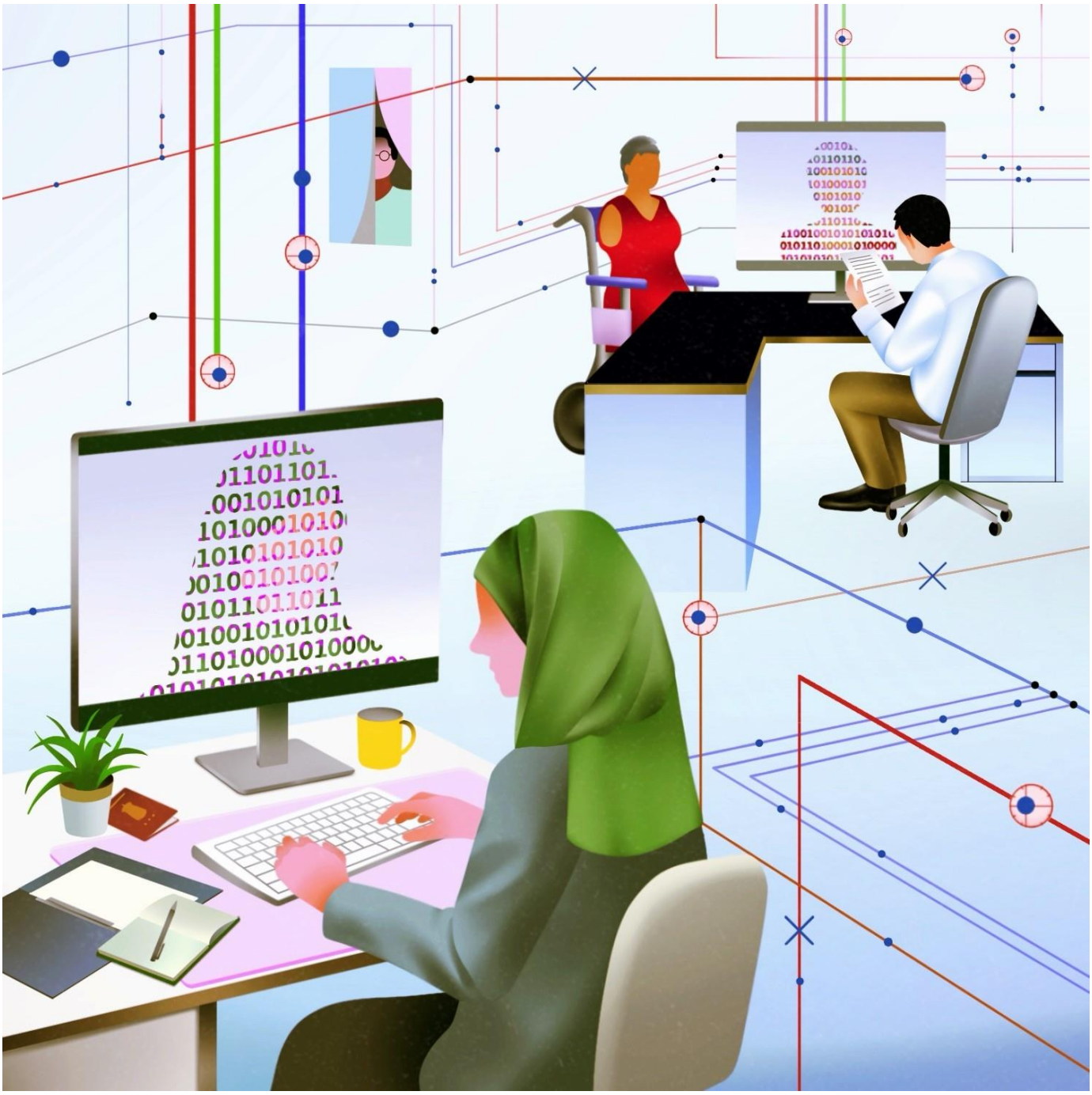


INDKODET URETFÆRDIGHED

OVERVÅGNING OG DISKRIMINATION I DANMARKS AUTOMATISEREDE VELFÆRDSSYSTEM

AMNESTY INTERNATIONAL

Amnesty International er bevægelse, der består af 10 millioner mennesker, som mobiliserer menneskeligheden i os alle, og som iværksætter kampagner for at opnå forandringer til gavn for menneskerettighederne overalt.

Vores vision er en verden, hvor magthaverne overholder deres løfter, respekterer folkeretten, og holdes ansvarlige.

Vi er uafhængige af enhver regering, politisk ideologi, økonomiske interesser eller religion, og er hovedsageligt finansierede af vores medlemmer, samt individuelle bidrag.

Vi tror på, at vi kan ændre vores samfund til det bedre, ved at handle i solidaritet og medfølelse med mennesker over hele kloden.

© Amnesty International 2024
Medmindre andet er angivet, har indholdet i dette dokument licens under et Creative Commons (tilskrivning, ikke-kommerciel, ingen afledninger, international 4.0) licens.
https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode
Besøg tilladelsessiden på vores hjemmeside for mere information: www.amnesty.org
Hvor materiale er tilskrevet en anden copyright-ejer end Amnesty, er materialet ikke underlagt Creative Commons Licensen.
Først udgivet i 2024
af Amnesty International Ltd
Peter Benenson House, 1 Easton Street
London WC1X 0DW, UK

Index: EUR 18/8738/2024
Originalsprog: Engelsk

amnesty.org

Forsidebillede: *© Ard Su*

AMNESTY INTERNATIONAL

SAMMENFATNING

Danmark hyldes ofte for at være et velfærdssamfund, som tilbyder et socialt sikkerhedsnet for sine borgere ved at bruge 26% af sit BNP på velfærdsydelser. Ved at øge digitaliseringen, har Danmark placeret sig som en eksemplarisk digital velfærdsstat, og anses ofte som en frontløber for digitalisering i Europa. Landets overgang til en digital velfærdsstat er opstået ved at skiftende regeringer har gjort en indsats for at strømline administrative opgaver, og sørge for at velfærd og andre væsentlige offentlige ydelser leveres effektivt, og fungerer ordentligt.

