse*.
- För det andra måste grunden för uppgiften ha *stöd i unionsrätten eller medlemsstaternas nationella rätt* (legalitetsprincipen).
- För det tredje måste behandlingen av personuppgifter vara *nödvändig* för att myndigheten ska kunna utföra uppgiften av allmänt intresse (art. 6.1 e dataskyddsförordningen).
- För det fjärde måste behandlingen vara *proportionerlig* i förhållande till syftet med behandlingen.
- Därutöver tillkommer krav för behandling av vissa särskilda kategorier av personuppgifter (framför allt känsliga personuppgifter och uppgifter om lagöverträdelser enligt artiklarna 9 och 10) samt krav enligt de grundläggande principerna som gäller för all personuppgiftsbehandling enligt artikel 5.

I de följande avsnitten går vi igenom dessa krav och vad de innebär vid utveckling av en AI-tjänst.

3.3 Vad krävs för att utveckling av en AI-modell ska vara en *uppgift av allmänt intresse*?

Alla uppgifter som myndigheter utför är, enligt regeringen, av allmänt intresse om myndigheten har tilldelats uppgiften av riksdagen eller regeringen i en författning. Därutöver ger det kommunala självstyret kommuner och regioner en möjlighet att ägna sig åt många olika typer av verksamheter så länge som verksamheterna uppfyller kraven enligt kommunallagen. Även privata aktörer kan utföra uppgifter av allmänt intresse.[20] Uppgifter av allmänt intresse kan vara av allmänt ekonomiskt intresse såsom posttjänster, icke ekonomiska tjänster såsom polisverksamhet och sociala tjänster såsom socialtjänst.[21] Allmänhetens rätt att få tillgång till allmänna handlingar kan, enligt skäl 154 dataskyddsförordningen, betraktas som ett allmänt intresse.

Få myndigheter torde ha en uttalad uppgift att utveckla AI-tjänster. Däremot kan utvecklingen av en AI-tjänst vara ett led i att effektivisera myndighetens verksamhet. Att effektivisera sin verksamhet ingår i myndigheternas generella uppdrag att arbeta kostnadseffektivt.[22] Utveckling av en AI-tjänst bör i det avseendet kunna jämställas med annan verksamhets- och it-utveckling. Förutsatt att AI-tjänsten ska användas inom myndighetens verksamhet bör utvecklingsarbetet anses ingå i de uppgifter av allmänt intresse som myndigheten utför, liksom rättsliga förpliktelser och myndighetsutövning som ligger inom myndighetens mandat.

Myndigheternas utveckling och tillämpning av AI-tjänster bör mot den bakgrunden kunna innefattas i de generella uppgifter som myndigheterna har i uppdrag att utföra. En annan fråga är förstås om utveckling av en egen AI-modell är ett kostnadseffektivt sätt att fullgöra uppgiften av allmänt intresse, jämfört med att t.ex. använda eller anpassa en befintlig AI-modell som finns tillgänglig på marknaden och träna den för den egna verksamheten. Det är en bedömning som enligt dataskyddsförordningen kan vägas in i frågan om den eventuella personuppgiftsbehandling som krävs är ”nödvändig” (se avsnitt 3.5 nedan).

3.4 Vad krävs för att utveckling av en AI-tjänst ska ha *stöd i nationell rätt*?

Grunden för den uppgift av allmänt intresse som myndigheten har bedömt innefatta utveckling av en AI-tjänst måste enligt artikel 6.3 dataskyddsförordningen ha stöd i unionsrätten eller medlemsstaternas nationella rätt. I Sverige innebär detta att uppgiften ska följa av lag eller annan författning, av kollektivavtal, eller beslut som meddelats med stöd lag eller annan författning (2 kap. 2 § dataskyddslagen). Uppgiften behöver

[20] Prop. 2017/18:105 s. 56 f.

[21] EU-kommissionen https://commission.europa.eu/topics/single-market/services-general-interest_sv.

[22] Enligt 1 kap. 3 § budgetlagen (2011:203) ska hög effektivitet eftersträvas i statens verksamhet och god hushållning iakttas. Av 3 § myndighetsförordningen (2007:515) framgår att myndigheters verksamheter ska bedrivas effektivt och att myndigheter ska hushålla väl med statens medel. Kommuner och regioner ska enligt 11 kap. 1 § kommunallagen (2017:725) ha en god ekonomisk hushållning i sin verksamhet.

inte direkt framgå av t.ex. lagtext. Det kan vara tillräckligt att den framgår av förarbetena till lagen, bestämmelsens syfte och den rättsliga kontexten som bestämmelsen befinner sig i.[23]

Den rättsliga grunden enligt artikel 6.3 är ofta fastställd i myndigheters instruktioner eller i annan reglering som styr verksamheten t.ex. hälso- och sjukvårdslagen och socialtjänstlagen (2001:453). På vissa områden kompletteras verksamhetsregleringen av en särskild författning som reglerar personuppgiftsbehandlingen, en så kallad registerförfattning såsom lagen (2001:454) om behandling av personuppgifter inom socialtjänsten. Den rättsliga grunden som fastställs i nationell rätt enligt artikel 6.3 omfattar då både verksamhetsregleringen och registerförfattningen. För kommunala myndigheter kan den rättsliga grunden ha stöd i beslut från kommun- eller regionfullmäktige.

En viktig utgångspunkt är att ju integritetskänsligare behandling som ska utföras desto tydligare stöd för uppgiften och behandlingen krävs, i form av tydlighet, precision, förutsebarhet och proportionalitet i den rättsliga grunden. En bedömning måste göras från fall till fall utifrån behandlingens och verksamhetens art. En mindre ingripande behandling, såsom att hantera elevers namn i skolverksamhet, kan baseras på en mer generell rättslig grund. Däremot krävs en mer preciserad grund för känsligare behandlingar, såsom hantering av känsliga uppgifter inom hälso- och sjukvården, för att intrånget i den personliga integriteten ska vara förutsebart. Vid betydande intrång i den personliga integriteten, vid övervakning eller kartläggning av personliga förhållanden behövs dessutom särskilt lagstöd enligt 2 kap. 6 § andra stycket regeringsformen.

Det finns inget hinder mot att lagstiftaren ger en myndighet generellt formulerade uppgifter men det begränsar möjligheterna för myndigheten att utföra mer integritetskänslig personuppgiftsbehandling. Om lagstiftaren inte samtidigt anger gränserna för vilken behandling av personuppgifter som får utföras för myndighetens uppgifter är det upp till myndigheten att göra en bedömning av vad som ryms inom de uppgifter som tilldelats myndigheten. Bedömningen ska utgå från den rättsliga grundens tydlighet, precision och förutsebarhet.

Om lagstiftaren å andra sidan tydligt anger vad myndigheten får (eller ska) göra och preciserar att det kan innefatta integritetskänslig personuppgiftsbehandling har myndigheten mindre utrymme att behandla personuppgifter utöver detta. En sådan precisering av uppgiften anses i regel uttömmande och begränsar då utrymmet för myndigheten att tolka sin uppgift till att omfatta andra åtgärder än de som lagstiftaren har preciserat för den generellt utformade arbetsuppgiften. Därtill kommer bedömningen av om den aktuella behandlingen av personuppgifter är nödvändigt och proportionerlig, vilket vi återkommer till i avsnitt 3.5 och 3.6 nedan.

[23] Prop. 2017/18:105 s. 188.

IMY har haft en relativt strikt syn på den rättsliga grunden ”uppgift av allmänt intresse” för myndigheternas personuppgiftsbehandling. I ett flertal fall har IMY underkänt generellt formulerat lagstöd med hänvisning till integritetsriskerna:

- Närvarokontroll av gymnasieelever med hjälp av ansiktsigenkänning ansågs sakna uttryckligt lagstöd även om närvarokontroll i sig var en uppgift av allmänt intresse.[24]

- Stockholms stifts tillsyn av präster och diakoner i s.k. avkragningsärenden i enlighet med lagen (1998:1591) om Svenska kyrkan ansågs inte ha rättsligt stöd enligt artiklarna 6.3 och 9.2 g dataskyddsförordningen för att behandla känsliga personuppgifter.[25]

- Sveriges kommuner och regioners hantering av den s.k. väntetidsdatabasen till vilken regionerna har en skyldighet att rapportera uppgifter om väntetider ansåg IMY visserligen utgjorde en uppgift av allmänt intresse. Men IMY ansåg att grunden i 6 kap. 3 § sjukvårdsförordningen (2017:80) med hänsyn till den integritetskänsliga behandlingen inte uppfyllde kraven på tydlighet, precision och förutsägbarhet.[26]

- Utlämnande av patientuppgifter mellan vårdgivare i syfte att träna en AI-modell ansåg IMY inte ha stöd i patientdatalagens bestämmelse om kvalitetsutveckling och kvalitetssäkring eftersom den bestämmelsen gäller endast vårdgivarens egen verksamhet.[27]

- IMY har ansett att hantering av allmänna handlingar och sekretessprövning utgör uppgifter av allmänt intresse och att regleringen i TF och OSL *”i princip är tillräckligt tydlig och precis för att ligga till grund för personuppgiftsbehandling i digitala verktyg som underlättar hanteringen av begäranden av allmänna handlingar och sekretessbedömningar.”*[28]

När det kommer till personuppgiftsbehandling i samband med myndigheters kamerabevakningar tycks dock IMY ha godkänt mer generellt formulerade rättsliga grunder. I stället har det ofta varit kravet på nödvändighet och proportionalitet som IMY anser begränsar möjligheterna för myndigheter att kamerabevaka.

- Kamerabevakning i skolor har IMY ansett vara en uppgift av allmänt intresse som har stöd i skyldigheten att skapa en trygg skolmiljö enligt 5 kap. 3 § skollagen.

[24] IMYs beslut 2019-08-20, dnr DI-2019-2221. Beslutet överklagades men såväl förvaltningsrätten som kammarrätten avslog överklagandet, Högsta förvaltningsdomstolen lämnade inte prövningstillstånd.

[25] IMYs beslut 2019-03-26, dnr DI-2018-15886.

[26] IMYs beslut 2024-12-18, dnr DI-2021-9333.

[27] IMY, Federerad maskininlärning mellan två vårdgivare, Slutrapport om Integritetsskyddsmyndighetens pilotprojekt med regulatorisk testverksamhet om dataskydd, 2023-03-14, dnr IMY-2023-2602, s. 22.

