##### Risikokatalog med afhjælpende foranstaltninger

**Juridisk værktøjskasse for Ansvarlig AI**

**Værktøj nr. 3**

Risikokatalog over risici vedrørende databeskyttelse i AI-løsninger – til brug for udarbejdelse af konsekvensanalyse vedrørende databeskyttelse

 

**1. Formål og anvendelsesområde**

Dette risikokatalog indeholder et overblik over en række typiske risici vedrørende databeskyttelse, der er forbundet med behandling af personoplysninger i AI-løsninger. Der er tale om risici for de *registrerede* og ikke for kommunen, f.eks. tab af omdømme eller bøder m.v. Risikokataloget indeholder også et forslag til, hvilke tekniske og organisatoriske foranstaltninger der kan være relevante at implementere for at håndtere de pågældende risici.

Formålet med risikokataloget er at give vejledning til kommuner, der skal lave en konsekvensanalyse vedrørende databeskyttelse i forbindelse med behandling af personoplysninger i en AI-løsning (**Værktøj nr. 2**). Kataloget kan bruges som inspiration til at identificere typiske risici vedrørende databeskyttelse i AI-løsninger og give forslag til tilhørende tekniske og organisatoriske afhjælpende foranstaltninger til at håndtere disse risici, så disse kan dokumenteres i konsekvensanalysen og der kan opnås et acceptabelt risikoniveau, inden AI-løsningen idriftsættes.

Risikokataloget har særligt fokus på typiske risici i AI-løsninger, der involverer automatisk beslutningstagen – enten i form af fuldautomatiske afgørelser eller beslutningsstøtte. Men kataloget kan også anvendes til andre former for brug af AI, f.eks. til modeller til snydbekæmpelse og profilering til brug for målrettet vejledning af borgerne m.v.

Risiciene og de tilhørende afhjælpende foranstaltninger er ikke udtømmende, og valg af risici og foranstaltninger skal tilpasses til de konkrete omstændigheder. Ikke alle risici og afhjælpende foranstaltninger er relevante i alle projekter, systemer m.v., og valget heraf vil bero på den konkrete konsekvensanalyse samt den specifikke teknologi, der er tale om, samt organisationens kompetencer, modenhed m.v. Risikokataloget er således til inspiration og vejledning – ikke en facitliste.

**2. Risikokatalogets risikotemaer**

Risikokataloget er overordnet opdelt i følgende risikotemaer vedrørende databeskyttelse og AI:

*Figur 1. Risikokatalogets risikotemaer*



**3. Risikokatalogets forslag til afhjælpende foranstaltninger/kontroller**

For hver enkelt risiko gives der i risikokataloget forslag til, hvordan man kan håndtere denne risiko, dvs. hvilke foranstaltninger eller kontroller, man med fordel kan overveje at implementere for at håndtere risikoen.

Der sondres overordnet mellem organisatoriske og tekniske foranstaltninger. De organisatoriske foranstaltninger omfatter f.eks. undervisning af medarbejdere og awarenessaktiviteter, udarbejdelse af kontroller, revision og audits, udarbejdelse af interne retningslinjer og procedurer vedrørende ansvarsfordeling og ledelsesgodkendelse m.v. De tekniske foranstaltninger vedrører f.eks. logning, pseudonymisering, tests og forskellige tekniske metoder til at gøre beslutninger baseret på AI transparente og forklarlige (såkaldt ”Explainable AI” eller ”XAI”).

Foranstaltningerne kan have et forskelligt fokus, alt efter om de tager sigte på (1) forebyggelse og beskyttelse af, at risikoen ikke realiseres, (2) opdagelse af, at risikoen er blevet realiseret eller (3) respons, dvs. handling når det er et konstateret, at risikoen er realiseret, f.eks. at man har opdaget en fejl i algoritmen.