Virkeligheden er imidlertid, at Danmarks overgang til et digitalt velfærdssystem nu betyder at modtagere af sociale ydelser risikerer at blive udsat for menneskerettighedskrænkelser.

I denne rapport behandles resultatet af Amnesty Internationals research af Danmarks sociale ydelsessystem, herunder den danske stats brug af algoritmer til kontrol af socialt bedrageri som basis for fordeling af sociale ydelser. Det sker igennem den offentlige instans, Udbetaling Danmark (UDK), samt igennem den selvejende institution, Arbejdsmarkedets Tillægspension (ATP), som har mandat fra den danske stat til at administrere sociale ydelser på vegne af UDK.

UDK blev oprettet i 2012 ved Lov om Udbetaling Danmark, som har til formål at centralisere udbetalingen af de sociale ydelser, som kommunerne står for, herunder børneydelser, pensionsydelser, boligstøtte, arbejdsløshedsunderstøttelse, barsels-og sygedagpenge. UDK/ATP oprettede "Den fælles Dataenhed" som fik til opgave at udvikle datadrevne bedragerikontolalgoritmer, der efter sigende skulle have til formål at identificere svigagtige ansøgninger om sociale ydelser, og udtage dem til yderligere undersøgelser i samarbejde med private virksomheder. Den Fælles Dataenhed forbinder eller sammenkører millioner af danske borgeres personlige data fra registre (offentlige databaser), som indeholder oplysninger om ydelsesmodtagere og deres familiemedlemmer, eller andre medlemmer af husstanden. Oplysningerne omfatter blandt andet modtagernes bopæl og flytninger, statsborgerskab, fødested, familieforhold og -omstændigheder, boligforhold og bygningstilstande, ansættelsesforhold, indtægter, skatteforhold, helbred og uddannelse. UDK bruger profileringsmodeller (som analyserer aspekter af et individs personlighed, adfærd, interesser og vaner for at forudsige eller tage beslutninger, der vedrører individet) til at analysere disse oplysninger med henblik på at identificere personer, der angiveligt har forøget risiko for at modtage uberettiget ydelser, og udtage dem til yderligere undersøgelser. Modellerne genererer en "undringsliste" med de mennesker som angiveligt har høj risiko (statistisk set) for svigagtigt eller uberettiget at modtage ydelser. I 2019 forlød det, at UDK brugte omkring 60 forskellige modeller for kunstig intelligens (AI) og maskinlæring (ML) til at identificere personer, som de mente var i forhøjet risiko for at modtage uberettigede ydelser.

Selvom brugen af teknologi i den offentlige sektor ofte fremstilles som værende objektiv og upartisk, er det praktisk talt umuligt at skabe værdineutrale teknologier, da de i sagens natur afspejler de underliggende love, regler, standarder, samt holdnings- og adfærdsmønstre i de miljøer, hvor de produceres. For så vidt angår den offentlige sektor, og i endnu højere grad når det bruges til at opdage svindel, tilskynder de ofte til diskrimination af visse grupper på baggrund af deres race, etniske tilhørsforhold, religion, migrantstatus, indtægt, køn, handicap og alder.

Denne rapport beskriver og stiller spørgsmålstegn ved UDK's og ATP's brug af bedrageri-algoritmer og hvordan de påvirker ydelsesmodtageres menneskerettigheder negativt, herunder deres ret til privatliv, ligebehandling og ikke-diskrimination, værdighed, social tryghed og klageadgang.

Rapporten viser også, hvordan modtagere af sociale ydelser udsættes for masseovervågning gennem brug af traditionelle og digitale overvågningsmekanismer. Den fremhæver de diskriminerende virkninger, der følger af UDK/ATP's brug af algoritmiske systemer til afsløring af bedrageri. Amnesty Internationals research har afdækket, hvordan denne diskriminerende effekt opstår indenfor de diskriminerende eller ulige strukturer, der findes i danske samfundsinstitutioner – fjendtlige danske love, regler, standarder, holdningsmønstre og adfærd, der skaber og promoverer "andetgørelse" ("othering"), eller forskelsbehandling af forskellige grupper. Det er en praksis, som ikke kun tilskynder til diskrimination af marginaliserede grupper, men som også giver mulighed for masseovervågning af befolkningen generelt.

Desuden fremhæver rapporten måder, hvordan UDK/ATP's brug af teknologi og digitalisering udelukker marginaliserede grupper fra at få adgang til de sociale sikringsordninger, og endvidere på negativ vis inkluderer eller tvangsinkluderer grupper, der ikke ønsker at være omfattet af ordningerne.

DANMARKS MENNESKERETTIGHEDSFORPLIGTELSER

I henhold til forskellige internationale menneskerettighedsinstrumenter og -standarder, EU-lovgivning og -standarder, samt nationale love og politikker, har Danmark en forpligtelse til at respektere og beskytte retten til blandt andet privatliv, databeskyttelse, ytringsfrihed, ligestilling og ikke-diskrimination, samt social sikring. Danmark har også et særligt ansvar for at respektere og beskytte menneskerettighederne for børn, kvinder og LGBTI-personer, mennesker med handicap, samt ældre. Staten har endvidere en forpligtelse til at sikre gennemsigtighed, ansvarlighed og retten til fyldestgørende oprejsning. Endelig har virksomheder et ansvar for at respektere menneskerettighederne i alle deres forretningsaktiviteter, og stater har en forpligtelse til at beskytte mod menneskerettighedskrænkelser inden for deres territorium og/eller jurisdiktion begået af tredjeparter, herunder erhvervsvirksomheder.