[28] IMY, Utlämnande av allmänna handlingar med hjälp av AI, Slutrapporten från Integritetskyddsmyndighetens regulatoriska sandlåda om dataskydd, dnr IMY-2024-5156 (cit. IMY-rapporten), s. 33. Att allmänhetens rätt att få tillgång till allmänna handlingar är ett allmänt intresse framgår även av skäl 154 i dataskyddsförordningen.

Någon särskild reglering av personuppgiftsbehandlingen ansåg IMY inte nödvändig men fann att viss kamerabevakning var i strid med kravet på nödvändighet och proportionalitet.[29]

- Räddningstjänsten i Östra Skaraborgs kamerabevakning av utrymmen där anställda befinner sig bl.a. omklädningsrum ansågs vara nödvändig för en uppgift av allmänt intresse, trots en generellt formulerad bestämmelse om verksamheten utan särskilda bestämmelser för personuppgiftsbehandling.[30]

Sammanfattningsvis behöver en myndighet som avser att utveckla en AI-tjänst identifiera den rättsliga grunden för verksamheten och bedöma om den är tillräckligt tydlig och förutsebar i förhållande till de integritetsrisker som utvecklingsarbetet kan innebära.

3.5 När är utveckling av en AI-modell *nödvändig* för en uppgift av allmänt intresse?

Utöver att utvecklingen av en AI-modell ska innefattas i uppgift av allmänt intresse som ska ha stöd i nationell rätt krävs att behandlingen av personuppgifter som kommer utföras vid utvecklingen är *nödvändig* för att fullgöra uppgiften av allmänt intresse.

EU-domstolen har uttalat att nödvändighetskravet innebär att behandlingen faktiskt svarar mot de mål av allmänt intresse som eftersträvas, utan att gå utöver vad som är nödvändigt för att uppnå dessa mål. Detta krav på nödvändighet är inte uppfyllt om det eftersträvade målet av allmänt intresse rimligen kan uppnås på ett lika effektivt sätt genom andra medel som i mindre utsträckning inskränker de registrerades grundläggande rättigheter.[31] Regeringen har ansett att kravet på nödvändighet inte innebär att den aktuella behandlingsåtgärden ska vara oundgänglig. Det är tillräckligt att den leder till effektivitetsvinster. Att behandlingen skulle kunna ske manuellt, dvs. utan tekniska hjälpmedel, medför därför normalt inte att automatisk behandling inte anses nödvändig.[32] Vår bedömning är att effektivitetsvinst är en faktor bland flera som ska vägas in i den generella nödvändighets- och proportionalitetsbedömningen.[33]

Bedömningen av huruvida en behandling av personuppgifter är nödvändig är beroende av hur uppgiften är utformad i nationell rätt. Det är i regel svårare att hävda att en viss behandling är nödvändig om uppgiften är allmänt beskriven och kan tänkas uppfyllas på olika sätt med olika grad av personuppgiftsbehandling. Av rättssäkerhetsskäl krävs därför att mer integritetskänslig behandling har stöd i en tydligt formulerad uppgift i föreskrift.

[29] IMYs beslut 2023-10-03, dnr DI-2021-5774.

[30] IMYs beslut 2021-06-09, dnr DI-2018-22697.

[31] EU-domstolens dom den 4 oktober 2024 i mål C-200/23 "Agentsia po vpisvaniyata" p. 109-111.

[32] Prop. 2017/18:105 s. 189.

[33] Jfr EU-domstolen som bl.a. har ansett att publicering och utlämnande av personuppgifter om trafikförseelser (prickningar) inte var nödvändigt för att förbättra trafiksäkerheten med beaktan av att uppgifternas känslighet och att det fanns andra handlingsalternativ som med mindre ingripande medel lika effektivt kunde uppnå målet (bl.a. informationskampanjer, förnyade förarprov). Se EU-domstolens dom 2021-06-21 i mål C-439/19 "Latvijas Republikas Seima", p. 111-113.

Vidare är bedömning av nödvändigheten av den aktuella personuppgiftsbehandlingen ett första steg inför proportionalitetsbedömningen. Om man finner att behandlingen inte är nödvändig finns det dock ingen mening med att bedöma huruvida behandlingen är proportionerlig. Eller annorlunda uttryckt, en behandling som inte är nödvändig kan aldrig vara proportionerlig.

Om utvecklingen av en AI-tjänst lika bra kan genomföras utan att behandla personuppgifter t.ex. med användning av fiktiva personuppgifter är inte kravet på nödvändighet uppfyllt. Om exempelvis träning av AI blir effektivare med användning av riktiga personuppgifter kan dock nödvändighetskravet vara uppfyllt. Bedömningen av nödvändighet måste även göras på uppgiftsnivå. Behandlingen får således inte omfatta uppgiftskategorier som inte ger något mervärde t.ex. för träning av en AI-modell. Det följer av principen om uppgiftsminimering (artikel 5.1 c dataskyddsförordningen). Detsamma gäller omfattningen; träningen bör således inte utföras med flera personuppgifter (antal registrerade, antal uppgifter om respektive registrerad) än vad som krävs för att uppnå syftet med träningen.

3.6 När är behandlingen av personuppgifter för utveckling av en AI-tjänst *proportionerlig*?

Behandlingen av personuppgifter måste vara proportionerlig, dvs. de konsekvenser som behandlingen innebär för de registrerade måste stå i rimlig proportion till den nytta för det allmänna intresse som eftersträvas med behandlingen.[34] Det följer av EU:s rättighetsstadga men framgår även av artikel 6.3 dataskyddsförordningen och för svenska myndigheter av 5 § FL. Kravet på proportionalitet gäller både den nationella rätten (eller unionsrätten) som utgör grunden för uppgiften och den behandling som en myndighet planerar att utföra för att fullgöra uppgiften.

3.6.1 Proportionalitetsbedömning av nationell rätt

Utformningen av den nationella rätten som utgör grunden för uppgiften är en fråga för den nationella lagstiftaren, kommunfullmäktige eller beslutsfattaren som beslutar om en uppgift som ska utföras (jfr 2 kap. 2 § dataskyddslagen). Regeringen har ansett att den personuppgiftsansvarige inte behöver göra någon proportionalitetsbedömning avseende regleringen av den rättsliga grunden för behandlingen, utan kan utgå från att svensk rätt uppfyller krav på att nationell rätt ska vara proportionerlig mot det legitima mål som eftersträvas.[35] En annan sak är att en myndighet måste bedöma om den aktuella behandlingen ryms inom den rättsliga grund som lagstiftaren tillhandahållit (och är nödvändig).

[34] Proportionalitetsprincipen såsom den kommer till uttryck i EU-rätten är inte tydligt definierad, särskilt inte när det gäller vid behandling av personuppgifter. I vissa fall anses den innefatta samtliga krav enligt artikel 52.1 EU:s rättighetsstadga dvs. begränsningen ska vara fastställd i lag (kallas även legalitetsprincipen), nödvändig och proportionerlig i förhållande till det eftersträvade målet (proportionalitetsprincipen i strikt mening). Här hanterar vi proportionalitetsprincipen i strikt mening. Se European Data Protection Supervisor (EDPS), Guidelines on assessing the proportionality of measures that limit the fundamental rights to privacy and the protection of personal data, 19 December 2019, s. 10.

[35] Prop. 2017/18:105 s. 50.

3.6.2 Proportionalitetsbedömning av behandlingen

Den myndighet som utvecklar en AI-tjänst måste göra en proportionalitetsbedömning utifrån den aktuella behandlingen av personuppgifter, vad den innebär för de registrerade och vilken nytta den kan bidra med till den uppgift av allmänt intresse som myndigheten ska utföra. Som nämnts ovan, blir det inte aktuellt att göra en proportionalitetsbedömning om myndigheten redan har bedömt att målet med behandlingen kan uppnås på ett mindre integritetskänsligt sätt än utvecklingen av en AI-tjänst.

3.6.3 Proportionalitet handlar om att väga intressen mot varandra

En proportionalitetsbedömning är en intresseavvägning mellan de registrerades intresse bl.a. av skydd för sina personuppgifter och syftet med åtgärden. Fördelarna med en viss behandling ska vara så stora att de väger tyngre än de konsekvenser som behandlingen kan få för de registrerade. Det är således nödvändigt att bedöma riskerna med en viss behandling vilket förutsätter en tydlig kartläggning av vilken personuppgiftsbehandling som kan komma att utföras vid utvecklingen av en AI-tjänst. Vilka personuppgifter kommer att behandlas, vilka är de registrerade, vilka kommer att få tillgång till personuppgifterna, varifrån inhämtas uppgifter och till vilka kommer uppgifterna att spridas, är några av de mest centrala frågorna att besvara.

3.6.4 Det "allmänna intresset" i en AI-tjänst

Fördelar med en behandling av personuppgifter – dvs. det allmänna intresset som en AI-tjänst ska stödja – brukar normalt sett vara relativt enkelt att identifiera på ett abstrakt plan men kräver en detaljerad beskrivning och analys för att kunna läggas till grund för en proportionalitetsbedömning. Även om det handlar om personuppgiftsbehandling i samband med utveckling av en AI-tjänst måste man i proportionalitetsbedömningen utgå från de fördelar som tillämpningen av den färdigutvecklade AI-tjänsten medför. Som exempel på sådana fördelar kan enligt vår uppfattning följande nämnas:

- Effektivare ärendehandläggning som frigör tid och resurser för handläggare till mer kvalificerade uppgifter.
- Bättre möjligheter att avslöja otillåten eller brottslig verksamhet som riktas mot myndigheten både av tredje man, såsom bidragsfusk, och av anställda.
- Bättre analyser av insamlade data som ger bättre underlag för myndigheten att utföra sina uppgifter.
- Bättre service för medborgare genom kortare handläggningstid eller anpassad information vid kontakter med myndigheten.
- Fördelar på en samhällsnivå av andra grundläggande rättigheter och friheter såsom bättre insyn i det offentliga, tillgång till allmänna handlingar, identifiering av skadligt innehåll på nätet m.m.

- Rättssäkrare handläggning genom mer enhetlig handläggning av ärenden.
- Kostnadsfördelar, minskade kostnader för it och/eller personal.

3.6.5 De registrerades intressen vid utveckling av en AI-tjänst

Riskerna med en personuppgiftsbehandling är i första hand konsekvenser för enskilda, både immateriella och materiella konsekvenser som omfattas av de enskildas grundläggande friheter och rättigheter. Det innefattar även eventuella positiva konsekvenser för de registrerade vars personuppgifter behandlas under utvecklingsarbetet. Skyddet för de enskildas rättigheter och friheter omfattar inte endast faktiskt konstaterade konsekvenser utan även risker för sådana konsekvenser som med viss sannolikhet kan tänkas följa av en behandling av deras personuppgifter.[36]

Som exempel på sådana risker som kan beaktas vid myndigheters utveckling och användning av en AI-tjänst kan följande nämnas.

