



**Diskriminerings
ombudsmannen**

Rapport 2022:1

Transparens, träning och data

**Myndigheters användning av AI och automatiserat
beslutsfattande samt kunskap om risker för diskriminering**



Transparens, träning och data

Myndigheters användning
av AI och automatiserat
beslutsfattande samt kunskap
om risker för diskriminering

Denna pdf är tillgänglig. Andra alternativa format, exempelvis Daisy, lättläst eller punktskrift, kan beställas vid behov. Kontakta då DO på order@do.se.

Diskrimineringsombudsmannen (DO)

© DO

R19 2022 PDF

ISBN 978-91-88175-27-4

Illustration: Stina Wirsén.

Layout: Kollijox AB.

Förord

Artificiell intelligens (AI) och annat automatiserat beslutsfattande används allt mer inom allt fler områden i samhällslivet. Ämnet har också hamnat högre upp på de europeiska likabehandlingsorganens dagordning, och flera av dessa har uppmärksammat att tillämpningen av AI och automatiserat beslutsfattande kan leda till risker för diskriminering och hinder för lika rättigheter och möjligheter. Mot denna bakgrund behöver DO mer kunskap inom området och verka för att andra aktörer systematiskt beaktar diskrimineringsperspektivet i arbetet med att implementera AI och automatiserat beslutsfattande.

I denna rapport har DO undersökt hur statliga förvaltningsmyndigheter använder AI och automatiserat beslutsfattande gentemot enskilda, och i vilken utsträckning som de beaktar risker för diskriminering och hinder för lika rättigheter i tillämpning av AI och automatiserat beslutsfattande. Rapporten ska ses som ett led i att öka medvetenheten om diskrimineringslagen och på vilka sätt som automatiserade beslut kan leda till diskriminering och utgöra hinder för lika rättigheter och möjligheter.

Solna den 17 februari 2022

Lars Arrhenius

Diskrimineringsombudsman

Innehåll

Förord	5
Summary	8
Inledning	13
Syfte och frågeställningar	14
Automatiserade beslutsprocesser och definitioner	15
Tidigare studier av automatiserat beslutsfattande i myndigheter	18
Likabehandlingsorgan uppmärksammar risker för diskriminering	21
Risker vid bristande kvalitet på använd data	23
Risker med bristande kvalitet vid träning av AI	24
Risker med bristande transparens och svarta lådor	24
Resultat	26
Statliga myndigheters användning av automatiserat beslutsfattande	26
Kunskap om risker för diskriminering och hinder för lika rättigheter och möjligheter	30

Avslutande iakttagelser	35
Referenser	39
Bilaga 1. Metod och urval	43
Bilaga 2: Enkätfrågor	47
Bilaga 3: Intervjuguide	56
Bilaga 4. Intervjuförteckning	57

Summary

Artificial intelligence (AI) and other types of automated decision-making are increasingly being used in ever more areas within society. This topic has also moved higher up on the agendas of equality bodies in Europe, with more of these having drawn attention to how the application of AI and automated decision-making may lead to risks of discrimination and barriers to equal rights and opportunities. In light of this, DO needs to know more about this area and to take action to ensure that other actors are systematically taking into account the discrimination perspective in their efforts to implement AI and automated decision-making.

In this report, DO has investigated how government agencies are using AI and automated decision-making and the extent to which they are taking into account risks of discrimination and barriers to equal rights in their application of AI and automated decision-making. The report is to be seen as a way to increase awareness of the Discrimination Act and the ways in which automated decisions can lead to discrimination and constitute barriers to equal rights and opportunities.

Aim

The aim of the survey is to gain a better understanding of how widespread the use of AI and automated decision-making is among government agencies and what knowledge these agencies have about conditions that can lead to discrimination or constitute barriers to equal rights and opportunities when it comes to automated decision-making.