"DYNELØFTERI": MONITORERING OG OVERVÅGNING AF ANSØGERE, YDELSESMODTAGERE OG TILKNYTTEDE VIRKSOMHEDER

Amnesty Internationals research viser, at den danske stat har indført lovgivning, der krænker privatlivets fred, og som gør det muligt at indsamle data fra ydelsesmodtagere og deres husstandsmedlemmer uden deres samtykke, med det formål at overvåge landets indbyggere og kontrollere for socialt bedrageri. Disse love tillader og styrer UDK/ATP's og kommunernes masseudtræk og behandling af personoplysninger om ydelsesmodtagere med henblik på afsløring af bedrageri, herunder love om sammenlægning af offentlige databaser, brug af algoritmer til bedragerikontrol på disse data, den uregulerede brug af sociale medier, samt den rapporterede brug af geolokaliseringsdata til efterforskning af bedrageri.

Vores research har vist, at indsamlingen og sammenlægningen af store mængder persondata i offentlige databaser, reelt har tvunget modtagere af sociale ydelser til at opgive retten til privatliv og databeskyttelse, for at kunne udøve deres ret til social sikring, og andre sociale og økonomiske rettigheder.

Indsamling og behandling af store mængder data (herunder følsomme oplysninger, der kan afsløre race og etnicitet, helbred, handicap og seksuel orientering), centralisering eller samkøring af databaser til bedragerieftersforskning uden borgernes samtykke, samt brugen af sociale medier, er en uforholdsmæssigt indgribende måde at afsløre bedrageri på. Desuden er det tvivlsomt hvor nyttigt det er. Myndighederne bør finde mindre indgribende metoder til at afsløre socialt bedrageri; metoder, som opfylder menneskerettighedernes krav om nødvendighed og proportionalitet.

Personer, der ansøger om eller modtager ydelser, udsættes også for »traditionelle« eller »analoge« former for overvågning og kontrol med henblik på at afsløre bedrageri. Disse bruges fortsat side om side med bedrageri-algoritmer, og omfatter kommunernes fortsatte revurdering af støtteberettigelse, sager om bedragerikontrol, eller rapporter fra andre offentlige myndigheder, herunder skattemyndigheder og politiet, samt anonyme anmeldelser fra offentligheden. Indgribende overvågning og kontrol fra medborgere, kommuner og andre offentlige myndigheder griber ind i, eller begrænser ydelsesansøgeres og -modtageres ret til privatliv, menneskelig værdighed og social sikkerhed, hvilket yderligere forstærker de menneskerettighedskrænkelser, som disse digitale og analoge former for overvågning muliggør.

Disse analoge former for kontrol og overvågning har, sammen med alt for omfattende digitale kontrolmetoder, skabt et skadeligt overvågningssystem. Konstant overvågning af ydelsesmodtagere har i tilgift en negativ indvirkning på folks mentale sundhed, og forårsager betydelig stress og angst.

STRUKTUREL DISKRIMINATION; FORHØJET RISIKO FOR ALGORTIMISK DISKRIMINATION

I Danmarks tilfælde befinder det sociale ydelsessystem sig i et allerede fjendtligt miljø for asylansøgere, mennesker, der har opnået flygtningestatus i Danmark, migranter, samt racialiserede grupper, hvilket øger risikoen for diskrimination af disse grupper på baggrund af deres race, etnicitet og religion. Denne diskrimination kommer til udtryk i den differentierede allokering af ikke-bidragspligtige børneydelser til personer, der har fået flygtningestatus i Danmark. Den danske stats indførelse af langvarige, tyngende og uforholdsmæssigt indgribende opholdskrav for personer, der ansøger om børneydelser, fører til diskrimination af personer, der har fået flygtningestatus i Danmark, især fra lande som Syrien, Afghanistan, Libanon og Irak. Derved begrænses deres menneskeret til fuld adgang til børneydelser (en ikke-bidragspligtig ydelse), på lige fod med andre grupper. Den begrænsede adgang til børneydelser, får negative konsekvenser for forældrenes evne til at opfylde deres børns basale behov. Problemet forstærkes af, at migrant- og flygtningeforældre ofte har svært ved at få adgang til det danske arbejdsmarked, på grund af sprogbarrierer og mangel på relevant viden om, og netværk i Danmark. Amnesty Internationals research viser også, at den diskrimination, der følger af UDK/ATP's algoritmer, først og fremmest sker fordi det danske samfund har strukturer, love, regler, institutioner, standarder og værdier, som forskelsbehandler. Disse diskriminerende strukturer er indlejret i UDK/ATP's algoritmiske modeldesign, og muliggør at kategoriseringer, der er baseret på forskelle eller "othering" skabes og fremmes. Mere specifikt kan det betyde, at love, regler, standarder og holdningsmønstre, der er designet eller etableret af de mest socialt dominerende grupper i Danmark, og som ser ud til at være neutrale eller »race-agnostiske«, i praksis have en diskriminerende effekt.

Amnesty International finder , at UDK/ATP's brug af algoritmer til at identificere ydelsesmodtagere og -ansøgere, der forventes at ville begå bedrageri, kan føre til både risikable og uproportionale indgreb overfor allerede marginaliserede grupper. Det drejer sig om mennesker, som UDK/ATP betegner som »andre«("others") i det danske samfund, fordi de har afvigende eller »usædvanlige« bolig- eller familieforhold eller »tilknytning til udlandet«.

Danske myndigheder fremstiller marginaliserede grupper som personer, der mere tilbøjelige til at begå bedrageri, eller som personer der ikke er berettiget til at modtage ydelser, og der er derved en risiko for, at disse grupper vil blive udtaget til nærmere undersøgelse for bedrageri på grund af deres familie- og boligforhold og udenlandske tilhørsforhold. Disse karakteristika eller variabler kan komme til at erstatte race, migrationsstatus og social og økonomisk status, og kan derved tilskynde til diskrimination af personer med disse karakteristika.