- Risken för diskriminering till följd av att en AI-tjänst ger ett diskriminerande svar på grund av att AI-modellen har tränats med ett icke representativt underlag.
- Risken för att felaktiga eller missvisande personuppgifter behandlas eller skapas i en AI-tjänst.
- Risken för att personuppgifter som behandlas sprids på ett obehörigt sätt t.ex. till personal som inte ska ha tillgång till dem, till myndigheter eller andra externa aktörer såsom leverantörer.
- Risken för identifiering av pseudonymiserade personuppgifter genom att analyser som utförs i en AI-tjänst urskiljer mönster som möjliggör identifiering genom s.k. urgallring (singling out).
- Risken för att analyser som utförs i en AI-tjänst innebär att enskilda kategoriseras på ett sätt som innebär att känsliga personuppgifter behandlas t.ex. uppgifter om etniskt ursprung eller religiös övertygelse.
- Risken för att analyser som utförs i en AI-tjänst innebär profilering av enskilda, eller anställda, bl.a. avseende personliga aspekter, arbetsprestationer, ekonomisk ställning, hälsa, personliga preferenser eller intressen, tillförlitlighet, beteende, vistelseort eller rörelsemönster.

Utöver risker och konsekvenser för enskilda måste en riskbedömning av en viss personuppgiftsbehandling även beakta omfattningen av behandlingen, dvs. antal registrerade som berörs av behandlingen, antal personuppgifter om respektive registrerad, vilken typ av personuppgifter som ska behandlas, hur länge behandlingen ska pågå (bevarande av uppgifterna), behandlingens geografiska omfattning och

[36] Se bl.a. skäl 75 dataskyddsförordningen.

begränsningar i andra rättigheter och friheter är rätten till skydd för personuppgifter. Dessa faktorer anses i sig vara riskfaktorer.[37]

3.6.6 En sammanvägd riskbedömning

I en sammanvägd riskbedömning ska de identifierade riskerna vägas in med beaktande av sannolikheten för att dessa risker realiseras och därmed leder till förutsedda konsekvenser.

Det finns gott om vägledningar för hur riskbedömningar vid personuppgiftsbehandling ska utföras. Utveckling av AI-tjänster som innefattar behandling av personuppgifter under utvecklingsfasen innebär många gånger sådana risker för de registrerade att det krävs att en konsekvensbedömning enligt artikel 35 dataskyddsförordningen upprättas. IMY har nyligen (februari 2025) publicerat en vägledning och mall för sådana konsekvensbedömningar.[38]

Vid utveckling av AI-tjänster är det viktigt att myndigheten har kompetens att kunna bedöma hur AI-tjänsten kommer att fungera. Det kan exempelvis vara lämpligt att ta hjälp av leverantörer för att få en tillräckligt tydlig förklaring hur personuppgifter kommer att behandlas och vilka risker som kan uppkomma vid användning av leverantörens tjänster.

3.7 När finns det rättsligt stöd för behandling av *känsliga personuppgifter* vid utveckling av en AI-tjänst?

3.7.1 Särskild proportionalitetsbedömning

Myndigheterna hanterar ofta stora mängder med personuppgifter som inte sällan omfattar känsliga personuppgifter, uppgifter om lagöverträdelser eller andra integritetskänsliga uppgifter. Dessa datamängder är ofta intressanta att använda som träningsdata för AI-modeller. Det kan därför vara svårt att undvika att behandla känsliga personuppgifter i samband med utveckling av en AI-tjänst. Att exempelvis träna en AI-modell med uppgifter som på ett korrekt sätt avspeglar realistiska förhållanden kan förutsätta att även känsliga personuppgifter behandlas. Om det inte är möjligt att anonymisera datamängden, exempelvis genom att använda fiktiva uppgifter måste det finnas ett tydligt stöd i lag eller annan författning för sådan behandling.

I grunden är det samma krav som gäller vid behandling av ”vanliga” personuppgifter som vid behandling av känsliga personuppgifter men med den skillnaden att kraven enligt legalitetsprincipen och proportionalitetsprincipen är högre. Det kan således krävas ett tydligare stöd i unionsrätten eller den nationella rätten och att det allmänna intresset väger ännu tyngre än riskerna med behandlingen. Behandling av känsliga personuppgifter är, som utgångspunkt, förbjuden enligt artikel 9.1 dataskyddsförordningen. Det krävs därför att något av undantagen i artikel 9.2 är

[37] Jämför EDPB ”balancing test” för den rättsliga grunden berättigat intresse i artikel 6.1 f dataskyddsförordningen. EDPB, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, Adopted on 17 December 2024, s. 23 ff.

[38] https://www.imy.se/publikationer/vagledning-vid-konsekvensbedomning/

tillämpliga. Flera av dessa undantag som kan användas av myndigheter förutsätter kompletterande lagstiftning i nationell rätt eller unionsrätten.

3.7.2 Stöd för behandling av känsliga personuppgifter i svensk rätt och unionsrätten

Särskilt stöd för känsliga personuppgifter finns vanligtvis i s.k. registerförfattningar såsom i 7 § lagen (2001:454) om behandling av personuppgifter inom socialtjänsten. Registerförfattningar innehåller vanligtvis begränsningar för hur känsliga personuppgifter och andra integritetskänsliga personuppgifter får behandlas, såsom bestämmelserna om sammanställningar av känsliga personuppgifter i 7 a och 7 b §§ i nämnda lag eller förbud mot sökning i syfte att få fram ett urval av personer grundat på känsliga personuppgifter i 26 a 4 § andra stycket skollagen (2010:800).

AI-förordningen innehåller ett särskilt stöd för att behandla känsliga personuppgifter för att undvika bias i samband med utveckling av AI-tjänster med hög risk, se artikel 10.5.[39] Bestämmelsen innehåller skyldigheter som i stor utsträckning är krav som redan följer av dataskyddsförordningen, bl.a. krav om att det ska vara ”absolut nödvändigt” att behandla känsliga personuppgifter och att införa skyddsåtgärder. Bestämmelsen bör därför kunna ge vägledning även vid utvecklingen av AI-tjänster som faller utanför definitionen i AI-förordningen av AI-tjänster med hög risk. Den kan dock inte utgöra rättsligt stöd för utveckling av AI-tjänster i sådana fall.

I 3 kap. 3 § dataskyddslagen finns ett generellt stöd för myndigheter att behandla känsliga personuppgifter i tre situationer; (i) om de känsliga personuppgifterna har lämnats till myndigheten och behandlingen krävs enligt lag, (ii) om behandlingen är nödvändig för handläggningen av ett ärende, samt (iii) i andra fall om det är nödvändigt med hänsyn till ett viktigt allmänt intresse och inte innebär ett otillbörligt intrång i den registrerades personliga integritet.[40] Av särskilt intresse för utveckling av en AI-tjänst för utlämnande av allmänna handlingar är den tredje situationen som är avsedd att tillämpas för s.k. faktiskt handlande (dvs. inte ärendehandläggning, jfr andra punkten). Att utveckla AI-tjänster torde i de flesta fallen utgöra faktiskt handlande. Vi ska därför i det följande avsnittet titta närmare på den bestämmelsen.

3.7.3 Undantagsregeln i 3 kap. 3 § första stycket 3

Det finns ingen definition eller närmare beskrivning av vad som utgör ett viktigt allmänt intresse eller hur ett sådant intresse skiljer sig från ”allmänt intresse”, varken i dataskyddsförordningen eller i svensk rätt. I samband med att bestämmelsen i 3 kap. 3 § dataskyddslagen infördes uttalade regeringen att det inte råder något tvivel om att det är ett viktigt allmänt intresse att myndigheterna kan sköta sitt uppdrag på ett korrekt, rättssäkert och effektivt sätt. Regeringen ansåg också att den grundlagsfästa rätten att ta

[39] Artikel 10.5. Europaparlamentet och Rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens och om ändring av förordningarna (EG) nr 300/2008, (EU) nr 167/2013, (EU) nr 168/2013, (EU) 2018/858, (EU) 2018/1139 och (EU) 2019/2144 samt direktiven 2014/90/EU, (EU) 2016/797 och (EU) 2020/1828 (förordning om artificiell intelligens)

[40] Bestämmelsen i 3 kap. 3 § dataskyddslagen bygger på, och hänvisar till, artikel 9.2 g dataskyddsförordningen som innehåller ett undantag från det generella förbudet att behandla känsliga personuppgifter när behandlingen är nödvändig för ett "viktigt allmänt intresse".

del av allmänna handlingar måste anses utgöra ett viktigt allmänt intresse och att bestämmelserna i bl.a. OSL, arkivlagstiftningen och förvaltningslagen innehåller sådana lämpliga och särskilda skyddsåtgärder som krävs enligt artikel 9.2 g dataskyddsförordningen som undantagen i 3 kap. 3 § dataskyddslagen baseras på.[41]

Mot den bakgrunden bör det mesta av vad myndigheter ägnar sig åt utgöra inte bara ett allmänt intresse utan även ett viktigt allmänt intresse, åtminstone enligt nämnda uttalanden i förarbetena till dataskyddslagen. Att behandla personuppgifter i samband med utvecklingen av en AI-tjänst i syfte att effektivisera eller annars förbättra en verksamhet som myndigheten är satt att utföra, kan därför enligt vår uppfattning anses utgöra ett viktigt allmänt intresse.

Inte heller innebörden av formuleringen ”otillbörligt intrång i den registrerades personliga integritet” ges någon tydlig förklaring i förarbetena och förekommer inte heller i dataskyddsförordningen. Regeringen anger att en proportionalitetsbedömning måste göras varvid vikt ska läggas vid bl.a. uppgifternas känslighet, behandlingens karaktär, den inställning de registrerade kan antas ha till behandlingen, den spridning uppgifterna kan komma att få och risken för vidarebehandling för andra ändamål än insamlingsändamålet. Detta innebär t.ex. att bestämmelsen inte ger stöd för att skapa integritetskänsliga sammanställningar av känsliga personuppgifter.[42] Det senare följer också av sökbegränsningen som har införts i andra stycket som innebär att vid behandling med stöd av bestämmelsen i första stycket är det förbjudet att utföra sökningar i syfte att få fram ett urval av personer grundat på känsliga personuppgifter.