*Figur 2. Fokus for afhjælpende foranstaltninger/kontroller*



**4. Skema med risici fordelt på risikotemaer samt relevante afhjælpende foranstaltninger/kontroller**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Nr.** | **Risiko** | **Forklaring af risikoen** | **Retsgrundlag** | **Relevante afhjælpende foranstaltninger/kontroller** | |
| **Type** | **Beskrivelse og fokus for tiltag** |
| **1. Rimelighed og forebyggelse af ulovlig forskelsbehandling**  **AI-løsningens ukorrekte output eller afgørelser fører til ulovlig forskelsbehandling af de registrerede (se også punkt 5 vedrørende datakvalitet)** | | | | | |
| 1.1 | Risiko for, at anvendelsen af systemet vil medføre ulovlig forskelsbehandling på grund af indlejrede skævheder og fordomme (bias). | Hvis systemets datasæt (både udvikling og drift) er præget af historisk (utilsigtet) skævhed, vil systemets beslutningsprocesser med høj sandsynlighed også medføre utilsigtet direkte eller indirekte ulovlig forskelsbehandling. Selvom rådata (inputdata) registreres korrekt, kan datasættet muligvis ikke være fuldt repræsentativt, eller analysen af dataene kan indeholde skjulte skævheder. Ulovlig forskelsbehandling kan også opstå som følge af en forkert udvikling og programmering af algoritmer. | Artikel 5, stk. 1, litra a | Organisatoriske foranstaltninger | * Gennemgående træning af AI-systemetmed forskellige træningsdata, inden systemet anvendes i praksis. (Forebyggelse og beskyttelse) * Sikring af repræsentative træningsdata. Test for brug af modellen på specifikke populationer eller udsatte grupper. (Forebyggelse og beskyttelse og Opdagelse) * Uddannelse af AI-udviklere og nøglepersoner, så de kan identificere og håndtere ulovlig forskelsbehandling i AI-systemer. (Forebyggelse og beskyttelse) * Overvejelse og beskrivelse af, hvordan ulovlig forskelsbehandling kan ske på det pågældende område, eventuelt med en definition af ”skævhed” og ”retfærdighed”, som anvendes under udvikling og anvendelse af systemet. Etablering af klare og målbare kriterier til at måle og teste de anvendte definitioner. (Forebyggelse og beskyttelse) * Regelmæssige kvalitetssikringskontroller og tests af AI-systemet for at sikre, at de registrerede behandles rimeligt og ikke forskelsbehandles, enten på grundlag af følsomme personoplysninger eller andet grundlag. (Forebyggelse og beskyttelse) * Indførelse af test og procedurer for overvågning af potentielle skævheder under udvikling og anvendelse af systemet. (Forebyggelse og beskyttelse og Opdagelse) * Hvis træningsdataene afspejler tidligere ulovlig forskelsbehandling, kan man udskifte træningsdataene og træne modellen på ny. (Respons) |
| 1.2 | Risiko for manglende hensyntagen til målgruppen ved behandling af personoplysninger om sårbare registrerede, herunder børn og handikappede m.v. | Når den dataansvarlige behandler personoplysninger om sårbare registrerede, herunder børn og handikappede m.v., vil der på grund af den øgede skævhed i magtfordelingen mellem den registrerede og den dataansvarlige være øget risiko for, at enkeltpersoner kan være ude af stand til på en nem måde at overskue eller modsætte sig behandlingen af deres oplysninger eller udøve deres rettigheder. | Artikel 5 og artikel 12-14 | Organisatoriske foranstaltninger | * I forbindelse med opfyldelse af oplysningspligten eller ved anden form for kommunikation har den dataansvarlige særligt fokus på, at oplysninger om behandlingen til sårbare registrerede skal være fyldestgørende, letforståelige og lettilgængelige under hensyntagen til denne gruppe. (Forebyggelse og beskyttelse) |
| **2. Transparens, forståelighed og systemtransparens**  **Det er ikke muligt fuldt ud at forklare AI-systemets output, dvs. hvordan systemet er nået frem til en afgørelse eller forslag hertil – beslutningsprocessen eller logikken i systemet er en ”black box”** | | | | | |
| 2.1 | Risiko for manglende forklarlighed og transparens i forhold til systemets beslutningsprocesser og datagrundlag. | Det er et krav, at den dataansvarlige skal kunne forklare, hvordan systemet når frem til en beslutning (output), dvs. mellemregningerne eller logikken i systemet samt datagrundlaget. Både de registrerede, der er adressat for en beslutning, medarbejdere hos den dataansvarlige selv samt tilsynsmyndigheder m.v. skal efter omstændighederne kunne få indsigt heri.  