

OKTOBER 2021

NÅR ALGORITMER SAGSBEHANDLER: TRANSPARENS I DET OFFENTLIGES BRUG AF PROFILERINGSMODELLER

Det er et grundlæggende princip i en retsstat, at borgere og offentlighed skal kunne kigge forvaltningen i kortene og få indsigt i sagsbehandling og afgørelser. I Danmark er der imidlertid ikke et samlet overblik over myndigheders brug af kunstig intelligens som for eksempel profileringsmodeller, og det udgør et retssikkerhedsproblem. Profileringsmodeller udfører desuden komplekse analyser, der kan risikere at gøre dem til en slags "black box", som mennesker ikke kan forstå. Dette notat redegør for udfordringerne ved den manglende transparens i og om det offentliges brug af modellerne.

INSTITUT FOR MENNESKERETTIGHEDER ANBEFALER, AT:

- Digitaliseringsstyrelsen, som led i det fællesoffentlige arbejde med kommuner og regioner, opretter et **offentligt register** over samtlige offentlige myndigheders brug af profileringsmodeller rettet mod borgere.
- Justitsministeriet med inddragelse af Datatilsynet og Digitaliseringsstyrelsen udsteder en vejledning om offentlige myndigheders brug af profileringsmodeller, hvori der stilles krav **om Artificial Intelligence-konsekvensanalyser** (AI-konsekvensanalyser), og at disse eller et uddrag heraf **offentliggøres** i registret.
- AI-konsekvensanalyserne udtrykkeligt forholder sig til, hvordan modellerne lever op til krav om **algoritmisk transparens**.

Danmark er blandt de førende i verden, når det gælder digitalisering af borgernes kontakt med stat, region og kommune.¹ Digitaliseringen omfatter i stigende grad også brug af kunstig intelligens, hvor store mængder data selv "lærer" at analysere opgaver og finde løsninger. Ansigtsgenkendelse, maskinoversættelser og rutevejledninger er alle eksempler på brug af kunstig intelligens, som de fleste kender, men teknologien kan også bruges i offentlige myndigheders sagsbehandling, når de skal træffe afgørelser om for eksempel ydelser, indsatser eller sanktioner overfor borgere.

I 2019 udgav regeringen "National strategi for kunstig intelligens", hvori man understreger vigtigheden af fortsat udvikling på området.² Strategiens vision er, at "Danmark skal gå forrest med ansvarlig udvikling og anvendelse af kunstig intelligens", og visionen er allerede nu udmøntet i en række signaturprojekter³ og forskningsprogrammer på både nationalt og EU-plan.

Trods det politiske fokus og de mange tiltag mangler der imidlertid et overblik over, hvordan offentlige myndigheders brug af kunstig intelligens griber ind i menneskerettighederne. Spørgsmålet er kun sporadisk blevet adresseret i Danmark, og det offentliges forskellige digitaliseringsstrategier kommer ikke nærmere ind på, hvilke følger det kan få for borgernes retssikkerhed, at dele af sagsbehandlingen udføres af algoritmiske profileringsmodeller.

Både i EU, Europarådet og FN⁴ er der udtrykt bekymring for, at kunstig intelligens kan underminere beskyttelsen af den enkeltes rettigheder. Brugen af profileringsmodeller kan fx skabe risici for diskrimination på en række forskellige måder.

Institut for Menneskerettigheder anser offentlige myndigheders brug af kunstig intelligens for at have stor betydning for borgernes rettigheder, og instituttets arbejde tager netop udgangspunkt i en rettighedsbaseret tilgang. Dette har som forudsætning, at det skal være muligt at få indblik i myndighedernes brug af profileringsmodeller og til en vis grad også indblik i, hvordan de virker.