Amnesty International har fundet, at ved at kategorisere ud fra »andetgørelse« ("othering") eller forskelligheder, risikerer man direkte eller indirekte at diskriminere lavindkomstgrupper, racialiserede grupper, migranter, flygtninge, etniske minoriteter, mennesker med handicap og ældre mennesker. Derved straffes disse grupper, fordi deres ret til ligebehandling og ikke-diskrimination krænkes, samtidig med at de risikerer at blive frataget retten til sociale sikringsydelser.

USÆDVANLIGE HUSSTANDS-, FAMILIE- OG OPHOLDSMØNSTRE

Et af hovedprincipperne bag UDK/ATP's modeller for bedragerikontrol er at identificere »usædvanlige« eller »atypiske« levevilkår eller -arrangementer; det vil sige atypiske forhold og usædvanlige opholdsmønstre, der bruges til at iværksætte yderligere undersøgelser om bedrageri. Selv om det ikke er klart defineret i loven, hvad der udgør »usædvanlige« eller »atypiske ordninger«, og der derved er risiko for at beslutninger tages på et vilkårligt grundlag, bruger algoritmerne oplysninger som husstandens størrelse, sammensætning og bevis for samliv, og er designet til at finde »statistiske afvigere«; det vil sige ydelsesmodtagere, hvis omstændigheder afviger tilstrækkeligt fra »normen«. Algoritmerne risikerer at diskriminere grupper på baggrund af deres race og etnicitet, klasse og samlivsstatus, fordi modellerne er forankret i sociale normer, der afspejler samfundets dominerende gruppers syn på, hvad en dansk husstand eller en familie er. De tager ikke højde for kontekstbetingede faktorer, som allerede eksisterende uligheder i det danske samfund, herunder klasse- og handicapbaserede uligheder, samt forskellige kulturelle normer, der konstant udvikler sig; eller forskelle mellem forskellige grupper, der påvirker levevilkår og husstandssammensætning. Hvis der ikke tages højde for relevante kontekstbetingede faktorer, eksisterende samfundsmæssige uligheder, og forskellige kulturelle normer i udvikling, kan brugen af disse modeller være uforholdsmæssigt indgribende overfor lavindkomstgrupper, mennesker med handicap, racialiserede mennesker, migranter og ældre mennesker. UDK/ATP bør derfor genoverveje hvordan de fremover bruger data om »usædvanlige« opholds-, familie- eller husstandsmønstre, som grundlag for bedragerikontrol.

TILKNYTINING ELLER FORBINDELSER TIL UDLANDET

Ud over at afdække usædvanlige mønstre i husstandssammensætningen, har Amnesty International fundet ud af, at input relateret til »udenlandsk tilknytning« bruges af UDK/ATP som en del af de algoritmiske profileringsmodeller, især for så vidt angår pensioner og fordeling af børneydelser. Det skyldes, at UDK/ATP er bekymret for, at ydelsesmodtagere bor i udlandet uden at informere myndighederne, og at de uretmæssigt tager deres velfærdsydelser med sig. For at forsøge at kontrollere for dette, bruger UDK/ATP en algoritme, der kaldes »Udlandsmodellen«. »Udlandsmodellen« genererer en score - en vurdering af - en modtagers ›udenlandske tilknytning‹ ved at skabe et relativt mål for hvor stærk en persons tilknytning er til de forskellige lande. Den dokumentation, der blev udleveret efter en forespørgsel i henhold til offentlighedsloven, giver kun delvis svar på modellens input og specifikke anvendelsesområder. Selv om det angives, at »udenlandsk tilknytning« ikke vil blive brugt direkte til at udpege sager om potentiel svindel, bruges det ikke desto mindre til at forfine søgningen efter sager, som UDK vil udtage til undersøgelse om svindel, og fungerer derfor som en de facto-indikator. Mere specifikt, bruges modellens output til at identificere grupper af støttemodtagere, der anses for at have »medium og stærk tilknytning« til lande uden for EØS, og prioriterer disse grupper til svindelefterforskning. Dette er opbygget som en relativ metrik, hvilket betyder, at »mellemstore og stærke bånd« er fastsat i forhold til andre modtagere af sociale ydelser, i stedet for at være defineret ud fra objektive kriterier. Dette understreger både hvordan algoritmens design bevirker at modtagerne vurderes i forhold til »normen« eller den socialt dominerende gruppe i Danmark, og samtidig understreger det behovet for større gennemsigtighed, idet algoritmen og output er en selvkonstrueret metrik (i modsætning til at være en metrik taget fra akademisk litteratur, som er blevet statistisk valideret af en gruppe eksperter). UDK bruger data indsamlet af Den Fælles Dataenhed Udland til brug for algoritmen; om borgeres udenlandske bopæl, ind- og udrejse i udlandet, civilstand, antal børn, fast ejendom eller køretøjer i udlandet, samt modtagne sociale ydelser. Brugen af statsborgerskab og andre kriterier for udenlandsk tilhørsforhold, bruges eksplicit mod mennesker fra lande uden for EØS, og er derfor udtryk for direkte diskrimination på baggrund af nationalitet, etnicitet og migrationsstatus. Dette krænker retten til ligestilling og ikke-diskrimination for racialiserede grupper. Det risikerer også at gribe ind i folks ret til social sikkerhed. Desuden afspejler datainput, der bruges til at skabe, udvikle og drive AI-systemer, ofte en historisk, systemisk, institutionel og samfundsmæssig diskrimination. Derfor risikerer indførelsen af algoritmer til bedragerikontrol at fastholde den historiske uretfærdighed mod marginaliserede samfund, herunder dem, der lever i fattigdom. Flere af modellerne til bekæmpelse af svindel, omfatter direkte input vedrørende en persons løn og indkomst. Ud fra den redigerede dokumentation, som Amnesty International har modtaget, er det ikke muligt at fastslå den præcise effekt af at medtage indkomst- eller lønindikatorer i algoritmerne; men hvis de inkluderes, vil systemerne skelne mellem modtagere af sociale ydelser på dette grundlag, og det medfører en uacceptabel risiko for, at modtagere med lavere indkomster forfordeles og eksplicit vælges ud. Derudover giver modellernes analytiske integritet anledning til betydelig bekymring , hvad enten det skyldes de input, der er indeholdt i modellen, brugen af selvkonstruerede og subjektive metrikker, eller i forhold til hvor repræsentative de datasæt er, som de er udviklet fra.