Hvis den dataansvarlige ikke er i stand til at forklare, hvordan systemet virker og kvaliteten af systemets output samt rationalet bag dette, skaber det risiko for, at fejl ved systemet ikke opdages eller at outputtet ikke kan kommunikeres til de registrerede og andre interessenter. | Artikel 5, stk. 1, litra a, artikel 13-15 og artikel 22 | Organisatoriske foranstaltninger | * Sikring af, at modellen kan forklares, herunder hvordan den når frem til (forslag til) beslutninger, logikken og de kriterier/data den lægger vægt på. Det skal sikres, at et menneske (1) kan forudsige, hvordan systemets output vil ændre sig, hvis input ændrer sig, (2) kan identificere de mest vigtige inputvariable og deres vægt og betydning for afgørelsens resultat, (3) identificere, hvornår outputtet sandsynligvis er forkert. (Forebyggelse og beskyttelse) * Dokumentation af datasættene og de processer, der fører til AI-systemets beslutninger, herunder træningsdata, testdata og algoritmer. Hvis der er tale om et regelbaseret system, bør programmeringsmetoden eller den måde, systemet er bygget på, dokumenteres. Hvis der er tale om et læringsbaseret AI-system, bør der ske dokumentation af metoden til oplæring af algoritmen, herunder hvilke inputdata der er indsamlet og udvalgt, og hvordan dette er sket. (Forebyggelse og beskyttelse) * Dokumentation af testmetoder og resultater af disse. (Forebyggelse og beskyttelse) * Sikring af, at enkeltpersoner, der er adressat for en fuldautomatisk afgørelse, kan få en individuel forklaring på grundlaget for og logikken bag afgørelsen. (Forebyggelse og beskyttelse) |
| Tekniske foranstaltninger | * AI-systemet designes, så det understøtter transparens og ikke udgør en uforklarlig "black box”. Forklarlighed udgør et grundprincip for valg af systemets matematiske model (algoritme), f.eks. ved brug af beslutningstræer i stedet for et neuralt netværk etc. (Forebyggelse og beskyttelse) * Brug af værktøjer og teknologier i relevant omfang til at understøtte forklarlighed (Explainable AI). (Forebyggelse og beskyttelse) |
| 2.2 | Risiko for, at registrerede ikke kender til AI-systemets karakter, kompetencer og begrænsninger. | Det udgør en risiko for brugerne af systemet, hvis de ikke har den information, det kræves for at forstå og interagere med AI-systemet i en tilfredsstillende grad, herunder vurdere systemets begrænsninger. | Artikel 5, stk. 1, litra a | Tekniske foranstaltninger | * Klar kommunikation af AI-systemets kapaciteter og begrænsninger til slutbrugerne. Dette kan omfatte en beskrivelse af AI-systemets nøjagtighed og dets begrænsninger. (Forebyggelse og beskyttelse) |
| **3. Formålsbegrænsning**  **Personoplysninger viderebehandles til modeludvikling eller anvendelse af systemet på en måde, der er uforenelig med de formål, personoplysningerne blev indsamlet til, i strid med princippet om formålsbegrænsning** | | | | | |
| 3.1 | Risiko for, at personoplysninger viderebehandles til modeludvikling eller anvendelse af systemet i strid med princippet om formålsbegrænsning. | Der er typisk videre rammer for at bruge personoplysninger indsamlet til forskellige formål til udvikling af en AI-model, når grundlaget for behandlingen er statistik. Omvendt er der snævrere rammer for brug af personoplysninger i nye sammenhænge – f.eks. på tværs af kommunale forvaltninger – når personoplysningerne skal anvendes til brug for at træffe afgørelser overfor de registrerede. Der er derfor risiko for, at man i strid med formålsbestemthedsprincippet behandler personoplysninger ved anvendelsen af den udviklede AI-model. | Artikel 5, stk. 1, litra b | Organisatoriske foranstaltninger | * Klar opdeling af behandling af personoplysninger i *udvikling* og *anvendelse* af modellen/systemet. (Forebyggelse og beskyttelse) * Dokumenterbare retningslinjer og procedurer for overholdelse af formålsbegrænsningsprincippet. (Forebyggelse og beskyttelse) |
| 3.2 | Risiko for viderebehandling af personoplysninger til andre formål end forskning og statistik (utilstrækkelig formålsbestemthed). | Når personoplysninger bliver indsamlet udelukkende med henblik på forskning og statistik efter de særlige bestemmelser herom i databeskyttelseslovens § 10, må oplysningerne ikke efterfølgende anvendes – herunder videregives – til andre formål end forskning og statistik, medmindre resultaterne er anonymiserede i sådan en grad, at de ikke kan føres tilbage til enkelte individer. | Artikel 5, stk. 1, litra b, og databeskyttelseslovens § 10, stk. 2 | Organisatoriske foranstaltninger | * Implementering af effektiv data governance og skarp opdeling af data i administrative og statistiske/forskningsmæssige formål. (Forebyggelse og beskyttelse) * Udarbejdelse af retningslinjer og procedurer for adgangskontrol og at personoplysninger alene må anvendes til forskningsmæssige og statistiske formål, og at alle resultater skal anonymiseres effektivt. (Forebyggelse og beskyttelse) * Overholdelse af kravene til videregivelse i bekendtgørelse nr. 1509 af 18. december 2019 om videregivelse af personoplysninger omfattet af databeskyttelseslovens § 10, stk. 1, og 2. (Forebyggelse og beskyttelse) |
| **4. Dataminimering og opbevaringsbegrænsning**  **AI-systemet udvikles, uden at der sker en tilstrækkelig vurdering af, om behandlingen af personoplysningerne til udvikling eller anvendelse af systemet er begrænset til det nødvendige, relevante og tilstrækkelige. Personoplysningerne desuden opbevaret i et længere tidsrum end nødvendigt i lyset af behandlingens formål. Dette medfører risiko for overtrædelse af princippet om dataminimering og princippet om opbevaringsbegrænsning.** | | | | | |