The survey is based on the following questions:

1. How widespread are various technologies within AI and automated decision-making at government agencies?
2. What knowledge do these government agencies have about risks of discrimination and barriers to access to equal rights and opportunities when designing and applying AI and automated decision-making?

Method

The report is based on reports from other discrimination bodies, a questionnaire and interviews with a sample of government agencies in Sweden. The sample has been based on agencies making decisions towards individuals and where we have found it to be probable that these agencies are using automated decision-making. The questionnaire was sent to 34 of the 249 government agencies in Sweden. The questionnaire was answered by 33 government agencies. The government agencies that participated in the survey vary in size and area of activity. In addition to the questionnaire, the report is also based on interviews with 18 people from six government agencies.

Results

The survey shows that automated decision-making is being used at 14 of the 34 government agencies questioned and that the automated decisions concern a large number of individuals. DO's assessment is that the government agencies that were interviewed have some insight into risks of discrimination that may be present, but that they primarily talk about ethics and integrity rather than a risk of discrimination in relation to automated decision-making. The government agencies describe how they actively consider GDPR, for example, but that they do not, to any great extent or systematically, take into account the discrimination perspective in their efforts to implement automated decision-making.

It also emerges from the survey that government agencies which use automated decision-making conduct risk analyses and quality assurance follow-ups but that these rarely consider the grounds of discrimination. Few authorities have policy documents for ensuring that individuals are not disadvantaged by automated decision-making. We conclude that a discrimination perspective is largely absent from the automated decision-making processes at the government agencies. In addition, few government agencies see the need to do more in order to reduce the risk of individuals being disadvantaged. Nevertheless, those government agencies we have interviewed stress that they would like to see an exchange of knowledge on these issues with DO.

DO is able to conclude that the government agencies are aware of the three most prominent risks of discrimination when applying AI and automated decision-making that have been pointed out by other discrimination bodies:

1. Risks when the data quality is insufficient.
2. Risks when the AI is insufficiently trained.
3. Risks when there is a lack of transparency.

However, the survey does not show how the government agencies are working to prevent these risks. On the basis of the questionnaire and interviews, DO is however able to conclude that the authorities need to increase their knowledge about risks of discrimination when applying automated decision-making.

Based on the survey, there are some observations that we would like to highlight:

- The government agencies have rarely considered the grounds of discrimination in policy documents, risk analyses and quality assurance follow-ups.

- A few of these fourteen government agencies have policy documents that are there to ensure individuals are not disadvantaged. Only four state that they consider the grounds of discrimination in this work.
- Ten of the fourteen government agencies conduct risk analyses in order to ensure that individuals are not being disadvantaged. Only four state that they consider the grounds of discrimination in this work.
- The majority of government agencies perform quality assurance follow-ups in order to ensure that individuals are not being disadvantaged. Only four state that they consider the grounds of discrimination in this work.
- Only nine of the fourteen government agencies inform the individual that decisions are being made using an automated process.
- The government agencies point to the risk of case officers giving too much weight to the automated decisions and losing the ability to critically appraise them.
- Few government agencies state that they need more knowledge and support with regard to how to reduce the risk of individuals being disadvantaged. However, the six government agencies that were interviewed ask for more knowledge about risks of discrimination and increased expertise in computer programming.

In summary, the survey shows that knowledge about risks of discrimination and barriers to equal rights and opportunities is unsatisfactory. This means that the government agencies need to increase their awareness of the prohibition of discrimination. The survey only encompasses government agencies but automated decision-making is also being used in other parts of the public sector and in the private sector. Consequently, awareness of the risks of discrimination that may arise when applying AI and automated decision-making may also need to be improved in other parts of society.