Tekniske evalueringer og revisioner er blevet et stadig mere populært værktøj til at vurdere algoritmers ydeevne og effekt, samt til at diagnosticere problematisk adfærd. Mens UDK leverede redigeret dokumentation om nogle få af deres svindelkontrolmodeller, blev alle anmodninger om data, der ville gøre det muligt for researchere at udføre bias- og fairness-tests, afvist. UDKs afvisninger viser en mangel på gennemsigtighed i forhold til at sikre tilgængeligheden af oplysninger, der ville gøre det muligt at granske og teste deres algoritmer.

DIGITAL EKSKLUSION OG TVUNGEN INKLUSION

Amnesty Internationals research har vist, at automatisering og digitalisering af ydelsessystemet, ikke kun gør det muligt at overvåge og kontrollere ydelsesansøgere og -modtagere, men at systemet også skaber en adgangsbarriere til sociale ydelser for visse marginaliserede grupper, herunder for kvinder på krisecentre og mennesker med handikap. Digitaliseringen risikerer også at udelukke ældre mennesker. Som følge heraf, risikerer systemet at begrænse disse gruppers ret til social sikring og ikke-diskrimination.

Amnesty International samarbejdede med LOKK (Landsorganisationen af Kvindekrisecentre i Danmark), som repræsenterer 46 kvindekrisecentre rundt om i Danmark, om at udforme en undersøgelse af UDK's systemets tilgængelighed for kvinder, der bor på krisecentre på grund af partnervold.

Amnesty Internationals research viser, at digitaliseringen af Danmarks sociale ydelsessystem har ført til, at mennesker med handicap, tvangsmæssigt eller ufordelagtigt er blevet en del af UDK's system. Ufordelagtig

inklusion defineres som »at blive tvunget til at blive inkluderet på dybt ufordelagtige vilkår«. Mennesker med handicap er tvangsmæssigt eller ufordelagtigt inkluderet, fordi de ikke har andet valg end at dele deres oplysninger med tredjeparter, for at få adgang til ydelsessystemet. Denne form for inklusion giver anledning til bekymringer om databeskyttelse og sikkerhed, fordi der er en risiko for, at de sagsbehandlere, der har adgang til deres personlige oplysninger kan misbruge dem. Adgang til systemerne bør derfor ikke udelukkende være digital; myndighederne bør tilbyde realistiske alternativer, der er inkluderende, tilgængelige og ikke-diskriminerende for de dårligst stillede og marginaliserede grupper, således at alle grupper i samfundet er i stand til at modtage sociale tilbud, uden at deres privatliv og databeskyttelse kompromitteres.

MANGLENDE STATSLIGT TILSYN, MANGLENDE GENNEMSIGTIGHED OG RISIKO FOR MANGLENDE ADGANG TIL EFFEKTIVE RETSMIDLER

Amnesty Internationals research viser, at den danske stat har uddelegeret fordelingen af ydelser til ATP, et selskab etableret som en selvejende institution under ATP-loven fra 1964. Amnesty International konstaterer også, at der er ikke er et tilstrækkeligt, uafhængigt tilsyn med UDK/ATP's data- og algoritmiske praksis, hvilket øger risikoen for menneskerettighedskrænkelser. Det er tydeligt, at der er huller i det eksisterende tilsyn med UDK/ATP-styringsstruktur, og manglen på proaktive undersøgelsesbeføjelser hos det danske datatilsyn, betyder at der er klare mangler i den danske stats pligt til at respektere og beskytte menneskerettighederne. Denne pligt består i at sikre, at der er et effektivt tilsyn med den offentlige myndighed UDK, samt med den selvejende institution ATP, som står til ansvar overfor staten. Endvidere finder Amnesty International, at ATP ikke ser ud til at gennemføre træning i at modvirke "automatiseringsbias" (= når sagsbehandlere stoler overdrevent på de automatiserede systemer), eller diskrimination for sit personale, offentliggøre konsekvensanalyser af databeskyttelse, eller gennemføre tilstrækkelige revisioner af sine algoritmer til bekæmpelse af bedrageri. Disse foranstaltninger ville kunne identificere og mindske risikoen for potentiel skade i forbindelse med brugen af algoritmiske systemer. Derudover har Amnesty International fundet, at ATP ikke udfører menneskeretlig rettidig omhu ("due diligence") for at identificere, afbøde og forhindre de skadelige virkninger af UDK/ATP's ydelsessystem, og dermed respekterer de ikke de internationale menneskerettighedsstandarder på området.

For så vidt angår de berørte personers adgang til effektive retsmidler, har Amnesty International indkredset, at der især på to områder er problemer. For det første mangler der gennemsigtighed og tydelig information om UDK/ATP's brug af de bedrageri-algoritmer, der bruges til at udvælge hvilke personer, der skal undersøges yderligere. For det andet indeholder forvaltningsloven ikke bestemmelser, der pålægger offentlige myndigheder at oplyse en person om, at sagen mod dem er baseret på en algoritme. Som følge heraf har en person, der er udvalgt til bedragerikontol af UDK/ATP's algoritmer, ikke mulighed for at vide, at de er genstand for en automatiseret proces, og de kan derfor ikke effektivt anfægte UDK/ATP's beslutninger.

KOMMENDE FORPLIGTELSER UNDER EU'S FORORDNING VEDRØRENDE KUNSTIG INTELLIGENS (AI-FORORDNINGEN).