Instituttet vurderer overordnet, at der er behov for regler om, at offentlige myndigheder skal påvise, hvordan de sikrer samtlige rettigheder og retsgarantier for borgerne, når de bruger algoritmiske profileringsmodeller. Dette gøres efter instituttets vurdering bedst ved gennemførelsen af såkaldte 'artificial intelligence'-konsekvensanalyser (AI-konsekvensanalyser), som vi redegør for i et beslægtet notat.⁵

FOKUS FOR NOTATET

En central retssikkerhedsmæssige udfordring er, at profileringsmodeller kan forhindre transparens i forhold til myndighedens sagsbehandling og afgørelser. Transparens i og om algoritmiske profileringsmodeller brugt af det offentlige er afgørende for muligheden for kontrol, og det er fokus for dette notat, hvor vi ser nærmere på såkaldt systemisk og algoritmisk transparens.

HVAD ER EN ALGORITMISK PROFILERINGSMODEL?

En algoritmisk profileringsmodel er en computerdrevet matematisk formel, som på baggrund af statistiske data eller variabler vurderer, om et objekt – eksempelvis en borger – har en bestemt egenskab, for eksempel om borgeren er berettiget til et tilskud.

Profileringsmodeller udvikles og trænes på baggrund af store mængder data om allerede afgjorte sager på området. Modellerne kan analysere og behandle data og finde sammenhænge og mønstre, der er langt mere komplekse, end hvad mennesker er i stand til, og de kan derfor være en stor ressource i sagsbehandlingen.

Profileringsmodeller kan både bruges til at træffe forvaltningsretlige afgørelser (fuldautomatisering) eller til at oplyse eller støtte myndighedens afgørelse (automatiseret beslutningsstøtte). Bruges modellen til beslutningsstøtte, kan dette enten ske for at afklare faktuelle omstændigheder som led i sagsoplysning, eller som støtte i den juridiske afvejning af en sag. Udfordringerne med transparens er til stede, uanset om modellen bruges fuldautomatiseret eller til automatiseret beslutningsstøtte.

Bruges modellen til beslutningsstøtte, kan dette enten ske for at afklare faktuelle omstændigheder som led i sagsoplysning eller som støtte i den juridiske afvejning af en sag. Udfordringerne med uigennemsigthed er til stede, uanset om modellen bruges fuldautomatiseret eller til automatiseret beslutningsstøtte.

Profileringsmodeller er allerede under udvikling og i brug i det offentlige Danmark. Blandt andet har Styrelsen for Arbejdsmarked og Rekruttering udviklet og implementeret en model til forudsigelse af langtidsledighed blandt nyledige, og Gladsaxe Kommune udviklede i 2018 – men implementerede ikke – en model til forudsigelse af mistriksel i børnefamilier. I begge tilfælde behandler modellen udvalgte data helholdvis om nyledige og om børnefamilier, og modellens resultater kan få betydning for den indsats, der igangsættes over for de identificerede ledige og børnefamilier.

TRANSPARENS SOM ET RETSSIKKERHEDSPRINCIP

Transparens – eller åbenhed – er et helt grundlæggende krav til offentlige myndigheder. Åbenheden skal sikre, at borgere og offentlighed kan kigge myndighederne efter i sømmene og stille dem til ansvar for deres handlinger.

Princippet indebærer, at myndighederne i videst muligt omfang skal være tilgængelige og åbne om deres behandling af sager, og er en integreret del af mange af de gældende forvaltningsretlige regler. For eksempel er princippet om åbenhed indbygget i reglerne om aktindsigt, som giver borgeren, og i visse tilfælde

offentligheden, indsigt i en sags dokumenter og oplysninger. Denne adgang til at efterse myndighedernes sagsbehandling udgør en kontrol med myndighedernes ageren og er en forudsætning for deres demokratiske legitimitet.

Myndighedernes begrundelsespligt sikrer ligeledes borgere et indblik i myndighedens virke. Begrundelsespligten indebærer, at en myndighed skal redegøre for, hvordan den er nået frem til en afgørelse. Begrundelsen skal formuleres på en måde, som gør det muligt for en tilsynsmyndighed at efterprøve, om afgørelsen fik det rette udfald. Denne bestemmelse tjener til at føre kontrol med myndigheder og til at give den enkelte borger klarhed over, hvordan afgørelsen blev truffet.