IMY har bedömt att undantagsregeln i 3 kap. 3 § första stycket 3 inte ger stöd för behandling av känsliga personuppgifter i den s.k. helhetstjänsten. Motiveringen är att helhetstjänstens separata vektoriserade och indexerade databas kan utgöra en sådan integritetskänslig sammanställning som enligt förarbetena och andra stycket inte är tillåten med stöd av 3 kap. 3 § första stycket 3.[43] Se vidare avsnitt 4.2 nedan.

Vår bedömning är att bestämmelsen i 3 kap. 3 § första stycket 3 inte ska tolkas så restriktivt som IMY ger uttryck för. Av förarbetena till bestämmelsen framgår att det krävs en särskild och noggrann proportionalitetsbedömning när känsliga personuppgifter ska behandlas vid faktiskt handlande av myndigheterna. Vidare anges att bestämmelsen inte ska användas slentrianmässigt i myndigheternas löpande verksamhet och att den personuppgiftsansvarige myndigheten måste göra en bedömning i det enskilda fallet om behandlingen utgör ett otillbörligt intrång i den registrerades personliga integritet. Det betyder dock inte att en bedömning behöver göras i förhållande till varje berörd individ. Värt att notera är att regeringen valde att inte begränsa bestämmelsen till behandlingen av enstaka uppgifter såsom utredningen föreslog.[44]

[41] Prop. 2017/18:105 s. 85 f.
[42] Prop. 2017/18:105 s. 194 f.
[43] IMY-rapporten s. 22.
[44] Prop. 2017/18:105 s. 80 ff.

Uttalandet att bestämmelsen inte ger stöd för att skapa integritetskänsliga sammanställningar av känsliga personuppgifter måste ses mot bakgrund av den proportionalitetsbedömningen, inte som ett fristående krav. Av det följer att inte alla sammanställningar som innehåller känsliga personuppgifter i sig kan betraktas som integritetskänsliga, utan att det krävs någon särskild egenskap hos sammanställningen. Det är enligt vår uppfattning rimligt att behandling som specifikt underlättar sammanställningar utifrån individers känsliga personuppgifter (ett försteg till den sökning som förbjuds i andra stycket) är en sådan egenskap. Regeringen anger att sökbegränsningen i andra stycket ”*omfattar alla tekniska åtgärder som innebär att uppgifter används för att strukturera eller systematisera information i syfte att få fram ett urval av personer grundat på känsliga personuppgifter*”.

Vi bedömer mot den bakgrunden att vektorisering och indexering av handlingar i syfte att i ett senare skede kunna finna relevanta handlingar inte tar sikte specifikt på känsliga personuppgifter och inte heller utgör sökning efter sådana uppgifter, utan handlar om en systematisering av informationen på generell nivå, vilken i sig inte resulterar i en integritetskänslig sammanställning av just känsliga personuppgifter.

Vi tolkar förarbetsuttalandena så att behandlingen inte får ske utan en sådan särskild proportionalitetsbedömning som krävs vid behandling av känsliga personuppgifter enligt artikel 9.2 g dataskyddsförordningen. Att bestämmelsen inte ska läggas till grund för att skapa integritetskänsliga sammanställningar av känsliga personuppgifter innebär enligt vår bedömning inte hinder för att känsliga personuppgifter behandlas för att effektivisera myndigheternas hantering av allmänna handlingar.

Inte heller sökbegränsningen i andra stycket bör enligt vår mening innebära något hinder mot utveckling av AI-tjänster så länge det inte innebär att känsliga personuppgifter struktureras eller på annat sätt behandlas på grund av sin karaktär. Vi återkommer till riskbedömningen av helhetstjänsten i avsnitt 4.

3.8 Vilka *uppgifter* finns det rättsligt stöd för att behandla vid utveckling av en AI-tjänst?

Utgångspunkten är att fler personuppgifter än vad som är nödvändigt för att uppnå syftet (ändamålen) med en viss behandling får inte behandlas (uppgiftsminimering, artikel 5.1 c). Vidare gäller att de personuppgifter som används för ändamålet utveckling av en AI-tjänst måste antingen vara insamlade för detta ändamål eller för ett ändamål som inte är oförenligt med utvecklingsändamålet (ändamålsbegränsning, artikel 5.1 b).

Det är myndigheten som utvecklar en AI-tjänst som i egenskap av personuppgiftsansvarig måste göra bedömningen av vilka personuppgifter som är nödvändiga att behandla för utvecklingsarbetet. Om uppgifterna är insamlade inom ramen för en verksamhet som är särskilt reglerad i en registerförfattning sätter den regleringen ofta gränserna för vilka personuppgifter som får samlas in och hur insamlade personuppgifter får användas, t.ex. lagen (2001:454) om behandling av personuppgifter inom socialtjänsten.

I många registerförfattningar anges att personuppgifter som behandlas i sakverksamheten får behandlas även för andra ändamål än de som särskilt anges i författningen, under förutsättning att uppgifterna inte behandlas på ett sätt som är oförenligt med de ändamål för vilket uppgifterna samlades in. Det förekommer dock registerförfattningar som saknar sådan bestämmelse eller där uppräkningen av tillåtna ändamål är uttömmande. I sådana fall kan utrymmet för att behandla uppgifter för att utveckla en AI-tjänst vara begränsat, såvida inte de i registerförfattningen angivna ändamålen kan anses inbegripa sådan behandling. Det förekommer exempelvis att som tillåtna ändamål anges t.ex. uppföljning, utvärdering, kvalitetssäkring, vilket bör kunna innefatta utveckling av AI-tjänster.[45]

Registerförfattningar saknas i stor utsträckning för kommunal verksamhet. Ett undantag är lagen (2001:454) om behandling av personuppgifter inom socialtjänsten som inte innehåller något stadgande om behandling för icke oförenliga ändamål men tillåter behandling för bl.a. tillsyn, uppföljning, utvärdering, kvalitetssäkring och administration av verksamhet (6 § jämfört med 2 §).

3.8.1 Uppgifter som kan bli föremål för utlämnande enligt offentlighetsprincipen bör som utgångspunkt kunna behandlas

Personuppgifter som kan bli föremål för utlämnande enligt offentlighetsprincipen kan i regel anses vara insamlade för bl.a. detta ändamål, dvs. att uppfylla myndighetens skyldigheter att på begäran lämna ut allmänna handlingar enligt bestämmelserna i 2 kap. TF. Alternativt kan sådan behandling anses ett icke oförenligt ändamål.[46] Ändamålet bör innefatta all behandling som är nödvändig dels för skyldigheten att organisera allmänna handlingar (4 kap. 1 § OSL), dels för skyldigheten att hantera en begäran om utlämnande (6 kap. OSL). Ändamålet hantering av allmänna handlingar enligt 2 kap. TF gäller således vid sidan av de ändamål som myndigheten samlar in uppgifter om för sin sakverksamhet och behöver därför inte vara förenligt med andra ändamål.

Att en myndighet behandlar personuppgifter för att effektivisera hanteringen av allmänna handlingar med hjälp av AI bör, enligt vår mening, som utgångspunkt innefattas i ändamålet som följer av myndigheternas skyldighet att organisera och lämna ut allmänna handlingar enligt TF och OSL. I vart fall bör en sådan behandling inte anses oförenlig med det ursprungliga ändamålet och därmed tillåten enligt artikel 6.4 dataskyddsförordningen.

Principen om ändamålsbegränsning bör därför som utgångspunkt inte innebära någon begränsning av vilka personuppgifter som får behandlas för att utveckla en AI-tjänst som effektiviserar utlämnandet av allmänna handlingar så länge som uppgifterna är

[45] I ett förslag till ny studiestödsdatalag har det generella och huvudsakliga ändamålet kompletterats med en bestämmelse om att uppgifterna får behandlas som ett led i Centrala studiestödsnämndens stödverksamhet om det är nödvändigt för "att göra dataanalyser och urval i syfte att förbygga, förhindra och upptäcka fel". Avsikten är att myndigheten ska kunna utnyttja AI för sådana dataanalyser. Se SOU 2025:95 s. 104 ff.

[46] Regeringen har ansett en myndighets utlämnande av personuppgifter med stöd av offentlighetsprincipen inte anses vara oförenligt med de ändamål för vilka uppgifterna ursprungligen samlades in. Offentlighetsprincipen framgår av grundlagen och får anses utgöra en integrerad del av all förvaltningsverksamhet som myndigheterna ägnar sig åt, prop. 1997/98:44, s. 44.

sådana som kan bli föremål för utlämnande enligt 2 kap. TF och att myndigheten inte är förhindrad enligt lag eller annan förordning att göra det. Som framgått ovan måste dock behandlingen vara nödvändig och proportionerlig.

I de fall utvecklingen av en AI-tjänst riskerar att medföra mer kännbara konsekvenser för de personer vars personuppgifter behandlas under utvecklingsarbetet, bör det ses som ett nytt ändamål som kräver särskilt rättsligt stöd. En bedömning behöver göras enligt de kriterier som anges i artikel 6.4, vilket bl.a. innefattar en bedömning av uppgifternas art, konsekvenser för de registrerade och skyddsåtgärder.

Vid utveckling av en AI-tjänst är det sällan nödvändigt att identifiera enskilda även om uppgifterna som används gör det möjligt. Däremot kan det vara viktigt att identifiera information som just en personuppgift eller en viss kategori av personuppgift, utan dess koppling till en identifierbar individ. Pseudonymisering eller användning av fiktiv data kan därför vara ett sätt att minimera eventuella konsekvenser för de registrerade, vilket skulle tala för att ändamålet är förenligt med det ursprungliga.

Exempel. Danska Datatilsynet har i en vägledning tagit upp följande exempel.[47]
En kommun beslutar att utveckla en AI-lösning för att förbättra sin hantering av tillgång till dokumentförfrågningar vad gäller svarstid, kvalitet och konsekvens. Lösningen ska på ett effektivt sätt kunna söka i filer och dokument och identifiera information som behöver anonymiseras samt stödja ärendebehandlingen. AI-lösningen tränas med hjälp av data som härrör från tidigare ärenden om tillgång till dokumenten. Det innebär att kommunen vidarebehandlar personuppgifter som ursprungligen samlats in för ett ändamål (ärendebehandling av begäran om tillgång till handlingar) för ett nytt ändamål (utveckling av en AI-lösning). Kommunen ska därför bedöma om det nya ändamålet är förenligt med det ursprungliga. Enligt Datatilsynets uppfattning finns det ett naturligt samband mellan behandling av personuppgifter som ett led i behandlingen av begäranden om tillgång till handlingar och utvecklingen av ett verktyg som är avsett att stödja samma behandling. Dessutom utförs fortsatt behandling av samma myndighet som ursprungligen samlat in uppgifterna. Det nya syftet – utvecklingen av AI-lösningen – kan därför anses vara förenligt med det ursprungliga syftet.