| 4.1 | Risiko for behandling af et unødvendigt omfang af personoplysninger til udvikling og anvendelse af systemet i strid med dataminimeringsprincippet. | Hvis AI-systemet udvikles, uden at der sker en tilstrækkelig vurdering af, om behandlingen af personoplysningerne til udvikling eller anvendelse af systemet er begrænset til det nødvendige, relevante og tilstrækkelige, medfører dette en risiko for overtrædelse af princippet om dataminimering. | Artikel 5, stk. 1, litra c | Organisatoriske foranstaltninger | * Overvejelse og dokumentation af, om brug af fiktive, anonymiserede og aggregerede oplysninger er tilstrækkelig til oplæring, modeludvikling og test. (Forebyggelse og beskyttelse) * Overvejelse og dokumentation af, om hver enkelt oplysningstype er relevant i hver fase af modeludviklingen, herunder ved inddragelse af personer med domænekendskab. (Forebyggelse og beskyttelse) * Brug af læringskurver til at dokumentere, hvornår yderligere kriterier ikke forbedrer modellens prædiktionsevne og præcision. (Forebyggelse og beskyttelse) * Brug af pseudonymisering af personoplysninger, herunder klare retningslinjer for, hvornår pseudonymisering er et krav. (Forebyggelse og beskyttelse) * Udarbejdelse og implementering af dokumenterede retningslinjer og procedurer for regelmæssig periodisk evaluering af proportionaliteten i databehandlingen med henblik på at eliminere datakilder, der ikke længere er nødvendige i lyset af formålet. (Forebyggelse og beskyttelse) * Risikoen forsøges reduceret ved at udarbejde og implementere retningslinjer i projektet for regelmæssig periodisk evaluering af proportionaliteten i databehandlingen med henblik på at eliminere datakilder, der ikke længere er nødvendige i lyset af formålet. * Brug af en uafhængig audit af modellen med henblik på at konstatere, om kriterierne (fortsat) er relevante. Brug af test med andre systemer med færre kriterier for at konstatere, om samme resultater kan opnås med brug af færre kriterier. (Opdagelse) * Sletning af kriterier, der ikke er relevante. (Respons) * Sletning af træningsdata, der ikke længere er relevante. (Respons) * Brug af en model med færre kriterier og grundig dokumentation af begrundelsen herfor. (Respons) |
| 4.2 | Risiko for, at personoplysninger opbevares i længere tid end formålet tilsiger (dataophobning). | Hvis den dataansvarlige indsamler personoplysninger til et bestemt formål, men fortsætter med at opbevare oplysningerne efter formålets ophør, udgør dette en unødvendig risiko for krænkelse af den registrerede. | Artikel 5, stk. 1, litra c og e | Organisatoriske foranstaltninger | * Udarbejdelse af slettepolitik for personoplysninger og dokumentation for overholdelse heraf. (Forebyggelse og beskyttelse) * Sletning af trænings- og testdata, der ikke længere er relevante. (Respons) * Gennemførelse af opfølgning på, at sletning faktisk sker i overensstemmelse med de fastsatte slettefrister. (Forebyggelse og beskyttelse og Opdagelse) |
| Tekniske foranstaltninger | * Implementering af automatiske sletningsprocedurer, der aktiveres ved formålets ophør. (Forebyggelse og beskyttelse og Respons) * Oprettelse af ”tombstones” eller sletteflag, så oplysninger slettes ved batchkørsler på foruddefinerede tidspunkter (hvis muligt). (Respons) |
| 4.3 | Risiko for uforholdsmæssig behandling i forbindelse med samkøring i kontroløjemed eller profilering. | Da samkøring i kontroløjemed per definition udgør en indgribende behandlingssituation, skærper en sådan behandling opmærksomheden på kravene i bl.a. andet forordningens artikel 5 om principper for behandling af personoplysninger, herunder dataminimeringsprincippet. Tilsvarende gælder efter omstændighederne profilering. Det skal sikres, at der ikke samkøres eller behandles flere oplysninger end nødvendigt i lyset af formålet. | Artikel 5, stk. 1, litra c | Organisatoriske foranstaltninger | * Fokus på og dokumentation af proportionalitetsafvejning, herunder navnlig i forhold til (1) om der skal foretages samkøring i kontroløjemed/profilering, (2) hvor mange oplysningstyper, der skal indgå i samkøringen/profileringen, samt (3) i hvilket omfang samkøringen/profileringen er forventelig for de registrerede. (Forebyggelse og beskyttelse) * Se om profilering tillige nedenfor i punkt 7. |
| **5. Datakvalitet**  **Fejl i modellen eller i inputdata fører til forkerte beslutninger (output) (se også punkt 1 om rimelighed og forebyggelse af ulovlig forskelsbehandling)** | | | | | |