Det er også et af grundprincipperne i EU's databeskyttelsesforordning, der blandt andet udmønter sig i oplysningsforpligtelser for myndigheder og indsigtsret for borgere, når myndigheder behandler deres personoplysninger. Disse regler skal sikre, at borgerne kan få indblik i, hvilke personoplysninger der behandles om dem af de enkelte myndigheder, og til hvilket formål. Herudover indeholder databeskyttelsesforordningen særlige regler om myndighedernes fuldautomatiserede brug af profileringsmodeller.

Transparens vedrører således mange forskellige rettigheder og giver indblik i forskellige dele af myndighedernes sagsbehandling.

TRANSPARENS I OG OM PROFILERINGSMODELLER

Når offentlige myndigheder i stigende grad bruger profileringsmodeller i sagsbehandlingen, kan det svække beskyttelsen af de grundlæggende rettigheder og principper, der sikrer åbenhed om og indsigt i myndighedens sagsbehandling. Overordnet kan behovet for transparens efter instituttets vurdering opdeles i to kategorier: systemisk og algoritmisk transparens.

Den første omhandler transparens om brugen af modellerne. Dette krav fokuserer på, hvorvidt der er tilstrækkelig åbenhed om, hvor mange myndigheder der bruger profileringsmodeller i sagsbehandlingen, og hvad de bruger dem til. Fortæller myndighederne, om det er en profileringsmodel eller et menneske, der har truffet afgørelsen eller en del heraf? Åbenhed om disse forhold er efter vores vurdering en væsentlig forudsætning for den enkelte borgers retssikkerhed.

Den anden kategori følger af den første og drejer sig om transparens i selve modellen. Åbenhed om brugen af profileringsmodeller forudsætter nemlig, at myndighederne til en vis grad er i stand til at redegøre for, hvordan modellen er udviklet, og hvilke opgaver den udfører. Det fremhæves ofte, at profileringsmodeller har et element af uigennemsiagtighed, fordi de udfører komplekse analyser, som de færreste mennesker har forudsætninger for at forstå, og at de derfor udgør en slags "black box". Grundlæggende rettigheder og principper, som beskrevet ovenfor, stiller imidlertid krav til et vist indblik i

profileringsmodellerne. Derfor skal myndighederne sikre sig, at modellen lever op til disse krav om transparens, før de tager den i brug.

SYSTEMISK TRANSPARENS: ÅBENHED OM BRUG

Når offentlige myndigheder i Danmark bruger profileringsmodeller som led i sagsbehandlingen, er der på nuværende tidspunkt ingen udtrykkelig pligt til at informere borgeren om brugen af modellerne, ligesom den bredere offentlighed heller ikke har indblik i, hvor mange danske myndigheder der gør brug af profileringsmodeller. Nogle myndigheder vælger at offentliggøre visse oplysninger om deres brug af profileringsmodeller, mens andre ikke stiller oplysninger til rådighed om deres brug.

Dette forhold udgør en udfordring for retssikkerheden, idet åbenhed er en forudsætning for, at myndighederne kan stilles til ansvar i eventuelle sager, ligesom også den bredere offentlighed og tilsynsmyndigheder har krav på indsigt i myndighedernes ageren.

For borgeren, for den bredere offentlighed og for tilsyns- og kontrolinstanser er det derfor væsentligt, at der er en offentligt tilgængelig oversigt over samtlige myndighedernes brug af profileringsmodeller.