Amnesty International mener, at de beviser, der er indsamlet til denne undersøgelse indikerer, at UDK/ATP's algoritmiske modeller bør falde ind under forbuddet mod "social scoring" (social bedømmelse ved hjælp af algoritmer) i forordningens artikel 5, stk. 1, litra c), hvilket ville betyde, at systemet skulle forbydes. Amnesty International mener, at dette system er en "social scorealgoritme", fordi systemet udpeger et eksplicit sæt af målinger, som udgør »sociale scores«, fordi de er relateret til en berørt persons troværdighed, dvs. sandsynligheden for at personen vil begå bedrageri. Systemet evaluerer og klassificerer fortsat mennesker, baseret på oplysninger om deres sociale adfærd eller personlige karakteristika, som ikke har noget at gøre med det oprindelige formål med at indsamle oplysningerne - og som fører til at de stilles ufordelagtigt, fordi de udtages til bedragerieftерforskning.

Medmindre UDK og ATP kan fremlægge tilstrækkelige beviser for det modsatte, mener Amnesty International, at systemet i sin nuværende form bør indstilles, indtil UDK og ATP fremlægger tilstrækkelige beviser for, at deres praksis ikke udgør "social scoring".

Amnesty International skrev til UDK og ATP, og forklarede hvorfor vi mener, at deres modeller til bekæmpelse af svindel udgør et socialt scoringssystem som beskrevet i EU's AI-forordning, og udbad os et svar. Amnesty International bad også UDK og ATP om at bevise med tilstrækkelig tydelighed, hvis de ikke mener at modellerne falder ind under definitionen på et socialt scoringssystem. UDK svarede, at dets algoritmiske praksis ikke udgør social scoring i henhold til artikel 5 i EU's AI-forordning, idet kontrollerne har et klart defineret formål, er forholdsmæssige, og har til formål at sikre korrekt betaling af sociale ydelser, samt fordi dets bedragerikontrol overholder gældende EU-lovgivning og national lovgivning. UDK og ATP har dog ikke givet Amnesty International detaljeret dokumentation eller uafhængige vurderinger af, at deres algoritmiske praksis ikke er et socialt scoringssystem i henhold til artikel 5 i EU's AI-forordning, og de har heller ikke givet os nogen dokumentation for, at deres praksis er nødvendig og forholdsmæssig. Den specifikke fortolkning af artikel 5, herunder forbuddet mod social scoring, vil blive afklaret i Europa-Kommissionens kommende vejledning i, hvad der udgør forbudt praksis i henhold til artikel 5. Europa-Kommissionen bør præcisere, at risikoscoringsalgoritmer, der fører til diskrimination af berørte personer, såsom UDK's bedrageri-algoritme, er forbudt i henhold til forordningen.

Alle algoritmiske systemer, der anvendes af UDK/ATP til at opdage svindel med sociale ydelser, og som vurderes til ikke at udgøre social scoring, klassificeres som højrisikosystemer i henhold til EU's forordning om kunstig intelligens 2024 (AI-forordningen), som trådte i kraft den 1. august 2024. I henhold til bilag III er højrisikosystemer » AI-systemer, der tilsigtes anvendt af offentlige myndigheder eller på vegne af offentlige myndigheder til at vurdere fysiske personers berettigelse til væsentlige offentlige bistandsydelser og -tjenester, herunder sundhedstjenester«. Som følge heraf, er de danske myndigheder som minimum forpligtet til at sikre, at UDK's og ATP's algoritmiske systemer til afsløring af bedrageri, opfylder kravene til risikostyring, gennemsigtighed og andre bestemmelser, der stilles til "deployer" (en virksomhed eller institution, der udruller systemet) og udbydere som beskrevet i AI-forordningen. Ovenstående bestemmelser om højrisikosystemer gælder først fra den 2. august 2030. Ikke desto mindre anbefaler Amnesty International, at de danske myndigheder sikrer, at de bliver implementeret så hurtigt som muligt for at sikre større gennemsigtighed i brugen af højrisikosystemer.

SVAR FRA MYNDIGHEDER OG VIRKSOMHEDER

Amnesty International delte resultaterne af vores research med UDK/ATP den 17. oktober 2024 og Beskæftigelsesministeriet (Styrelsen for Arbejdsmarked og Rekruttering eller STAR) den 18. oktober 2024 forud for rapportens offentliggørelse, og udbad os deres skriftlige svar på bestemte påstande i rapporten. UDK svarede på påstandene i rapporten den 30. oktober 2024 og den 1. november 2024, mens Beskæftigelsesministeriet svarede den 1. november 2024. Amnesty International skrev også til virksomheden NNIT den 23. oktober 2024 om deres omtale i denne rapport, og bad om deres svar på yderligere spørgsmål. NNIT svarede den 1. november 2024. Alle svar er, hvor det er relevant, afspejlet i rapportens tekst.

I deres svar afviste de danske myndigheder bestemte konklusioner i denne rapport. Hvor det er relevant, er deres svar afspejlet i rapportens fulde tekst.

CENTRALE ANBEFALINGER

Researchen i nærværende rapport bygger på Amnesty Internationals tidligere publikationer, herunder *Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal* (2021), *Trapped by Automation: Poverty and Discrimination in Serbia's Welfare State* (2023), og *Use of Entity Resolution in India: Shining a Light on How New Forms of Automation Can Deny People Access to Welfare* (2024).

VI ANBEFALER, AT DE DANSKE MYNDIGHEDER

- Sikrer at Udbetaling Danmark/ATP indstiller brugen af algoritmer, der bedømmer eller klassificerer mennesker baseret på oplysninger om deres sociale adfærd eller følsomme personlige karakteristika, eller "stedfortrædere" (proxy) herfor, på en måde som fører til at deres menneskerettigheder krænkes.