3.8.2 Historisk data från tidigare sökningar och sekretessbedömningar

Tidigare manuellt utförda sökningar efter relevanta handlingar för utlämnande och manuellt utförda maskeringar för sekretessbedömning kan vara av stort intresse vid utveckling av AI-tjänster för utlämnande av allmänna handlingar. Sökningar och sammanställningar av tidigare relevanta handlingar torde inte innebära några särskilda integritetsproblem förutsatt att inte känsliga personuppgifter har använts som sökord.

Tidigare utförd maskering i allmänna handlingar, och då särskilt handlingarna i sin omaskerade form tillsammans med maskeringsinformationen, för sekretessbedömning kan däremot medföra särskilda integritetsrisker vid återanvändning för träning av en

[47] Danska Datatilsynet, Offentlige myndigheders brug af kunstig intelligens Inden I går i gang, Oktober 2023, s. 15 f.

AI-modell eftersom en handläggare tidigare har bedömt att uppgifterna omfattas av sekretess och vidare strukturerat informationen för att exakt avgränsa dessa uppgifter. Det kan förstås handla om information som ansetts omfattas av sekretess på andra grunder än integritetsskäl men oavsett sekretessgrund handlar om det om information som måste behandlas med särskild försiktighet.

Om det är möjligt att begränsa hanteringen av de sekretessmarkerade uppgifterna genom att ersätta det faktiska informationsinnehållet med angivande av vilken typ av information som har markerats t.ex. ”beskrivning av sjukdomshistoria” i stället för redogörelse för en person sjukdom, begränsas integritetsriskerna. Detsamma gäller om behandlingen som sker med hjälp av AI kan begränsas till sammanhanget där uppgifterna förekommer utan att tolkning av informationsinnehållet sker. I det senare fallet behöver AI möjligen inte ens tillgång till den sekretessmarkerade informationen. Det kan vara tillräckligt att den maskerade informationen typiskt sätt förekommer i ett fält eller i ett visst sammanhang, t.ex. efter en bestämd textfras, för att AI ska kunna föreslå att liknade text ska maskeras i andra dokument.

Vi utesluter inte att det kan finnas rättsligt stöd för att i klartext behandla känsliga personuppgifter och andra integritetskänsliga uppgifter från tidigare maskerade handlingar (se avsnitt 3.7 ovan), men det förutsätter att det inte finns någon annan mindre integritetskänslig metod som lika effektivt löser uppgiften (nödvändighet) och att integritetsriskerna inte väger tyngre än nyttan (proportionalitet).

3.8.3 Försiktighet bör iakttas om personuppgifter från externa källor inhämtas och vid samkörning

Andra personuppgifter än de som myndigheten själv har samlat in eller skapat i sin sakverksamhet bör endast med stor försiktighet behandlas i samband med utvecklingen av en AI-tjänst. Uppgifter från internet som samlas in med s.k. webscraping bör inte förekomma. Detsamma gäller om utvecklingsarbetet innebär att personuppgifter från flera register, databaser eller verksamhetsgrenar samkörs även om sådan samkörning inte uttryckligen är förbjuden, särskilt om det handlar om uppgifter från andra myndigheter. Även personuppgifter från s.k. publicitetsregister som t.ex. Bolagsverkets register bör användas med försiktighet.

4. AI-TJÄNSTER INOM PROJEKTET ”RÄTT TILL INSYN”

4.1 Maskeringstjänsten och helhetstjänsten

Inom projektet ”Rätt till insyn” har de deltagande kommunerna tillsammans med Atea undersökt möjligheterna att använda en AI-tjänst för att effektivisera hanteringen vid utlämnande av allmänna handlingar. Två olika lösningar har diskuterats och analyserats. I den mer omfattande **helhetstjänsten** var avsikten att automatisera stora delar av processen vid utlämnande av allmänna handlingar; från begäran av den enskilde till utlämnandet av den sekretessbedömda handlingen. En central del i helhetstjänsten är stödet för att, utifrån en enskild begäran identifiera och söka fram sådana allmänna handlingar i myndighetens diarier och andra informationssamlingar

som svarar mot begäran, den s.k. söktjänsten. Under arbetets gång har av olika anledningar fokus övergått till en mer begränsad tjänst som ska fungera som stöd för handläggare vid maskering av allmänna handlingar som ska lämnas ut, den s.k. **maskeringstjänsten**.

Maskeringstjänsten syftar således till att identifiera och ge förslag på uppgifter i på förhand identifierade handlingar som bör omfattas av sekretess och därför ska maskeras vid utlämnande. Maskeringstjänsten är alltså ett handläggarstöd, vilket innebär att det är en handläggare som i slutändan beslutar om vad som ska omfattas av sekretess och inte.

Även vid användningen av helhetstjänsten är det tänkt att en handläggare beslutar om sekretess utifrån det förslag som lämnas av tjänsten. Helhetstjänsten är således i denna del och i likhet med maskeringstjänsten, ett handläggarstöd. Skillnaden mellan tjänsterna avser övriga delar. Vid användningen av helhetstjänsten är även tolkning av begäran från den enskilde och sökning av relevanta dokument tänkt att automatiseras och ske utan inblandning av kommunens personal. Det finns dock inget som hindrar att även dessa delar omfattas av en manuell kontroll och därför fungerar som stöd för kommunens personal.

4.2 IMY:s rapport "Utlämnande av allmänna handlingar med hjälp av AI"

Några av de dataskyddsrättsliga frågor som uppkommer i samband med den typ av AI-tjänst som projektet utvecklar tillsammans med Atea har analyserats av IMY i en s.k. regulatorisk sandlåda. IMY har redovisat sina slutsatser i rapporten "Utlämnande av allmänna handlingar med hjälp av AI".[48] I rapporten har IMY valt att fokusera på tre frågor; finns det rättslig grund för behandlingen, hur fördelar sig personuppgiftsansvaret samt behovet av lämpliga säkerhetsåtgärder.

IMY har i rapporten gjort följande bedömningar av dessa frågor.

- Det kan finnas rättslig grund för behandling av personuppgifter i den mer begränsade maskeringstjänsten. Övervägande skäl talar dock emot att behandlingen av personuppgifter i helhetstjänsten är nödvändig och proportionerlig i dagsläget, särskilt vad gäller behandling av känsliga personuppgifter.

- Mycket talar för att personuppgiftsansvaret är avgränsat till den enskilda kommunen, medan AI-leverantören (Atea) agerar som personuppgiftsbiträde. Inom Lidingö kommun finns det omständigheter som talar för att var och en av kommunens nämnder är separat personuppgiftsansvariga för den behandling som sker när maskeringstjänsten används.

- En konsekvensbedömning avseende dataskydd bör göras innan maskeringstjänsten tas i bruk. IMY rekommenderar även ett antal säkerhetsåtgärder och betonar vikten

[48] Integritetsskyddsmyndigheten, Utlämnande av allmänna handlingar med hjälp av AI. Slutrapport från Integritetsskyddsmyndighetens regulatoriska sandlåda om dataskydd.

av ”human-in-the-loop” för att säkerställa att sekretessmarkeringar i behandlade dokument är korrekta innan handlingar lämnas ut.

I vår rapport fokuserar vi på frågan om rättsligt stöd för utveckling av AI-tjänster varför vi inte kommer att gå in närmare på frågan om personuppgiftsansvar och övriga frågor som IMY har bedömt.

4.2.1 Rättslig grund för maskeringstjänsten

IMY bedömer, som nämnts ovan, att maskeringstjänsten medför begränsade integritetsrisker och att den har rättsligt stöd i dataskyddsförordningen även om IMY tillägger att det krävs noggranna överväganden för att säkerställa att den aktuella personuppgiftsbehandlingen är nödvändig och proportionerlig.

Skälet för denna bedömning är främst att maskeringstjänsten endast ska fungera som ett handläggarstöd och att det inte finns något som tyder på att den skulle kunna leda till kvalitetsförluster i sekretessprövningen. De integritetsrisker som förekommer vid personuppgiftsbehandlingen vid användningen av maskeringstjänsten bedömer IMY som jämförliga med de risker som kan förekomma vid sekretessprövning som utförs av en handläggare ”med hjälp av ett manuellt digitalt maskeringsverktyg”.[49]

IMY bedömer mot den bakgrunden att mycket talar för att det finns stöd i den rättsliga grunden ”uppgift av allmänt intresse” enligt artikel 6.1 e för behandling som krävs för maskeringstjänsten och att behandlingen av känsliga personuppgifter som kan förekomma i maskeringstjänsten kan baseras på 3 kap. 3 § 3 dataskyddslagen.

Vi delar i stort IMY:s bedömning vad gäller maskeringstjänsten.

4.2.2 Rättslig grund för helhetstjänsten

Vad gäller helhetstjänsten bedömer IMY att övervägande skäl talar för att den personuppgiftsbehandling som sker i samband med användningen av helhetstjänsten i dagsläget saknar rättslig grund. Anledningen är att det enligt IMY finns betydande osäkerheter kring helhetstjänstens förenlighet med dataskyddsförordningens nödvändighets- och proportionalitetsprinciper och att det befintliga rättsliga stödet för behandlingen som finns i TF och OSL inte är tillräckligt förutsebart och preciserat i relation till de risker som finns för den personuppgiftsbehandling som sker i samband med användningen av helhetstjänsten.

IMY menar visserligen att det kan finnas ett stöd för helhetstjänsten i den rättsliga grunden uppgift av allmänt intresse i artikel 6.1 e dataskyddsförordningen men att det saknas stöd för att behandla känsliga personuppgifter. Bestämmelserna i TF och OSL (bl.a. myndigheternas skyldigheter att vidta åtgärder för att underlätta sökande efter allmänna handlingar och ge enskilda möjligheter att använda tekniska hjälpmedel för att kunna ta del av allmänna handlingar) är för allmänt hållna, inte tillräckligt

[49] Integritetsskyddsmyndigheten, Utlämnande av allmänna handlingar med hjälp av AI. Slutrapport från Integritetsskyddsmyndighetens regulatoriska sandlåda om dataskydd, s. 20.

förutsebara och preciserade i relation till de risker som finns med personuppgiftsbehandlingen som sker i samband med användningen av helhetstjänsten.