For den enkelte borger, der har krav på at blive oplyst om, hvorvidt og hvordan en profileringsmodel er brugt i en konkret afgørelse, bør en sådan oplysning desuden være en naturlig del af kommunikationen til borgeren om afgørelsen. Det kan for eksempel fremgå af det brev, borgeren modtager, når afgørelsen er truffet, om afgørelsen er truffet udelukkende eller med støtte fra en profileringsmodel, og hvordan modellen i givet fald er indgået i afgørelsen eller selv er nået frem til et givent resultat om borgeren.

I forhold til oprettelsen af en samlet offentligt tilgængelig oversigt kan danske myndigheder lade sig inspirere af erfaringer fra landene omkring os. I Amsterdam og Helsinki har man for nylig oprettet registre om myndighedernes brug af kunstig intelligens i byerne – i en erkendelse af, at myndighedernes brug af kunstig intelligens stadig bliver mere udbredt, og at det er vigtigt at bevare borgernes tillid til offentlige myndigheders virke i takt med denne udvikling.⁶

Også EU-Kommissionens har i et udkast til en forordning for kunstig intelligens foreslået etablering af en offentlig EU-database for visse højrisiko AI-systemer, herunder offentlige myndigheders brug af profileringsmodeller.⁷

OFFENTLIGE AI-REGISTRE I AMSTERDAM OG HELSINKI

Amsterdam i Holland og Helsinki i Finland blev i efteråret 2020 nogle af de første byer i verden, der åbnede offentligt tilgængelige registre over offentlige myndigheders brug af kunstig intelligens – såkaldte AI-registre. Registerne giver overblik over, hvilke AI-systemer der anvendes af de offentlige myndigheder i byerne. For hvert system er der fremlagt oplysninger om: (1) datasæt, (2) databehandling, (3) sikring mod diskrimination, (4) menneskelig kontrol og (5) generelle risici. Registerne giver også et grundlæggende indblik i, hvad algoritmer er, og hvordan de påvirker hverdagen for byens borgere.⁸

Der er god inspiration at hente i eksemplerne fra Holland og Finland, men instituttet vurderer, at offentlige myndigheders brug af profileringsmodeller rejser nogle rettighedsmæssige risici, som kalder på mere fyldestgørende oplysninger om modellernes brug. Dette kalder vi systemisk transparens.

Systemisk transparens dækker de informationer, der ”omkredser” og supplerer algoritmisk transparens (se nedenfor). I praksis er tekniske oplysninger (algoritmisk transparens) om modellens funktion nemlig ikke altid tilstrækkelige for, at borgere og offentlighed kan forstå modellen og dens brug.

SYSTEMISK TRANSPARENS: FIRE TYPER AF INFORMATIONER

Modellens træning: Informationer om det datasæt, der er brugt til udvikling af modellen.

Modellens kvalitet og risici: Informationer om modellens fejlrate og ydeevne, når den anvendes på nye sager og risici i forhold til rettigheder og retssikkerhed.

Modellens anvendelse: Information om, hvad modellen skal bruges til, og om den er egnet til formålet samt overordnede og letforståelige beskrivelser af modellens funktion, type og design.

Kontrol og tilsyn med modellen: Rammer for tilsyn og oplysninger om, hvordan modellen forklares for borgeren.

Et samlet overblik over offentlige myndigheders brug af profileringsmodeller vil efter instituttets vurdering være med til at styrke retssikkerheden på området.

Oversigten kan efter instituttets vurdering oprettes af Digitaliseringsstyrelsen, der har til formål netop at skabe en sammenhængende offentlig service, der giver borgeren et overblik over det offentlige Danmarks digitalisering blandt andet som led i det fællesoffentlige arbejde med kommunerne og regionerne.⁹ Oversigten

kan sikre systemisk transparens ved, at myndighederne offentliggør hele eller dele af de AI-konsekvensanalyser, som instituttet anbefaler, at der bliver indført regler om.

ALGORITMISK TRANSPARENS: INDSIGT I HVORDAN MODELLEN TRÆFFER AFGØRELSER

Offentligt tilgængelige oplysninger om myndighedernes brug af profileringsmodeller er vigtig for retssikkerheden, men det giver ikke mening, hvis modellerne er så komplekse, at det er vanskeligt eller umuligt at få indsigt i, hvordan de virker.