| 5.1 | Risiko for fejl i modellen (algoritmen) i AI-løsningen fører til forkerte afgørelser m.v. | Hvis den udviklede AI-model er fejlbehæftet, vil resultaterne fra modellen (output) ligeledes være fejlbehæftet. Dette medfører, at sagsbehandlernes beslutningsgrundlag er ukorrekt, og der vil således være en risiko for, at den registrerede får en forkert afgørelse, rådgivning, service eller ydelse m.v.  Hvis AI-modellen f.eks. trænes til at forudsige et bestemt resultat, kan der være risiko for, at modellen ikke løbende trænes eller opdateres til at tage hensyn til kriterier med høj signifikans, som endnu ikke er kortlagt. Resultatet vil være, at modellen ikke har lige så stor forklaringskraft og prædiktionsevne, som den kunne have haft. | Artikel 5, stk. 1, litra a og d, og artikel 22, stk. 1 | Organisatoriske foranstaltninger | * Udarbejdelse af klar dokumentation og begrundelse for formålet med modellen, valg af model og de kriterier og deres vægtning, som indgår i modellen. (Forebyggelse og beskyttelse) * Beskrivelse af modellens begrænsninger i lyset af formålet, så alle relevante interessenter er opmærksom på, under hvilke omstændigheder den bør bruges. (Forebyggelse og beskyttelse) * Sikring af, at denne modelspecifikation godkendes af den relevante ledelse. (Forebyggelse og beskyttelse) * Fastsættelse af målbare kriterier og terminologi for modellens performance og nøjagtighed, som den skal leve op til og evaluering på, om det sker tilfredsstillende. Det skal fremgå, hvordan nøjagtighed måles og kontrolleres. (Forebyggelse og beskyttelse) * Udarbejdelse af beskrivelse og vurdering af mulige fejl og deres konsekvenser for de registrerede. (Forebyggelse og beskyttelse) * Test og revision af modellen og dens præcision og data før go live samt med et passende interval efter go live eller ved ændringer. Gennemførte tests, resultater heraf samt foranstaltninger truffet i medfør heraf skal kunne dokumenteres. (Forebyggelse og beskyttelse) * Gentræning/-kalibrering af modellen eller valg af en anden og bedre egnet model og grundig dokumentation for ændringen. (Respons) * Hvis der skal foretages centrale afvejninger – f.eks. mellem datakvalitet og dataminimering – skal ledelsen inddrages. (Forebyggelse og beskyttelse) * Træning og gentræning af AI-udviklere for at sikre, at de er opdateret på seneste viden. (Forebyggelse og beskyttelse og Respons) * Overvejelse af, om konstaterede fejl giver anledning til genoptagelse af sager. (Respons) * Dokumenterbare retningslinjer og processer for kravstillelse og due diligence for eksterne leverandører af AI-løsninger. Krav om, at løsningen skal overholde på forhånd fastsatte krav om statistisk sikkerhed og krav om regelmæssig testning. (Forebyggelse og beskyttelse) |
| 5.2 | Risiko for behandling af ukorrekte, utilstrækkelige og ikke ajourførte data fører til forkerte afgørelser (utilstrækkelig rigtighed). | Ukorrekte eller ikke ajourførte inputdata kan føre til, at AI-løsningens resultater er ukorrekte og at den registrerede dermed får en forkert rådgivning, service eller ydelse m.v. (”Garbage in – garbage out”). Det gælder både ved træning og anvendelse af modellen. Tilsvarende gælder ukorrekte, utilstrækkelige eller ikke ajourførte udledte data. | Artikel 5, stk. 1, litra d | Organisatoriske foranstaltninger | * Etablering af en beskrivelse af håndtering af forskellige datatyper, herunder en beskrivelse af, hvordan alle personoplysninger, der anvendes til træning, test og anvendelse af AI-systemet, er så korrekte, relevante, repræsentative, komplette og ajourførte som muligt. (Forebyggelse og beskyttelse). * Indførelse af foranstaltninger med det formål at sikre, at de anvendte data er fuldstændige, ajourførte og korrekte, samt om der er behov for yderligere data, f.eks. for at forbedre nøjagtigheden eller eliminere fejl. (Forebyggelse og beskyttelse) * Klar information til de registrerede om de personoplysninger, der behandles, og hvordan AI-systemet fungerer, således at de registrerede kan rette eventuelle unøjagtige oplysninger og forbedre datakvaliteten. (Forebyggelse og beskyttelse, Opdagelse og Respons) |
| **6. De registreredes rettigheder**  **AI-systemet understøtter ikke i tilstrækkeligt omfang de registreredes rettigheder.** | | | | | |
| 6.1 | Risiko for manglende muligheder for effektiv rettighedsudøvelse for de registrerede. | Personoplysninger kan indgå i (1) træningsdata ved udvikling af systemet, (2) testdata ved test af systemet, (3) data i selve modellen, (4) inputdata ved anvendelse af systemet samt (5) outputdata ved anvendelse af systemet, f.eks. en afgørelse eller forudsigelse. De registreredes rettigheder skal kunne håndteres i forhold til alle disse anvendelser af personoplysninger. | Artikel 15-22 | Organisatoriske foranstaltninger | * Implementering af dokumenterbare retningslinjer og procedurer for håndtering af de registreredes rettigheder ved modeludvikling og -anvendelse. (Forebyggelse og beskyttelse) * Træning af medarbejdere i identifikation og håndtering af rettighedsanmodninger. (Forebyggelse og beskyttelse) |
| 6.2 | Risiko for manglende eller utilstrækkelig opfyldelse af oplysningspligten. | De registrerede skal som udgangspunkt underrettes om den dataansvarliges behandling af deres personoplysninger. Det gælder som udgangspunkt også ved træning/udvikling og anvendelse af en AI-model. Oplysningspligten gælder som udgangspunkt også, hvis den dataansvarlige ønsker at viderebehandle personoplysningerne til et andet formål end det, hvortil de er indsamlet. | Artikel 12-14 | Organisatoriske foranstaltninger | * Udvælgelse og implementering af metode og retningslinjer til opfyldelse af oplysningspligten – både ved indsamling og hvis formålet med behandlingen skifter. (Forebyggelse og beskyttelse) * Implementering af arbejdsgange for dokumentation af oplysningspligtens overholdelse.(Forebyggelse og beskyttelse) |