4.3 Helhetstjänsten – ett koncept för hur en helt automatiserad tjänst tekniskt skulle kunna utformas

Helhetstjänsten utgår från tanken om en helautomatiserad tjänst som gör det möjligt för enskilda att begära allmänna handlingar från en kommun utan att kommunens personal behöver vara inblandad. Projektet som för närvarande har valt att gå vidare med den mer begränsade maskeringstjänsten har inte närmare utvecklat konceptet med helhetstjänsten.

I stora drag kan en tjänst som automatiserar utlämnandet av allmänna handlingar delas upp i följande delar

- Begäran; AI-teknik hjälper den enskilde att formulera sin begäran och informerar om vilka informationsmängder som är tillgängliga för den automatiserade tjänsten.
- Sökning efter relevanta dokument (söktjänsten); AI-teknik används för att underlätta sökningen och för att tolka begäran och jämföra den med vad som sannolikt är relevanta handlingar eller uppgifter.
- Sekretessprövning och maskering (maskeringstjänsten); AI-teknik används för att avgöra om informationen i de framsökta handlingarna eller sammanställningen omfattas av sekretess och föreslår vilka uppgifter som ska maskeras.
- Utlämnande; handlingar eller uppgifter lämnas ut med eventuellt maskerade uppgifter.

Eftersom projektet har valt att inte gå vidare med helhetstjänsten finns det inte någon närmare beskrivning av hur tjänsten skulle kunna utformas. Det är därför svårt att göra en fullständig bedömning av den rättsliga grunden för personuppgiftsbehandlingen som skulle kunna förekomma vid användning av tjänsten. I det följande utgår vi därför från de risker som IMY har tagit upp i sin rapport.

4.4 Riskbedömning för helhetstjänsten

IMY lyfter fram ett antal riskfaktorer som ligger till grund för IMY:s bedömning av helhetstjänsten.[50] Det gäller främst risker förknippade med sökning och sammanställning av handlingar och uppgifter som är relevanta vid enskilds begäran om utlämnande av allmänna handlingar, dvs. det som vi i denna rapport kallar för söktjänsten. Nedan går vi igenom de riskfaktorer som IMY redovisat och lämnar vår egen bedömning.

[50] Integritetsskyddsmyndigheten, Utlämnande av allmänna handlingar med hjälp av AI. Slutrapport från Integritetsskyddsmyndighetens regulatoriska sandlåda om dataskydd, s. 21.

Helhetstjänsten förutsätter att kommunens diarier kopieras till en separat databas där informationen vektoriseras och indexeras.

Det är inte helt klart hur en automatiserad sökning efter relevanta handlingar och uppgifter ska utformas. Vår bedömning är dock att den omständigheten att handlingar och uppgifter kopieras till en separat databas inte i sig bör innebära några särskilda integritetsrisker förutsatt att behandlingen av personuppgifter i den separata databasen ligger inom ramen för vad som är tillåtet och omfattas av tillräckliga skydds- och säkerhetsåtgärder. Vi noterar att liknande behandling regelmässigt sker i myndigheternas system när dokument indexeras för att möjliggöra traditionell fulltextsökning.

Att informationen indexeras i detta sammanhang kan antas innebära att innehållet i ett dokument (s.k. färdig elektronisk handling med en informationsmängd som i förväg är avgränsad), beskrivs med indexord i syfte att underlätta sökning och sammanställning av dokumentet. Sådan indexering innebär i sig inte några särskilda integritetsproblem förutom risken att indexord i undantagsfall skulle kunna innebära en form av profilering och att indexorden skulle kunna underlätta sammanställning av personuppgifter som annars inte skulle vara möjlig.

Vektorisering[51] resulterar i ett informationsinnehåll som i viss utsträckning är anonymiserat. Det är dock inte avsikten i söktjänsten. Syftet är i stället att underlätta att hitta allmänna handlingar som är relevanta vid en begäran från en enskild. Det kan antas göra att flera handlingar blir mer lättillgängliga för utlämnande enligt offentlighetsprincipen vilket i och för sig ökar allmänhetens insyn men också ökar integritetsriskerna, se avsnitt 5.2 nedan).

Uppgifterna i diarierna måste bearbetas kontinuerligt för att kunna identifiera samband, utföra sannolikhetsanalyser och dra slutsatser; utan helhetstjänsten behandlas dessa uppgifter i stället främst genom statisk lagring i arkivsyfte.

Det är som sagt inte helt klart hur AI kan underlätta sökning efter relevanta dokument men vår bedömning är att det troligen inte innefattar att uppgifter i diarier och kommunens databas behöver behandlas kontinuerligt. Handlingar indexeras (på handlingsnivå) och innehållet kan eventuellt även komma att vektoriseras för att underlätta sökning för relevanta sammanställningar av uppgifter.

[51] Vektorisering av handlingar innebär att text m.m. omvandlas till numeriska representationer (vektorer) så att maskininlärningsmodeller kan bearbeta informationen. Det är en grundläggande teknik inom naturlig språkbehandling (NLP) och används för att analysera, jämföra och extrahera information från text t.ex. för sökning och klassificering av text, skapa semantiska representationer för förståelse och analys. Möjligheten att återskapa den ursprungliga informationen beror på exakt hur denna vektorisering går till, exempelvis vilken embedding-modell som används för att skapa vektorer och på vilken språklig granularitet (ord, meningar, stycken eller hela dokument) denna arbetar.

Det är svårt att förutse vad helhetstjänsten innebär för integritetsrisker i förhållande till så kallade potentiella handlingar.

Om söktjänsten som vi menar är en del av helhetstjänsten utformas så att tjänsten kan söka efter relevanta uppgifter i handlingar eller databaser t.ex. i epost-systemet i syfte att göra sammanställningar av uppgifter som begärs av den enskilde, s.k. potentiella handlingar enligt 2 kap. 6 § andra stycket TF, är det troligt att betydligt fler sådana potentiella handlingar blir tillgängliga för utlämnande. De kan medföra bättre insyn för allmänheten men även större integritetsrisker. Se vidare nedan avsnitt 5.2.

Andra tänkbara risker är att helhetstjänsten vid sökning missar relevanta handlingar eller inkluderar irrelevanta handlingar. Dessa risker kan dock mildras till exempel genom att man låter handläggare granska det som helhetstjänsten tar fram för utlämning. Men i så fall kan det dock frågas hur noggrann en sådan granskning behöver vara och om resultatet är effektivt och proportionerligt.

Det är troligen inte möjligt att utforma en AI-tjänst för sökning efter relevanta handlingar och uppgiftssamlingar som inte gör fel. Det är dock troligt att tjänsten med tiden kan förbättras med träning bl.a. med human-in-the-loop och utvecklingen av AI. Risken för att en AI-tjänst tar fram handlingar och uppgifter som inte är relevanta eller utelämnar handlingar och uppgifter som är relevanta vid en begäran finns även när en handläggare med hjälp elektroniskt diarium eller databaser utför sökningen. Risken bör därför jämföras med sådan ”manuell” sökning och det är inte uteslutet att en AI-tjänst ger bättre resultat. En lämplig lösning torde dock vara att AI-tjänsten ger förslag som kontrolleras av kommunens personal och tjänsten därmed fungerar som ett handläggarstöd. Se vidare avsnitt 4.5.

En risk med en tjänst som helhetslösning – men även med maskeringstjänsten – är att den innebär att känsliga personuppgifter måste identifieras i de allmänna handlingarna. Även om syftet är att eventuellt maskera dessa uppgifter kan denna behandling innebär en risk i sig. Det kan jämföras med den behandling av känsliga personuppgifter som kan behöva utföras för att upptäcka och förhindra bias i AI-modeller med hög risk enligt artikel 10.5 AI-förordningen. Som vi har nämnt tidigare kan de skyddsåtgärder som anges där beaktas vid utvecklingen av söktjänsten.

Helhetstjänsten innebär att AI kan komma att användas för att tolka en fråga från en enskild och hjälpa denne att utforma en begäran om allmänna handlingar, t.ex. i form av en chat. Även en sådan tjänst kan innebära att personuppgifter behandlas och att särskilda integritetsrisker uppstår. Det finns emellertid inte utrymme i denna rapport att utreda dem särskilt.

4.5 Hur kan skyddsåtgärder begränsa integritetsriskerna i helhetstjänsten?

Helhetstjänsten var inte i fokus för projektet som utvärderades av IMY och IMY:s bedömning var därför inte heller mer än översiktlig. IMY:s slutsats, som nämnts ovan, var att den rättsliga grunden för helhetstjänsten inte är tillräckligt förutsebar och preciserad i relation till integritetsriskerna med helhetstjänsten, särskilt vad avser behandling av känsliga personuppgifter i söktjänsten.

I förhållande till den mer begränsade maskeringstjänsten föreslår IMY ett antal säkerhetsåtgärder bl.a. för att motverka att handläggare som använder sig av maskeringstjänsten i för stor utsträckning förlitar sig på AI-modellens bedömningar. Bland de föreslagna säkerhetsåtgärderna som avser de särskilda riskerna vid användning av AI finns bl.a. åtgärder för[52]

- en tydlig styrning och ansvarsfördelning för användningen av AI,
- uppföljning av AI-tjänstens resultat,
- mänsklig kontroll (human-in-the-loop) och utbildning av användarna så att de är medvetna om AI-tjänstens förmågor och begränsningar för att undvika övertro till tjänstens kapacitet,
- begränsningar av AI-tjänstens funktionalitet så att den endast utför det som är relevant för syftet med tjänsten t.ex. genom att begränsa kommandon och instruktioner som användaren kan ge tjänsten,
- säkerställa att AI-modellen inte utför instruktioner som kan finnas i texter som AI-tjänsten är avsedda att hantera, samt
- minimera mängden data som AI-tjänsten får tillgång till.

Vi vill lyfta fram följande överväganden vid utvecklingen av en AI-tjänst såsom helhetstjänsten.

- Användning av fiktiv data och pseudonymiserade personuppgifter vid träning av AI-modellen.
- Begränsa AI-tjänsten till verksamhetsområden och informationsmängder där det i förväg går att bedöma vilka uppgifter, särskilt personuppgifter, som kan förekomma t.ex. vid mer standardiserade dokument.
- Säkerställ att AI-tjänsten inte kan göra sammanställningar, sökningar eller analyser med utgångspunkt i känsliga personuppgifter eller andra integritetskänsliga personuppgifter.

[52] Se vidare Integritetsskyddsmyndigheten, Utlämnande av allmänna handlingar med hjälp av AI. Slutrapport från Integritetsskyddsmyndighetens regulatoriska sandlåda om dataskydd, s. 33 f.