Myndigheder har pligt til at nå frem til en lovlige og rigtig afgørelse. For at påvise, at det er sket – eller for, at en borger kan påvise og en domstol kan dømme, at det ikke er sket – skal både myndighed, borger og domstol kunne få indblik i baggrunden for afgørelsen. Dette gælder naturligvis også, når afgørelsen træffes med brug af profileringsmodeller.

For at forhindre, at profileringsmodellerne fører til diskriminerende praksis, kan det desuden være nødvendigt både at få indblik i baggrunden for den konkrete afgørelse og kendskab til, hvordan modellen generelt danner sine resultater, samt hvilke data modellen er trænet på.

Algoritmisk transparens dækker over adgangen til information om, hvordan en model træffer sine vurderinger. Nogle modeller er umiddelbart transparente i deres design, uden at man har behov for tekniske redskaber for at få indblik i modellen. Andre modeller kan imidlertid være så komplekse, at selv personer med teknisk fagkundskab ikke kan forstå dem. Grænserne for, hvor kompleks en model bør blive er indgående behandlet i forskningslitteraturen og rejser mange spørgsmål om udvikling og design. Blandt andet rejser det spørgsmål om valget mellem uigennemsigtige og komplekse modeller, som (oftere) når frem til korrekte afgørelser om borgerne, og transparente og simple modeller, der til gengæld (oftere) kommer med forkerte vurderinger om borgeren.

Dette kan skabe dilemma mellem modellens kompleksitet og dens algoritmiske transparens – eller sagt på en anden måde, et dilemma mellem på den ene side modellens adgang til at udføre sin opgave, og på den anden side menneskers mulighed for at få indsigt i dens funktion og resultater.

I sin kerne er dilemmaet derfor mellem en adgang til at forstå og efterprøve, om modellens afgørelse er korrekt på den ene side, og sikre, at modellen i flest mulige tilfælde når frem til et korrekt resultat på den anden side. Dilemmaet er retssikkerhedsmæssigt vigtigt, fordi modeller kan videreføre og forstærke risikoen for diskrimination og usaglighed og skabe andre typer af udfordringer, end man kender det fra manuel sagsbehandling.

Samlet set er det instituttets vurdering, at myndighederne bør stile efter at anvende den mest muligt algoritmisk transparente model i sagsbehandlingen og under alle omstændigheder redegøre for sit valg, herunder brugen af tekniske metoder for at gøre uigennemsigtige profileringsmodeller transparente – såkaldt xAI.

Derfor anser instituttet det for nødvendigt, at myndighederne udtrykkeligt i AI-konsekvensanalyserne forholder sig til udfordringerne ved algoritmisk transparens og redegør for deres valg af teknologi og imødegåelse af retssikkerhedsmæssige risici ved valget. Det er efter instituttets opfattelse vigtigt, at myndigheden dokumenterer, at den har forholdt sig til problemstillingen og anvender teknologien på en måde, der sikrer borgerens rettigheder og retssikkerhed.

ANBEFALINGER: HUSK RETTIGHEDERNE

Danmark er verdensmester i digital forvaltning. Offentlige myndigheder er ambitiøse, når det gælder udviklingen af digitale løsninger – men de har ikke taget stilling til, hvordan disse løsninger skal afspejle den fortsatte sikring af menneskerettigheder og retsprincipper.

En rettighedsbaseret tilgang til offentlige myndigheders brug af profileringsmodeller er særligt vigtig, da det gælder afgørelser om borgerne. Der kan være tale om afgørelser, som har stor indvirkning på borgernes retsstilling som for eksempel afgørelser om udsatte borgeres behov for støtte. Offentlige myndigheder er forpligtet til at handle sagligt og til at sikre beskyttelsen af borgernes rettigheder og retssikkerhed, også når sagerne automatiseres.