Inledning

Inom allt fler samhällsområden ökar användningen av artificiell intelligens (AI) och annat automatiserat beslutsfattande i snabb takt. Flera europeiska likabehandlingsorgan har uppmärksammat förekomsten av AI och automatiserat beslutsfattande samt de risker för diskriminering som kan uppstå när algoritmer används vid databearbetning och beslutsfattande.¹ Flera andra aktörer har på senare tid även lyft de etiska risker som är förenade med AI och automatiserat beslutsfattande, däribland EU-kommissionen som i april 2021 kom med förslag till en förordning om harmoniserade regler för artificiell intelligens.²

Ett av Diskrimineringsombudsmannens (DO:s) tidigare projekt var ”Utveckling av kunskap om automatiserad databehandling med algoritmer och stordata och diskriminering eller risker för diskriminering, 2018–2020”, och där konstaterade vi att diskrimineringslagen är tillämplig vid användningen av AI och automatiserat beslutsfattande.³ Vidare framkom att många aktörer är medvetna om etik- och integritetsrisker, men tycks ha lägre kunskap om risken för att bryta mot diskrimineringslagen vid tillämpningen av AI och automatiserat beslutsfattande. Den ökande medvetenheten om etiska risker behöver alltså inte inkludera kunskap om diskrimineringslagens krav.

Expansionen av AI och automatiserat beslutsfattande innebär att DO behöver bevaka och hantera en utveckling som kan leda till diskriminering eller risker för diskriminering. Mot denna bakgrund

1 Se exempelvis: FADA (2020); Equinet (2020); Institut for menneskerettigheter (2021); Kilden Kjønnsforskning.no (2020); Zuiderveen Borgesius (2018).

2 Europeiska kommissionen (2021).

3 Diskrimineringsombudsmannen (2018): LED 2018/387. Resultat från projektet återfinns i Diskrimineringsombudsmannen (2019a; 2019b; 2019c).

behöver DO förstärka sin kunskap för att kunna säkerställa att individer inte diskrimineras eller riskerar att diskrimineras. Det är också angeläget att uppmärksamma andra myndigheter på de risker för diskriminering och hinder för lika rättigheter och möjligheter som finns vid tillämpningen av AI och automatiserat beslutsfattande.

Syfte och frågeställningar

Syftet med denna kartläggning är att få en bättre bild av hur utbredd användningen av AI och automatiserat beslutsfattande är bland statliga myndigheter, och vilken kunskap dessa myndigheter har om förhållanden som kan leda till diskriminering eller utgöra hinder för lika rättigheter och möjligheter när det gäller automatiserat beslutsfattande.⁴

Kartläggningen utgår från följande frågor:

1. Hur utbredd är förekomsten av olika teknologier inom AI och automatiserat beslutsfattande vid statliga myndigheter?
2. Vilken kunskap har dessa myndigheter om risker för diskriminering och hinder för tillgång till lika rättigheter och möjligheter vid utformning och tillämpning av AI och automatiserat beslutsfattande?

4 Se direktiv LED 2020/487.

Målet är att kartläggningen ska kunna användas i DO:s främjande och förebyggande arbete:

1. som underlag inför dialoger och kunskapsspridande insatser med myndigheter och eventuellt andra aktörer⁵
2. som underlag inför planering av eventuella insatser inom detta område
3. för att ge DO uppslag till eventuella fördjupande studier.

Rapporten bygger på en sammanställning av resultat från andra studier om automatiserat beslutsfattande och diskriminering, en enkät till 34 statliga myndigheter och intervjuer med 18 tjänstepersoner på sex myndigheter. För mer information om metod och urval, se bilaga 1.

Automatiserade beslutsprocesser och definitioner

Automatiskt beslutsfattande omfattar i DO:s definition algoritmiska och automatiserade beslutsprocesser, med eller utan AI. Enligt denna definition kan automatiserat beslutsfattande ske både med och utan AI. När AI ingår sker det automatiska beslutsfattandet i system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder, med viss grad av självständighet.⁶

Automatiserat beslutsfattande hos myndigheter kan till exempel användas för att fatta beslut om arbetsmarknadsutbildningar, tillfällig föräldrapenning, inkomstskatt, och ekonomiska ersättningar.