- Användning av tidigare maskerade dokument som har bedömts inte innebära några särskilda integritetsrisker eller användning av historisk data som har maskats för att undvika behandling av integritetskänsliga personuppgifter.

- Säkerställ att maskeringstjänsten inte använder AI-modellen för att generera en ny, maskerad, handling, utan endast för att generera instruktioner om vilka delar i den ursprungliga handlingen som ska maskeras. Maskeringstjänsten kan sedan använda traditionell databehandling för att utföra dessa instruktioner.

Därutöver kan de åtgärder som leverantörer och tillhandahållare av AI-system med hög risk är skyldiga att vidta enligt AI-förordningen, utgöra en god vägledning, se artiklarna 9-15 och 26-27 i AI-förordningen.[53] Av särskild betydelse har bestämmelserna om data och dataförvaltning (artikel 10), mänsklig kontroll (artikel 14) och konsekvensbedömning (artikel 27).

4.6 Finns det rättsligt stöd för helhetstjänsten?

Det finns inte utrymme inom ramen för denna rapport att göra en fullständig analys av under vilka förutsättningar en helhetslösning för utlämnande av allmänna handlingar kan vara förenlig med dataskyddsförordningen.

Vi gör dock följande bedömningar.

- Utveckling av en AI-tjänst i syfte att effektivisera utlämnande av allmänna handlingar är en uppgift av allmänt intresse och ligger inom ramen för en kommuns uppgift och skyldighet för detta ändamål (se artikel 6.1 e dataskyddsförordningen). På ett generellt plan bör därför behandlingen av personuppgifter som är nödvändig för ett sådant ändamål vara tillåten enligt dataskyddsförordningen. En förutsättning är dock att behandlingen inte innefattar behandling av några andra uppgifter än de som behandlas vid utlämnande av den allmänna handlingen utan en AI-tjänst och att behandlingen inte samtidigt sker för några andra ändamål.

- Att behandla personuppgifter för att effektivisera utlämnande av allmänna handlingar bör inte vara ett ändamål som anses oförenligt med ändamålet med offentlighetsprincipen såsom det kommer till uttryck i 2 kap. TF och 4 kap. 1 § OSL (se avsnitt 3.8).

- Det blir svårt att undvika att behandla känsliga personuppgifter och andra integritetskänsliga personuppgifter. Vi bedömer dock att utrymmet för att behandla känsliga personuppgifter enligt det generella undantaget i 3 kap. 3 § första stycket 1 dataskyddslagen i många fall kan vara tillräckligt vid utveckling av en AI-tjänst (se avsnitt 3.7 ovan). Det finns dock en osäkerhet i hur detta undantag ska tolkas och vi fördjupar oss därför i denna fråga i avsnitt 5.1 nedan.

[53] Vi har i denna rapport inte utrett tillämpningen av AI-förordning på myndigheter som utvecklar AI-tjänster och huruvida sådana tjänster kan vara förbjudna eller utgör "AI-system med hög risk".

- Integritetsriskerna med en helhetstjänst måste identifieras och noga analyseras. Skyddsåtgärder måste vidtas för att minimera de identifierade integritetsriskerna. En konsekvensbedömning enligt artikel 35 dataskyddsförordningen behöver med största sannolikhet upprättas.

- Den främsta integritetsrisken med helhetstjänsten är troligen att den kan komma att öka tillgången till allmänna handlingar, vilken visserligen ökar allmänhetens insyn men samtidigt gör det svårt att bedöma hur personuppgifter i allmänna handlingar kommer kunna sammanställas, analyseras och lämnas ut. Vi fördjupar oss i denna risk i avsnitt 5.2 nedan.

5. ÖVERGRIPANDE OBSERVERATIONER

5.1 Myndigheterna behöver tydligare rättsligt stöd för utveckling av AI-tjänster

Vi har bedömt att det finns ett utrymme enligt dataskyddsförordningen för myndigheter att utveckla AI-tjänster men att det krävs en noggrann bedömning av nödvändighet och proportionalitet. Det är många gånger en komplicerad bedömning av riskerna med AI för de personer vars uppgifter behandlas jämfört med fördelarna för myndighetens verksamhet. Vid *utvecklingen* av AI är sällan de enskildas identitet av intresse. Integritetsriskerna bör därför som utgångspunkt vara mer begränsade och ofta lägre än vid *tillämpningen* av AI-tjänster.

Myndigheternas personuppgiftsbehandling förutsätter stöd i unionsrätten eller nationell rätt genom t.ex. angivande av ändamålen för vilka personuppgifter som får behandlas. Ju integritetskänsligare behandling desto tydligare rättsligt stöd krävs. Vi kan konstatera att den befintliga regleringen för behandling av känsliga personuppgifter i dataskyddslagen inte ger önskvärd tydlighet även om förarbetena, enligt vår uppfattning, ger ett visst stöd för utveckling av AI-tjänster för utlämnande av allmänna handlingar.

För att underlätta utvecklingen av AI-tjänster inom den offentliga sektorn menar vi att det kan krävas förtydliganden i lag, vilka underlättar myndigheternas bedömningar och sätter gränserna för hur personuppgifter får behandlas. Det kan exempelvis handla om förtydliganden av eller kompletterande ändamålsbestämmelser, översyn av sökbegränsningar i registerförfattningar eller en justering av bestämmelsen i 3 kap. 3 § 3 dataskyddslagen.

Det huvudsakliga problemet ligger i rättsligt stöd för att behandla känsliga personuppgifter. Det kan därför vara ett alternativ att komplettera bestämmelserna i 3 kap. dataskyddslagen med en bestämmelse som ger ett generellt stöd för myndigheternas behandling av känsliga personuppgifter vid utveckling av AI-tjänster. En sådan bestämmelse måste förenas med flera begränsningar och förutsättningar samt krav på vissa skyddsåtgärder. En förebild kan finnas i bestämmelsen i artikel 10.5 AI-förordningen som ger rättsligt stöd för att behandla känsliga personuppgifter för att säkerställa upptäckt och korrigering av bias. Vi har redan konstaterat att den bestämmelsen kan ge vägledning för bedömningen av rättsligt stöd för behandling av

känsliga personuppgifter i samband med utvecklingen av AI-tjänster (se p. 0 och 0 ovan). Genom att införa en liknande bestämmelse i svensk rätt som avser utveckling av AI-tjänster kan det rättsliga stödet för myndigheternas behandling av känsliga personuppgifter klargöras.

En bestämmelse som ger tydligare stöd för att behandla känsliga personuppgifter skulle exempelvis kunna innehålla följande förutsättningar och begränsningar.

- Myndigheten ska visa att utvecklingen inte kan ske utan känsliga personuppgifter t.ex. med syntetiska eller fingerade uppgifter.
- Behandlingen får endast omfatta personuppgifter som myndigheten redan behandlar i sin sakverksamhet och får inte vara oförenlig med de ändamålen som uppgifterna har samlats in för eller förbjuden enligt annan lag eller författning.
- De känsliga personuppgifterna som används under utvecklingsfasen får inte lämnas ut eller göras tillgängliga för någon utanför myndigheten; möjligen med undantag till andra myndigheter som samverkar i utvecklingsarbetet och i sådana fall med vissa begränsningar.
- De känsliga personuppgifterna får endast användas under utvecklingsfasen och ska omfattas av tekniska begränsningar som förhindrar vidareutnyttjande.
- Behandlingen måste omfattas av särskilda tekniska och organisatoriska skyddsåtgärder för att minimera risken för missbruk och obehörig åtkomst t.ex. behörighetskontrollsystem.

Det finns självfallet nackdelar med en bestämmelse som har syftet att ge ett generellt stöd för samtliga myndigheter att behandla känsliga personuppgifter vid utveckling av AI-tjänster. En nackdel är att det är svårt att förutse vilken behandling som blir tillåten och därmed vilka integritetsrisker som bestämmelsen kan medföra. En annan nackdel är att bestämmelsen kan behöva innehålla så många begränsningar och förutsättningar att den blir för restriktiv och att dess tillämpningsområde blir för begränsat. Dessutom kan begränsningar och förutsättningar som en sådan bestämmelse måste innehålla i sin tur medföra tolknings- och tillämpningssvårigheter. Alternativet är att se över befintliga bestämmelser som styr respektive myndighets behandling av känsliga personuppgifter t.ex. i registerförfattningar. Dessvärre saknas många gånger sådan reglering för kommunal verksamhet.

5.2 AI-tjänster kan medföra att offentlighetsprincipen utökas väsentligt

5.2.1 Ökad insyn och tillgång kan ge större integritetsrisker

Syftet med att utveckla och använda en AI-tjänst vid utlämnande av allmänna handlingar är bl.a. att underlätta och effektivisera sökningen efter relevanta handlingar och eventuellt även uppgifter som ingår i handlingarna. Det innebär snabbare handläggning och minskad arbetsbörda för den utlämnande myndigheten. Det är rimligt att anta att en sådan AI-tjänst även kommer att medföra större möjligheter att ta

fram och sammanställa informationsmängder innehållande personuppgifter än som vad annars skulle vara möjligt. Allmänhetens insyn ökar därmed liksom myndighetens möjligheter att använda samma AI-tjänst för internt bruk. Baksidan är att detta kan medföra större integritetsrisker i och med att personuppgifter i myndighetens diarier och databaser kan komma användas på ett sätt som inte går att förutse och spridas i större omfattning än tidigare.

Det går visserligen att till viss del hantera problem med att en AI-tjänst ”överlevererar” t.ex. genom att begränsa vilka handlingar som AI-tjänsten ska ha tillgång till men svårigheterna uppkommer i förhållande till s.k. potentiella allmänna handlingar som kan innebära sammanställningar av uppgifter och analyser av samband mellan uppgifter som inte går att förutse. Det är också en av de integritetsrisker som IMY tar upp i sin rapport om Utlämnande av allmänna handlingar med hjälp av AI vid bedömning av helhetstjänsten.[54] Vi ska därför nedan titta närmare på den frågan.

5.2.2 Potentiella handlingar

Offentlighetsprincipen enligt 2 kap. TF omfattar inte endast s.k. färdiga elektroniska handlingar med ett låst informationsinnehåll utan även sammanställningar av uppgifter i s.k. potentiella handlingar.[55] De potentiella handlingarna är således handlingar som inte existerar i slutlig form vid tiden för sökandens begäran utan kräver att uppgifter sammanställs med tekniska hjälpmedel. Det kan t.ex. handla om sammanställningar av uppgifter som tas fram ur register och databaser med hjälp av olika sökbegrepp.