| 6.3 | Risiko for manglende rettidig opfyldelse af retten til at gøre indsigelse med behandlingen. | Den registrerede kan efter omstændighederne have ret til at gøre indsigelse mod behandling af sine personoplysninger, herunder ved profilering baseret på disse bestemmelser. Senest på tidspunktet for den første kommunikation med den registrerede skal denne udtrykkeligt gøres opmærksom på indsigelsesretten, og oplysninger herom skal meddeles klart og adskilt fra alle andre oplysninger. | Artikel 21 | Organisatoriske foranstaltninger | * Samme som for punkt 6.2 ovenfor. |
| 6.4 | Risiko for manglende evne til at forklare logikken bag fuldautomatiske afgørelser. | Hvis AI-løsningen indebærer fuldautomatiske afgørelser, herunder ved brug af profilering, skal den registreredes oplyses herom samt modtage meningsfulde oplysninger om logikken heri samt betydningen og de forventede konsekvenser af en sådan behandling for den registrerede. Den registrerede har også ret til indsigt i disse oplysninger. | Artikel 13-15 og artikel 22 | Organisatoriske foranstaltninger | * Samme som punkt 6.2 ovenfor. * Se også foranstaltninger under punkt 2. Transparens, forståelighed og systemtransparens. |
| **7. Fuldautomatiske afgørelser og profilering**  **Risiko for manglende menneskelig kontrol og overvågning af systemets afgørelser og risiko for de facto fuldautomatiske afgørelser i beslutningsstøttesystemer** | | | | | |
| 7.1 | Risiko for manglende menneskelig kontrol og overvågning af AI-systemets afgørelser (output). | AI-systemer, der udfører fuldautomatiske afgørelser, har særlige risici for manglende menneskelig kontrol og overvågning af, om systemets afgørelser (output) er korrekte. Det gælder i særlig grad ved dynamiske modeller, der ændrer sig i takt med, at de modtager ny erfaring (data) fra sager. | Artikel 5 og artikel 22 | Organisatoriske foranstaltninger | * Løbende overvågning og mulighed for menneskelig overvågning og indgriben før/efter AI-løsningen har truffet en afgørelse. (Forebyggelse og beskyttelse og Respons) * Ved den manuelle kontrol med outputtet af systemet er der adgang til relevant viden om, hvorvidt den forudsatte kvalitet af outputtet er tilstrækkelig og i overensstemmelse med fastsatte kvalitetskrav, f.eks. at der er en kvalitetsscore af outputtet tilgængelig for brugeren, der kan forklares. * Udførelse af stikprøver af de automatiske afgørelser, således at et menneske kan vurdere, om afgørelserne er korrekte. (Opdagelse) * Overvejelse af genoptagelse af sager, hvis afgørelserne ikke er korrekte. (Respons) |
| 7.2 | Risiko for manglende kontinuitet, hvis AI-systemet skal stoppes. | Der kan være risiko for, at den forretningsmæssige opgave ikke kan løses, hvis AI-systemet af en eller anden grund – typisk på grund af en fejl eller en sikkerhedshændelse – skal stoppes. Det gælder særligt i fuldautomatiske afgørelsessystemer. | Artikel 5, artikel 22 og artikel 32 | Organisatoriske foranstaltninger | * Implementering af en fallback-plan, så sagsbehandlere kan overtage beslutningstagen og sagsbehandlingen fortsættes, hvis systemet af forskellige årsager må stoppes. (Forebyggelse og beskyttelse) |
| 7.3 | Risiko for de facto fuldautomatiske afgørelser i et beslutningsstøttesystem. | Der kan være risiko for, at den menneskelige sagsbehandler ikke foretager en reel vurdering og efterprøvelse af AI-systemets forslag til en afgørelse, dvs. at sagsbehandleren ukritisk lægger systemets forslag til afgørelse til grund (”rubber stamping”). Dette kan have flere årsager, f.eks. at sagsbehandleren ikke kan forstå, hvordan AI-systemet er nået frem til beslutningen (manglende transparens). Den manglende efterprøvelse kan også skyldes, at sagsbehandleren ”stoler på” AI-systemet og dermed stopper med at bruge sin egen kritiske sans og undlader at overveje, om beslutningen kan være forkert (”automation bias”). | Artikel 5 og artikel 22 | Organisatoriske foranstaltninger | * Klar og dokumenteret grænsedragning mellem fuldautomatiske afgørelser og beslutningsstøtte i de situationer, hvor der anvendes en kombination af de to. Dokumentation af overvejelserne af den passende fordeling mellem systemets og menneskets arbejdsopgaver. (Forebyggelse og beskyttelse) * Uddannelse af medarbejdere med henblik på at sikre, at de forstår, hvordan AI-systemet fungerer og kommer frem til beslutninger og systemets begrænsninger, samt give medarbejderne klare kriterier for, hvornår medarbejderne forventes at ændre systemets beslutninger. (Forebyggelse og beskyttelse) * Implementering af dokumenterede retningslinjer og procedurer, der sikrer, at systemets forslag til afgørelser efterprøves af personer med den fornødne kompetence og mulighed for at ændre afgørelsen. (Forebyggelse og