Det er afgørende, at både den bredere offentlighed og den enkelte borger har kendskab til myndighedernes brug af profileringsmodeller og ved, hvad disse modeller kan og ikke kan, samt hvilke udfordringer de rejser i forhold til borgernes rettigheder. Også tilsynsmyndigheder skal have et indgående kendskab til disse forhold. Det gælder også en vis teknisk fagkundskab om profileringsmodellernes opbygning og kunnen.

Institut for Menneskerettigheder efterlyser et større fokus på de mange rettighedsmæssige udfordringer, som udviklingen af en digital forvaltning stiller os overfor. En helt grundlæggende forudsætning for dette er åbenhed om myndighedernes anvendelse af for eksempel profileringsmodeller.

LÆS MERE I VORES RAPPORT

I rapporten 'Når algoritmer sagsbehandler – Rettigheder og retssikkerhed i offentlige myndigheders brug af profileringsmodeller' udgivet i oktober 2021 kortlægger Institut for Menneskerettigheder, hvordan borgernes rettigheder og retssikkerhed påvirkes, når offentlige myndigheder benytter sig af kunstig intelligens. Rapporten er en af de første, der tager udgangspunkt i danske myndigheders praksis og dansk lovgivning.

Dette notat gengiver dele af rapportens konklusioner.

Læs hele rapporten på menneskeret.dk/algoritmer.

SLUTNOTER

¹ Digitalstyrelsen (2020), Ny FN-måling: Danmark er fortsat verdensmestre i offentlig digitalisering, tilgængelig på: <https://digst.dk/nyheder/nyhedsarkiv/2020/juli/ny-fn-maaling-danmark-er-fortsat-verdensmestre-i-offentlig-digitalisering/>

² Finansministeriet og Erhvervsministeriet (2019): National strategi for kunstig intelligens tilgængelig på: https://www.regeringen.dk/media/6537/ai-strategi_web.pdf

³ Digitalstyrelsen (2019), Kommuner og regioner skal afprøve kunstig intelligens for at løfte kvaliteten i den offentlige service, tilgængelig på: <https://digst.dk/nyheder/nyhedsarkiv/2019/oktober/kommuner-og-regioner-skal-afproeve-kunstig-intelligens-for-at-loefte-kvaliteten-i-den-offentlige-service/>

⁴ Kommissionens hvidbog (2020), " On Artificial Intelligence - A European approach to excellence and trust" Tilgængelig på: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf , Europarådets resolution, CM/Rec(2020)1, tilgængelig på: <https://rm.coe.int/09000016809e1154> , FN's Højkommissariat for Menneskerettigheder (2019), "FN's Specialrapportør for ekstrem fattigdom", rapport A/74/48037.

⁵ Institut for Menneskerettigheder, (2021) "Når algoritmer sagsbehandler: Klare rammer for det offentliges brug af profileringsmodeller", <https://menneskeret.dk/udgivelser/klare-rammer-offentliges-brug-profileringsmodeller>

⁶ News powered by cision, (2020) Helsinki and Amsterdam first cities in the world to launch open AI register, tilgængelig på: <https://news.cision.com/fi/city-of-helsinki/r/helsinki-and-amsterdam-first-cities-in-the-world-to-launch-open-ai-register,c3204076>

⁷ Kommissionens udkast til forordning om harmoniserede regler for kunstig intelligens COM(2021) 206 final, artikel 60

⁸ Registerne er tilgængelige her: <https://algoritmeregister.amsterdam.nl/en/ai-register/> og her: <https://ai.hel.fi/en/ai-register/>

⁹ For særligt udvalgte projekter om myndighedernes brug af kunstig intelligens (såkaldte signaturprojekter), har Digitaliseringsstyrelsen desuden oprettet et overblik her: <https://digst.dk/strategier/kunstig-intelligens/signaturprojekter/>