En utgångspunkt enligt offentlighetsprincipen är att enskilda ska ha samma tillgång till allmänna handlingar som myndigheten själv har, den s.k. likställighetsprincipen. Men enskildas rätt till potentiella handlingar går även längre än den rätt som myndigheten själv har att söka och sammanställa allmänna handlingar. Enskildas rätt att få tillgång till potentiella handlingar gäller nämligen oavsett om myndigheten själv har behov av den begärda sammanställningen i sin egen verksamhet. Myndighetens skyldigheter avseende utlämnande av potentiella handlingar är dock begränsade i följande avseenden.

- Rätten till potentiella handlingar gäller endast under förutsättning att myndigheten kan ta fram sådana handlingar med ”rutinbetonade åtgärder” enligt 2 kap. 6 § andra stycket TF. Högsta förvaltningsdomstolen har i ett fall ansett att en arbetsinsats på 4-6 timmar inte utgjorde rutinbetonade åtgärder.[56]

- Om sammanställningen, dvs. den potentiella handlingen, innehåller personuppgifter begränsas rätten till potentiella handlingar enligt den s.k. begränsningsregeln i 2 kap. 7 § TF till att endast gälla sammanställningar som

[54] Integritetsskyddsmyndigheten, Utlämnande av allmänna handlingar med hjälp av AI. Slutrapport från Integritetsskyddsmyndighetens regulatoriska sandlåda om dataskydd, s. 23.

[55] En upptagning är enligt 2 kap. 3 § TF en handling som endast med tekniska hjälpmedel kan läsas eller avlyssnas eller uppfattas på annat sätt.

[56] HFD 2015 ref. 25.

myndigheten får göra enligt lag eller förordning. Begränsningar av myndighetens rätt att göra sammanställningar finns oftast i s.k. registerförfattningar i form av sökbegränsningar. Det krävs dock att det är förbjudet för myndigheten att söka eller sammanställa vissa uppgifter t.ex. känsliga personuppgifter. Bestämmelser som anger för vilka ändamål en myndighet får behandla personuppgifter anses inte utgöra en begränsning.[57]

- Myndigheter är inte heller skyldiga att lämna ut allmänna handlingar i elektronisk form, varken färdiga elektroniska handlingar eller potentiella handlingar. Det följer av det s.k. utskriftsundantaget i 2 kap. 16 § TF. Men det finns inte heller något förbud att lämna ut handlingar elektroniskt om det inte uttryckligen är förbjudet eller om utlämnandet inte i sig är oförenligt med dataskyddsförordningen. Idag sker troligen de flesta utlämnanden av allmänna handlingar elektroniskt.

- Sekretessbestämmelser kan begränsa tillgången till allmänna handlingar. Av särskild betydelse i detta sammanhang är bestämmelsen om s.k. dataskyddssekretess i 21 kap. 7 § OSL som innebär att sekretess råder om det kan antas att personuppgifter i allmänna handlingar kommer att behandlas i strid med bl.a. dataskyddsförordningen efter utlämnandet.

5.2.3 Färdiga elektroniska handlingar

Även tillgången till s.k. färdiga elektroniska handlingar kan komma att öka genom användning av AI-tjänster. Tillgången till färdiga elektroniska handlingar är inte begränsad på samma sätt som tillgången till potentiella handlingar är. Begränsningen som följer av att myndigheten inte behöver lägga ned mer än rutinbetonade åtgärder gäller inte, 2 kap. 6 § andra stycket TF. Inte heller gäller begränsningsregeln enligt 2 kap. 7 § TF. En myndighet kan därför vara skyldig att lägga ned omfattande arbete för att ta fram och lämna ut färdiga elektroniska handlingar även om myndigheten själv inte har rätt att göra sådana sökningar. Det är tillräckligt att handlingarna är tillgängliga för myndigheten genom ett tekniskt verktyg som myndigheten själv använder i sådan form att den kan läsas eller avlyssnas eller uppfattas på annat sätt (2 kap. 6 § första stycket TF).

5.2.4 Integritetsriskerna vid utlämnande av allmänna handlingar

AI innebär väsentligt utökade möjligheter att söka, sammanställa och analysera stora informationsmängder och t.ex. hitta samband som annars endast skulle vara möjliga att finna med stora manuella resurser. Det är också svårt att förutse vilka resultat som en AI-modell kan leverera genom att hantera uppgifter i allmänna handlingar.

Likställighetsprincipen och begränsningsreglerna i 2 kap. TF medför att enskilda kan använda AI-tjänster som myndigheter utvecklar på ett sätt som är svårt att förutse. Det beror bl.a. på att en AI-tjänst troligen gör det lättare att ta fram sammanställningar som ligger inom gränsen för ”rutinbetonade åtgärder”. En begäran om tillgång till allmänna

[57] Prop. 2007/08:160 s. 69 f. och prop. 2014/15:148 s.53.

handlingar behöver inte heller begränsas till de ändamål som myndigheten använder dem för. Om det inte krävs omfattande åtgärder kan således en enskild begära att myndigheten använder en av myndigheten utvecklad AI-tjänst för att göra integritetskänsliga sammanställningar och analyser som myndigheten själv inte får genomföra p.g.a. ändamålsbestämmelser. Det finns, som nämnts ovan, endast förbud mot vissa sökningar som begränsar vad myndigheten är skyldig att göra vid en sådan begäran. Vidare kan den tekniska utformningen av en AI-tjänst begränsa vilka sökningar som myndigheten har möjlighet och skyldighet att utföra för att tillmötesgå en begäran om utlämnande av allmänna handlingar.

Sökbegränsningar som gäller för en myndighet – antingen i form av förbud mot vissa sökningar eller p.g.a. tekniska begränsningar – bör därför inarbetas i en AI-modell som myndigheten använder för att dels säkerställa att myndigheten inte utför några otillåtna sökningar i sin egen sakverksamhet, dels att begränsa integritetsriskerna vid utlämnande av allmänna handlingar.

5.2.5 Komplicerade sekretessbedömningar

Resultatet av införande av en AI-tjänst för utlämnande av allmänna handlingar kan medföra att den information som kan tas fram med hjälp av en AI-tjänst och begärs ut kan vara mycket integritetskänslig. Det kan kräva att myndigheten behöver genomföra mer komplicerade sekretessbedömningar bl.a. enligt bestämmelsen om s.k. dataskyddssekretess i 21 kap. 7 § OSL. Det gäller förstås även sammanställningar som tas fram med hjälp av en AI-tjänst och som kan omfattas av sekretess på andra grunder än skydd för personuppgifter.

5.2.6 Integritetsriskerna med offentlighetsprincipen behöver beaktas vid utvecklingen av AI-tjänster

Mot denna bakgrund bör integritetsriskerna som kan följa av kombinationen av myndigheternas användning av AI och offentlighetsprincipen beaktas i utvecklingsarbetet. Det kan exempelvis innebära följande överväganden.

- Begränsa om möjligt AI-tjänsten till färdiga elektroniska handlingar, dvs. sådana handlingar som har ett låst innehåll.
- Sökbegränsningar som gäller enligt lag ska byggas in i en AI-modell och säkerställa att icke tillåtna sökningar genomförs av AI-modellen.
- Begränsningar av sökningar och sammanställningar som kan medföra integritetskänsliga resultat och som inte behövs för myndighetens ändamål bör byggas in i AI-modellen även om de inte är förbjudna enligt lag.

5.2.7 Bestämmelserna i 2 kap. TF kan behöva ses över

Integritetsriskerna med offentlighetsprincipen och vilka följderna av utvecklingen av ny teknik har diskuterats sedan digitaliseringen inom den offentliga sektorn tog fart på 1980-talet. E-offentlighetsutredningen avfärdade år 2010 exempelvis en generell rätt att

få ut allmänna handlingar i elektronisk form av integritetsskäl och säkerhetsaspekter. Av samma skäl kan införande av AI för sökning och sammanställningar av uppgifter som finns lagrade i myndigheternas handlingar och databaser framstå som tveksam. E-offentlighetsutredningen konstaterade bl.a. följande.[58]

"Möjligen förhåller det sig så, att viss information som i dag är offentlig kanske inte borde vara det om det införs en generell och oinskränkt skyldighet att lämna ut allmänna handlingar i elektronisk form. Befintliga regler om vad som kan hållas hemligt är inte alltid utformade med beaktande fullt ut av den tekniska utvecklingen. Information som i dagsläget inte omfattas av sekretess skulle kanske framstå som mer känslig om den får en alltför stor allmän spridning, vilket skulle kunna bli fallet med en lagfäst rätt att få ut allmänna handlingar i elektronisk form. […] En sådan lagfäst rätt skulle också få konsekvenser för de menbedömningar som görs i syfte att avgöra om uppgifter ska sekretessbeläggas eller inte. Plötsligt är det inte längre tillräckligt att bedöma den skada som uppstår om en uppgift i sig själv röjs. Det skulle även krävas en svårbedömd uppskattning av den skada en uppgift kan orsaka i relation till andra uppgifter. "

Samhälls- och teknikutvecklingen inklusive AI har troligen gjort denna slutsats mer aktuell idag än 2010, bl.a. med hänsyn till utvecklingen av kommersiella tjänster som tillhandahåller omfattande databaser med allmänna handlingar som har hämtats ut från myndigheter och domstolar med stöd av offentlighetsprincipen.

Ökad insyn i myndigheternas verksamhet ger dock många fördelar. Den ger enskilda och medier ökade möjligheter att granska myndigheternas arbete bl.a. hur de hanterar personuppgifter, vilket skulle kunna gagna integritetsskyddet. Den kan också innebära att enskilda bidrar till att myndigheternas verksamhet t.ex. finner lösningar och samband som kan vara till fördel för myndigheterna.

Kombinationen av en omfattande offentlighetsprincip och ett starkt grundlagsskydd för publicering har lett till en konflikt mellan de svenska grundlagarna och integritetsskyddet enligt Europakonventionen och EU-rätten.[59] AI-utvecklingen kommer troligen driva på denna konflikt och eventuellt rubba den intresseavvägning mellan insyn och integritetsskydd, vilken finns i det nuvarande regelverket. De befintliga bestämmelserna är inte anpassade för de förmågor som AI för med sig. Det kan därför finnas ett behov av att göra en översyn av bestämmelserna i TF, OSL och registerförfattningar.

---

[58] SOU 2010:4 s. 271.

[59] Se bl.a. Högsta domstolens beslut 25 februari 2025 i mål Ä 3457-24 och Ä 3169-24 och EU-domstolens domar i mål C-740/22 ("Endemol Shine") och C-439/19 ("Latvijas Republikas Saeima").