beskyttelse) * Understøttelse af, at beslutningsprocessen og ”logikken” i systemet er forklarlig og transparent. (Forebyggelse og beskyttelse) * Løbende dokumenteret revision og statistik på, hvor mange forslag til afgørelser der omgøres af sagsbehandlerne – en meget lav omgørelsesprocent kan være tegn på, at medarbejderne ikke foretager en reel vurdering af systemets forslag til afgørelser. (Opdagelse) * Udførelse af stikprøver af afgørelser med henblik på at efterprøve, om den menneskelige sagsbehandler har foretaget en reel vurdering og korrekte afgørelser. (Opdagelse) * Re-design af AI-systemet med henblik på at gøre forslagene til beslutninger mere transparente og forklarlige. (Respons) * Overvej genoptagelse af sager, hvis afgørelserne ikke er korrekte. (Respons) |
| 7.4 | Risiko for manglende overblik over udledte personoplysninger ved profilering. | Der kan være risiko for, at den dataansvarlige ikke er klar over, at der udledes følsomme personoplysninger om de registrerede. Ved profilering kan der potentielt udledes følsomme personoplysninger ud fra kombinationen af almindelige, ikke følsomme personoplysninger. Det kan f.eks. være muligt at udlede en persons sundhedstilstand ud fra oplysninger om fødevareindkøb. | Artikel 5, 9, 13-14 og 22 | Organisatoriske foranstaltninger | * Hvis der udledes følsomme personoplysninger ved profileringen, skal den dataansvarlige sikre, at behandlingen ikke er uforenelig med det oprindelige formål, at den dataansvarlige har identificeret et retligt grundlag for behandlingen af de følsomme personoplysninger samt at den registrerede oplyses om behandlingen, medmindre en af undtagelserne til oplysningspligten finder anvendelse. (Forebyggelse og beskyttelse) |
| 7.5 | Risiko for utilstrækkelig beskyttelse eller uproportional behandling ved brug af fuldautomatiske afgørelser over for børn. | Databeskyttelsesforordningen pålægger dataansvarlige yderligere forpligtelser i forbindelse med behandling af børns personoplysninger. Databeskyttelsesforordningen rummer ikke et forbud mod brug af fuldautomatiske afgørelser over for børn, men det er hovedreglen, at det ikke skal ske. Der kan dog være visse situationer, hvor dataansvarlige er nødt til at træffe afgørelser, der alene er baseret på automatisk behandling, herunder profilering, som har retsvirkning for eller på tilsvarende vis betydeligt påvirker børn, f.eks. for at beskytte deres velfærd. I så fald kan behandlingen foretages på grundlag af undtagelserne i artikel 22, stk. 2, litra a), b) og c), alt efter hvad der er relevant. | Artikel 22 | Organisatoriske foranstaltninger | * Udarbejdelse af interne retningslinjer for, hvordan behandlingen foregår og børnenes rettigheder sikres. * Udarbejdelse af adfærdskodeks for behandlingen, der omfatter beskyttelsesforanstaltninger for børn. |
| **8. Sikkerhed og robusthed**  **Utilstrækkelig sikkerhed i AI-systemet kan medføre brud på persondatasikkerheden. Manglende robusthed for systemet kan medføre, at systemet ikke kan performe i overensstemmelse med målsætningerne.** | | | | | |
| Brud på persondatasikkerheden kan vedrøre dataene (input og output), modellen eller den underliggende it-infrastruktur, dvs. både software og hardware. Hvis et AI-system angribes, kan både data og systemet ændres, således at systemet træffer anderledes, forkerte beslutninger, eller så det helt lukkes ned. For identifikation af risici og tilhørende afhjælpende foranstaltninger vedrørende temaet ”Robusthed og sikkerhed” kan du med fordel anvende følgende værktøjer og retningslinjer:   * KL’s materiale vedrørende risikostyring på KL og KOMBIT’s videncenter eller egne vejledninger, * Digitaliseringsstyrelsen, Erhvervsstyrelsen og Center for Cybersikkerhed: Vejledning: Tiltag til at sikre brugen af kunstig intelligens, januar 2020 * ISO/IEC 27001:2013, Anneks A, samt * ISO/IEC 29134:2017 og ISO/IEC 29151:2017, der i katalogform beskriver flere af de mulige sikkerhedsforanstaltninger, der kan benyttes. | | | | | |
| **9. Governance og ansvarlighed (accountability)**  **Manglende eller utilstrækkelig afklaring af ansvar for og ledelsesmæssig godkendelse af udvikling og brugen af AI-systemet kan medføre risiko for overtrædelse af reglerne. Tilsvarende gælder manglende eller utilstrækkelig dokumentation for centrale valg og fravalg samt retningslinjer og procedurer eller uddannelse af centrale medarbejdere.** | | | | | |