- Sætter det nuværende system på pause, indtil der er tilvejebragt tilstrækkelig dokumentation til at foretage en gennemgribende og endelig vurdering af systemet, samt af hvorvidt forbuddet mod "social scoring" som defineret i AI-forordningen, kan anvendes.
- Sikrer en stærk national implementering af AI-forordningen i respekt for menneskerettighederne, samt en stærk, effektiv og charterbaseret fortolkning af forbudte og risikable teknologier i henhold til AI-forordningen. Det bør ske så hurtigt som muligt, og senest inden for de lovbestemte frister, herunder bestemmelserne i artikel 9, 10, 11, 13, 26, 27, 85, 86 og 87 i EU's AI-forordning, for at kunne håndtere de risici, som højrisikosystemer kan udgøre for de grundlæggende rettigheder, og for at sikre større gennemsigtighed i brugen af højrisikosystemer.

VI ANBEFALER, AT DET DANSKE BESKÆFTIGELSESMINISTERIUM OG UDBETALING DANMARK

- Etablerer et klart, utvetydigt og juridisk bindende forbud mod brug af data om statsborgerskab, »udenlandsk tilknytning« eller nationalitet, eller "stedfortrædere" (proxy) for disse, når der udføres risikovurderinger ved bedragerikontrol.
- Gennemgår og ændrer Udbetaling Danmark/ATP's standarder, politikker og love, der understøtter risikoprofilering i Udbetaling Danmark/ATP's bedrageri-algoritmer, og som kan fastholde diskrimination baseret på indkomst, race, etnicitet, religion, migrationsstatus, køn, handicap eller alder, samt sikrer, at de er i overensstemmelse med relevante internationale menneskerettighedsstandarder.
- Sikrer, at Udbetaling Danmark/ATP stopper deres praksis med masseudtræk, -behandling og -udnyttelse af borgernes data til svindelkontrol, samt brugen af sociale medier.
- Sikrer, at der er fuld gennemsigtighed i Udbetaling Danmark/ATPs arbejde, og at de giver fyldestgørende information til de berørte personer om den underliggende logik, betydning og forventede konsekvenser af beslutninger, selv om de ikke er fuldt automatiserede, uanset graden af menneskelig involvering i beslutningsprocessen.
- Sikrer, at Udbetaling Danmark/ATP og kommunerne på en klar, forståelig og detaljeret måde informerer ydelsesansøgere og -modtagere om, at en algoritme har udtaget dem til bedragerikontrol.
- Sikrer, at Udbetaling Danmark/ATP offentliggør tydelige oplysninger om de input til bedragerikontrol, som de bruger til at foretage risikovurderinger for bedrageri og fejl (herunder offentliggørelse af regelmæssige rapporter og nøglestatistikker); oplysninger om, hvordan bedrageri-modellerne fungerer; oplysninger om gennemførte evalueringer af personalets optræden og fordomme; samt oplysninger om, at der er foretaget analyser af hvordan menneskerettighederne og retten til databeskyttelse bliver påvirket, både før og under brugen af systemer til bedragerikontrol, herunder mens oplysningerne behandles i systemerne.
- Sikrer, at Udbetaling Danmark/ATP og kommunerne gennemfører uafhængige konsekvensanalyser af UDK-systemets indvirkning på menneskerettighederne, og på databeskyttelsen. Disse konsekvensanalyser bør som minimum omfatte en evaluering af hvordan brugen af bedrageri-algoritmer kan diskriminere marginaliserede grupper, herunder lavindkomstgrupper, racialiserede grupper (inklusive migranter og mennesker, der har fået flygtningestatus i Danmark), etniske minoriteter, mennesker med handicap og ældre mennesker.
- Sikrer, at Udbetaling Danmark/ATP og kommunerne sørger for, at sagsbehandlere modtager yderligere uddannelse og kapacitetsopbygning efter behov, så de er i stand til at håndtere og forebygge problemer med diskrimination, f.eks. "automatiseringsbias".
- Sikrer, at Udbetaling Danmark sørger for at kvinder på krisecentre, ældre og mennesker med handicap ikke længere ekskluderes som følge af digitaliseringen af Udbetaling Danmarks ydelsessystem, ved at sikre, at systemet i praksis er fuldt tilgængeligt via ikke-digitale midler for grupper, der ikke kan bruge den nødvendige teknologi.
- Sikrer, at Udbetaling Danmark/ATP giver personer, der ansøger om sociale ydelser tydelige og tilgængelige oplysninger om, hvordan der træffes afgørelser i deres sager, og om hvordan de kan anke sådanne afgørelser. Endvidere bør det sikres at ansøgere får støtte til at anke afgørelsen, herunder juridisk og/eller økonomisk støtte, hvor det er nødvendigt.
- Kræver at virksomheder, der udvikler AI-produkter, sørger for rettidig omhu på menneskerettighedsområdet, så de har mulighed for at identificere og håndtere de faktiske eller potentielle overtrædelser af menneskerettighederne, der kan opstå på ethvert tidspunkt i

forsyningskæden eller produktets livscyklus, som beskrevet i FN's Vejledende Principper for Menneskerettigheder og Erhvervsliv.

VI ANBEFALER, AT DET DANSKE FOLKETING

- Reviderer og ændrer § 2, stk. 1, nr. 7, i bekendtgørelse af lov om en børne- og ungeydelse LBK nr. 724 af 25/05/2022 og § 5 a i bekendtgørelse af lov om børnetilskud og forskudsvis udbetaling af børnebidrag (LBK nr. 63 af 21/01/2019) for at fjerne urimelige og langvarige opholdskrav, der begrænser adgangen til børneydelser for personer, der har fået flygtningestatus i Danmark.
- Gennemgår og ændrer den danske forvaltningslov, så den indeholder bestemmelser om automatiseret beslutningstagning, der garanterer, at ansøgere og personer, der vil anke en afgørelse har adgang til effektive retsmidler.
- Vedtager lovgivning, der etablerer en uafhængig offentlig myndighed, der skal føre tilsyn med UDK/ATP, og som overvåger UDK/ATP's brug af AI-systemer, for at styrke ansvarligheds-mekanismer og øge beskyttelsen af menneskerettighederne. Dette omfatter etablering af en uafhængig myndighed, der har tilsyn med Udbetaling Danmark/ATP's aktiviteter i overensstemmelse med artikel 70 i EU's AI -forordning.