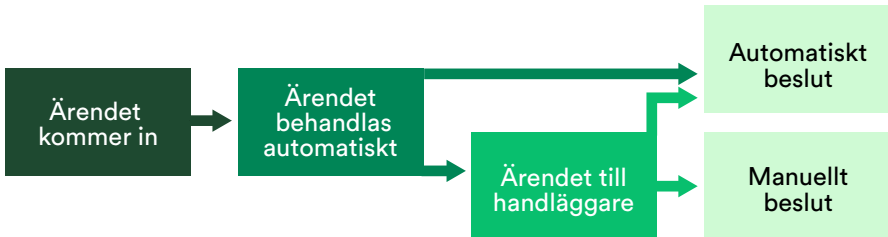
5 Resultaten i kartläggningen ska användas inom uppdraget LED 2020/484: Dialog och kunskapsspridning om AI och risker för diskriminering. Diskrimineringsombudsmannen (2020a).

6 Jämför med Riksrevisionen (2020).

Automatiserade beslutsprocesser förekommer i huvudsak i tre varianter hos myndigheter:

1. helt automatiserat beslutsfattande som oftast handlar om relativt enkla bedömningar
2. automatiserat beslutsfattande för beslutsstöd, där inkommande data bearbetas automatiskt och är en del av beslutsunderlaget men där en handläggare fattar det slutgiltiga beslutet
3. automatiserade processer för att identifiera ärenden. Det kan handla om att identifiera ärenden exempelvis för kontroll eller manuell hantering.⁷

Figur 1: Förenklad illustration över de steg ett ärende går inom myndigheterna till beslut.



Beskrivning: Ett ärende kommer in och behandlas automatiskt. Resultatet kan bli att ärendet fortsätter behandlas automatiskt eller väljs ut för att gå till en handläggare för kontroll. Handläggaren kan efter kontrollen och eventuella justeringar återföra ärendet till en automatiserad beslutsprocess eller fatta ett manuellt beslut.

7 Dessa tre tillvägagångsätt identifierades i de intervjuer som vi har genomfört med myndigheter. Det kan finnas flera tillvägagångsätt, men dessa tre är några av de vanligaste. Jämför Riksrevisionen (2020).

Definitioner av centrala begrepp

Kartläggningen berör inte alla former av digitalisering utan fokuserar på automatiserat beslutsfattande – både helt automatiserat beslutsfattande såväl som automatiserade beslutsprocesser som används till beslutsstöd för handläggare. Den omfattar både AI och regel- och villkorsstyrda former av automatiserade beslut.

Automatiserat beslutsfattande

Automatiskt beslutsfattande omfattar i vår definition algoritmiska och automatiserade beslutsprocesser, med eller utan AI. I begreppet inkluderar vi både helt automatiserat beslutsfattande och automatiserade processer som används som beslutstöd för handläggare.

Artificiell intelligens

Artificiell intelligens omfattar i vår definition system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder – med viss grad av självständighet – för att uppnå särskilda mål.

Profilering

Profilering är en automatiserad personuppgiftsbehandling som avser att framkalla information om en enskild persons egenskaper, beteende och personliga preferenser. Automatiserat beslutsfattande kan ske med eller utan profilering.

Missgynnande

Ett missgynnande innebär att ett beslut får en negativ följd för en individ, det vill säga en skada, en nackdel eller ett obehag. Det kan handla om att en person kommer i ett sämre läge eller går miste om en förbättring, en förmån eller service. Det spelar ingen roll om det finns en avsikt att missgynna eller inte; det är effekten eller resultatet som avgör om det är ett missgynnande.