| 9.1 | Risiko for manglende eller utilstrækkelig afklaring af ansvar for og ledelsesmæssig godkendelse af udvikling og brug af AI-systemet. | Hvis der ikke er en klar og kommunikeret fordeling af ansvar for og ledelsesmæssig godkendelse af udvikling og brug af AI-systemet, kan det medføre, at relevante beslutninger om AI-systemets udvikling, drift eller berostillelse er forkerte eller slet ikke tages. Tilsvarende gælder beslutninger vedrørende implementering og opfølgning på aftalte foranstaltninger. | Artikel 24-25 og artikel 5, stk. 2 | Organisatoriske foranstaltninger | * Implementering af dokumenterede retningslinjer, beslutningsgange og procedurer, der sikrer ansvaret og ansvarligheden for AI-systemet og dets resultater i hele systemets livscyklus. Disse mekanismer skal klart regulere, hvornår ledelsen skal godkende centrale milepæle, herunder beslutning om opstart, udvikling af modellen, idriftsættelse og revision samt centrale performancekrav. (Forebyggelse og beskyttelse) * Dokumenterbare retningslinjer og procedurer for adskillelse af udviklings- og produktionsmiljø og dokumentation af, at adskillelsen er sket. (Forebyggelse og beskyttelse) * Beskrivelse af modellens begrænsninger i lyset af formålet, så alle relevante interessenter er opmærksom på, under hvilke omstændigheder den bør bruges. (Forebyggelse og beskyttelse) * Dokumentation af valg og ændringer af AI-systemet, således at den dataansvarlige kan dokumentere valg og fravalg under udvikling, drift og opdatering af systemet og begrundelserne herfor. (Respons) * Dokumentation af ledelsens godkendelse af udvikling, test og anvendelse af AI-systemet. (Forebyggelse og beskyttelse) * Tilslutning til adfærdskodekser for revisionsprocesser vedrørende maskinlæring eller certificering/standardisering. (Forebyggelse og beskyttelse) |
| 9.2 | Risiko for manglende eller utilstrækkelig uddannelse af centrale medarbejdere. | Hvis medarbejderne ikke er tilstrækkeligt uddannede i at håndtere AI-systemer og dets særegne databeskyttelsesretlige risici, er der risiko for medarbejderfejl og dermed overtrædelse af reglerne. | Artikel 24-25 og artikel 5, stk. 2 | Organisatoriske foranstaltninger | * Uddannelse af relevante medarbejdere, herunder AI-udviklere, projektleder og øvrige nøglepersoner, og dokumentation af uddannelsens indhold. (Præventiv) * Inddragelse af de rette kompetencer fra start og undervejs i AI-systemets livscyklus, herunder AI-udviklere, it-arkitekter, ledelse, proceskonsulenter, DPO’en og personer med domænekendskab (faglige fyrtårne / erfarne sagsbehandlere). (Forebyggelse og beskyttelse, Opdagelse og Respons) |
| 9.3 | Risiko for utilstrækkelig kravstillelse, due diligence og tilsyn med eksterne leverandører af AI-løsninger. | Hvis den dataansvarlige ikke får stillet tilstrækkelige krav om databeskyttelse til løsningens design, herunder foretager den fornødne forudgående kontrol med systemets overholdelse af reglerne (due diligence), er der risiko for, at løsningen ikke overholder databeskyttelsesreglerne. Tilsvarende gælder, hvis den dataansvarlige ikke løbende fører tilsyn med leverandørens overholdelse af kravene under drift. | Artikel 24-25 og artikel 5, stk. 2 | Organisatoriske foranstaltninger | * Dokumenterbare retningslinjer og procedurer for kravstillelse, due diligence af og tilsyn med eksterne leverandører af AI-løsninger. (Forebyggelse og beskyttelse) * Opnåelse af kontraktlige garantier for tredjepartsalgoritmer om, at audit og testning er udført, og at algoritmen er i overensstemmelse med vedtagne standarder. (Forebyggelse og beskyttelse) * Tilsyn af leverandørers overholdelse af kravene. (Forebyggelse og beskyttelse og Opdagelse) |
| 9.4 | Risiko for uhensigtsmæssigt lav kvalitet eller utilsigtede hændelser pga. utilstrækkelig testning, revision og overvågning af AI-systemet. | Hvis der ikke foretages tilstrækkelig og dokumenterbare tests, revision og overvågning af AI-systemet både under udvikling og løbende i drift, er der risiko for manglende opdagelse af fejl og utilstrækkelig opfyldelse af kravene til systemet. | Artikel 24-25 og artikel 5, stk. 2 | Organisatoriske foranstaltninger | * Etablering og vedligeholde af dokumenterbare retningslinjer og processer for test af AI-systemet før go live og med et passende interval efter go live samt ved ændringer. Gennemførte tests, resultater heraf samt foranstaltninger truffet i medfør heraf skal kunne dokumenteres. (Forebyggelse og beskyttelse) * Gennemførelse af en algoritmisk audit af AI-systemet, modellen eller kildekoden fra en uafhængig ekstern tredjepart, f.eks. revisor, advokat eller sikkerhedsekspert m.v. Dvs. en audit for at påvise, at systemet rent faktisk fungerer efter hensigten og ikke giver diskriminerende, fejlagtige eller uberettigede resultater. (Opdagelse) |
| 9.5 | Risiko for manglende evne til at reagere rettidigt på fejl og hændelser vedrørende AI-systemet. | Der er risiko for overtrædelse af reglerne, hvis den dataansvarlige ikke evner at reagere rettidigt på fejl og hændelser og håndtere disse med de fornødne foranstaltninger. Det gælder f.eks. indsigelser eller klager fra de registrerede eller egentlige brud på persondatasikkerheden. | Artikel 24-25, artikel 33-34 og artikel 5, stk. 2 | Organisatoriske foranstaltninger | * Dokumentation af alle hændelser, klager og indsigelser fra de registrerede vedrørende AI-systemet og evaluering heraf. Overvejelse af, om øvrige registrerede kan være påvirkede af fejl. Dokumentation af alle foranstaltninger, der træffes i medfør heraf. (Opdagelse og Respons) * Overvejelse af genoptagelse af sager, hvis der er begået fejl i AI-systemet. (Respons) |