VI ANBEFALER, AT DATATILSYNET

- Udøver sin tilsynsmyndighed i henhold til artikel 29 i den danske databeskyttelseslov og artikel 58 i GDPR, og pålægger Udbetaling Danmark/ATP og kommuner at udlevere oplysninger om deres datapraksis, samt eventuelt gennemførte konsekvensanalyser af databeskyttelse.
- Sikrer, at Udbetaling Danmark/ATP og kommunerne overholder alle relevante bestemmelser i databeskyttelsesloven og databeskyttelsesforordningen, herunder artikel 5 og 6 om behandling af oplysninger, der er defineret i disse bestemmelser.

VI ANBEFALER, AT KOMMUNERNE

- Sørger for at sagsbehandlere modtager yderligere uddannelse og kapacitetsopbygning, således at de er i stand til at håndtere og forebygge problemer som "automatiseringsbias", diskrimination og krænkelse af velfærdsmodtageres værdighed og privatliv, i deres arbejde med ydelsesmodtagere.
- Gennemfører uafhængige konsekvensanalyser for at sikre menneskerettighederne og databeskyttelsen i arbejdet med at undersøge bedrageri. Konsekvensanalyser bør som minimum omfatte en evaluering af algoritmers diskriminerende effekt på bedragerikontrol af marginaliserede grupper, herunder lavindkomstgrupper, racialiserede grupper og mennesker med handicap.
- Sikrer, at personer, der ansøger om at modtage sociale ydelser, får tydelig og tilgængelig information om, hvordan der træffes beslutninger i deres sager, hvordan man anker sådanne beslutninger, samt, om nødvendigt, sørger for, at ansøgere får støtte til at anke en afgørelse, herunder juridisk og/eller økonomisk støtte.

VI ANBEFALER, AT EUROPA-KOMMISSIONEN

- Sikrer, at den kommende vejledning fra Europa-Kommissionen om den praktiske gennemførelse af de forbudte former for praksis, der henvises til i AI-forordningen, skaber juridisk klarhed, og adresserer relevante AI-baserede former for social scoringspraksis i hele EU, herunder diskriminerende systemer til afsløring af bedrageri, samt risikoprofilering i sager, der drejer sig om social beskyttelse.

VI ANBEFALER, AT ATP

- Tager omgående skridt til at sikre, at ATP ikke bidrager til menneskerettighedskrænkelser eller-misbrug gennem sin involvering i UDK's ydelsessystem, og adresserer eventuelle menneskerettighedskrænkelser, når de forekommer, herunder om nødvendigt ved at samarbejde om at afhjælpe dem.
- Fremlægger bevis for, at sagsbehandlere i bedrageri-enhederne har den nødvendige kompetence og autoritet til at gribe ind i undersøgelsen og beslutningsprocesserne, når en person udtages til undersøgelse for bedrageri af UDK/ATP's algoritmer.

- Giver sagsbehandlere den nødvendige træning og kapacitetsopbygning, så de er rustet til at håndtere og forebygge problemer som "automatiseringsbias" og diskrimination.
- Foretager proaktiv og løbende " due diligence" på menneskerettighedsområdet i hele de algoritmiske teknologiers livscyklus, - både før og efter udrulning og implementering af nye systemer – således at de kan indkredse eventuelle risici i udviklingsfasen, og straks kan opfange menneskerettighedskrænkelser og andre skader, når teknologierne er blevet implementeret.
- Offentliggør de tiltag, institutionen har iværksat, for at opdage, forebygge og afbøde menneskerettighedskrænkelser og -risici i sine forretningsaktiviteter, herunder gennem sit engagement og forretningsforbindelser med UDK.

AMNESTY INTERNATIONAL ER EN GLOBAL MENNESKERETTIGHEDS-BEVÆGELSE. HVIS ÉN PERSON BLIVER URETFÆRDIGT BEHANDLET, BETYDER DET NOGET FOR OS ALLESAMMEN.

KONTAKT OS

amnesty@amnesty.dk

+45 33 45 65 65

TAG DEL I SAMTALEN

www.facebook.com/Amnesty.dk

@Amnestydk

INDKODET URETFÆRDIGHED

OVERVÅGNING OG DISKRIMINATION I DANMARKS AUTOMATISEREDE VELFÆRDSSYSTEM

Denne rapport undersøger brugen af bedrageri-algoritmer, baseret på kunstig intelligens (AI) i det danske velfærdssystem. Velfærdssystemet, Udbetaling Danmark, underminerer menneskerettighederne ved at bruge analoge og digitale former for overvågning til at afsløre svindel med sociale ydelser. Analoge overvågningsmetoder omfatter andre borgeres, kommuners og andre offentlige myndigheders kontrol, mens digitale former for overvågning omfatter brug af bedrageri-algoritmer, der bruges til at udvælge visse personer til yderligere undersøgelser. Denne gennemgribende overvågning begrænser ydelsesmodtagernes ret til privatliv, samt til menneskelig værdighed og social sikkerhed.

UDK's algoritmer bruger oplysninger som husstandens størrelse, sammensætning og bevis for samliv til at finde personer og grupper, hvis omstændigheder afviger fra »normen«, og som sandsynligvis vil begå bedrageri. Disse karakteristika eller variabler kan fungere som "stedfortræder" (proxy) for race, migrationsstatus og socioøkonomisk status og kan tilskynde til diskrimination.

Der forekommer også diskrimination som følge af digitaliseringen af det danske ydelsessystem. Det skyldes, at det for nogle marginaliserede grupper, herunder kvinder i krise og mennesker med handicap, ikke er muligt at få adgang til digitale tjenester på egen hånd, og derved risikerer digitalisering at begrænse deres ret til social sikring på grund af deres køn og handicap.

INDEX: EUR 18/8738/2024
NOVEMBER 2024
SPROG: DANSK

amnesty.org

AMNESTY INTERNATIONAL