Tidigare studier av automatiserat beslutsfattande i myndigheter

Det finns få tidigare utredningar som beskriver förekomsten av automatiserat beslutsfattande och AI inom offentlig sektor i Sverige. Vi har tagit del av tre utredningar som rör hur svenska myndigheter tillämpar AI, automatiskt beslutsfattande och beslutsstöd: Statistiska centralbyråns (SCB:s) rapport *Artificiell intelligens i Sverige (2019)*⁸, Myndigheten för digital förvaltning (DIGG:s) rapport *Främja den offentliga förvaltningens förmåga att använda AI (2020)*⁹ och Riksrevisionens granskningsrapport *Automatiserat beslutsfattande i statsförvaltningen – effektivt, men kontroll och uppföljning brister (2020)*.¹⁰

Rapporterna visar en begränsad förekomst av AI och automatiserat beslutsfattande i myndigheter. SCB rapporterar att 10,2 procent av statliga, regionala och kommunala myndigheter använde någon form av AI under 2019.¹¹ I rapporten konstateras att större delen av användandet återfanns inom kommunsektorn och regionernas hälso- och sjukvård.¹² Även DIGG:s kartläggning rapporterar en begränsad förekomst bland statliga myndigheter.¹³ Bland de 14 statliga myndigheter som ingick fann DIGG drygt 10 AI-lösningar i produktion, drygt 20 AI-satsningar under utveckling och ytterligare drygt 10 som planerades. I rapporten

8 SCB (2019). Se även en EU-övergripande undersökning 2021 vilken visar att Sverige då hade fem AI-applikationer på nationell nivå, Europeiska kommissionen, AI Watch (2020).

9 Myndigheten för digital förvaltning (2019).

10 Riksrevisionen (2020).

11 SCB (2019).

12 SCB undersökte enbart AI, inte andra former av automatiserat beslutsfattande.

13 Myndigheten för digital förvaltning (2020).

konstaterar DIGG att den offentliga sektorn i stort har effektiviserats i och med införandet av AI, men att arbetet med frågor om rättssäkerhet och etik behöver utvecklas.

Riksrevisionen granskade i vilken mån statliga myndigheters automatiserade beslutsfattande är effektivt och rättssäkert, och man identifierade tretton myndigheter som tillsammans har 112 miljoner helt automatiserat beslutsprocesser. Totalt fattades 137 miljoner automatiserade beslut 2019. Av dessa var det sex myndigheter som använde AI. Men enbart tre av dessa sex använde AI för automatiserat beslutsfattande. De tre myndigheterna var Arbetsförmedlingen, Skatteverket och Universitets- och högskolerådet. Riksrevisionen konstaterade i rapporten att användningen är effektiv och att vissa aspekter av rättssäkerhet, såsom likabehandling, har förbättrats jämfört med manuell hantering av ärenden. I Riksrevisionens granskning framkommer också en del brister när det gäller myndigheternas hantering av ärenden med hög risk för fusk och fel, med en alltför begränsad uppföljning av automatiserade beslut för att säkerställa att besluten är korrekta. De ser även att myndigheterna riskerar att inte behandla liknande fall på samma sätt i kontrollarbetet. Utöver detta anser Riksrevisionen att dokumentationen av den automatiserade handläggningen inte är tydlig och läsbar. En kritisk faktor enligt Riksrevisionen är även behovet av kompetens för att kunna översätta lagstiftning till programkod. Flera av de myndigheter som ingick i Riksrevisionens granskning efterfrågade stöd när det gäller etik och juridik.¹⁴

I SCB:s och DIGG:s rapporter saknas övergripande uppgifter om användningen av andra former av automatiserat beslutsfattande än AI, såsom villkors- och regelstyrda beslutsprocesser. Det saknas kartläggningar av förhållandet mellan helt automatiserat beslutsfattande och

14 Riksrevisionen (2020).

automatiserat beslutsstöd.¹⁵ Riksrevisionens granskningsrapport behandlar enbart helt automatiserade beslut.

Det finns få exempel hos likabehandlingsorgan på att myndigheter utsatt individer för risk för diskriminering genom automatiserat beslutsfattande. DO konstaterade i ett tillsynsbeslut från 2021 att Försvarsmakten överträdde diskrimineringsförbudet genom att automatiskt utesluta personer som uppgett att de har diagnosen adhd eller autism från möjligheten att prövas för sin lämplighet att göra militärtjänst. I detta fall sållades personerna bort direkt när de fyllt i ansökningsblanketten och där angivit att de hade en kognitiv funktionsnedsättning.¹⁶ Det finns ytterligare exempel från privat sektor, exempelvis bedömde den finska Diskriminerings- och jämställdhetsnämnden att ett kreditinstitut bröt mot diskrimineringslagen genom att använda ett poängsättningsystem för kreditvärdigheten som kunde missgynna enskilda personer utifrån kön, språk, bostadsort och ålder. Personen som anmälde detta fick en lägre kreditvärdighet med schablonen än vad hen skulle ha fått annars.¹⁷

15 Se SCB (2019); Myndigheten för digital förvaltning (2019); Riksrevisionen (2020).

16 Se DO:s tillsyner: Til 2018/511 och Til 2020/3.

17 Beslut i finska Diskriminerings- och jämställdhetsnämnden (2018).

Likabehandlingsorgan uppmärksammar risker för diskriminering

Som vi konstaterat har flera likabehandlingsorgan under senare tid uppmärksammat frågan om risker för diskriminering vid tillämpningen av AI och automatiserat beslutsfattande.

Equinet, ett europeiskt nätverk för likabehandlingsorgan, har i olika underlag pekat ut risker för diskriminering som kan uppstå vid tillämpningen av automatiserat beslutsfattande.¹⁸ De lyfter bland annat fram bristande insyn, svarta lådan-fenomenet¹⁹ och brist på kompetens om diskriminering och mänskliga rättigheter i utvecklingen av automatiserat beslutsfattande. Equinet bedömer att antidiskriminering och lika rättigheter och möjligheter bör bli en naturlig del i all utveckling av AI och automatiserat beslutsfattande. Equinet föreslår också att likabehandlingsorgan ska få de resurser och möjligheter som krävs för att ha en effektiv tillsyn inom detta område.

European Union Agency for Fundamental Rights (FRA) är europeiska unionens oberoende byrå för grundläggande rättigheter, och de konstaterar i en rapport att det saknas kunskap om hur grundläggande rättigheter påverkas av de nya teknologierna.²⁰ Rapporten bygger på ett

18 Equinet (2020a; 2020b; 2020c).

19 Svarta lådan-fenomenet handlar om datorprogram som självständigt gör kopplingar och ser mönster och samband mellan olika data, där det kan vara svårt även för programmerare att beskriva hur och varför dessa program drar vissa slutsatser. Se även Naarttijärvi (2017); Diskrimineringsombudsmannen (2020); Pasquale (2015).

20 FRA (2020).

stort antal intervjuer med informanter i både offentlig sektor och privata företag i olika delar av Europa, och en central iakttagelse är att majoriteten av de intervjuade i princip förstår att automatiserat beslutsfattande kan leda till diskriminering. Samtidigt är det få som anser att deras egna automatiserade beslutsprocesser skulle kunna leda till diskriminering. Av intervjuerna framkom även att det fanns en övertygelse om att det räcker med att plocka bort direkta kopplingar till kön, ålder och etnicitet för att systemen inte ska leda till diskriminering.

Även Institut for menneskerettigheder i Danmark tar upp riskerna för diskriminering vid tillämpningen av AI och rekommenderar offentliga myndigheter att granska algoritmiska profileringsmodeller för att säkerställa att dessa inte leder till diskriminering.²¹ De påtalar vikten av att myndigheter informerar om när profileringsmodeller används vid beslutsfattandet. Vidare konstaterar de att myndigheterna måste ha kunskap om och utbildning i hur profileringsmodeller kan utvecklas och tillämpas på ett sådant sätt att individers rättigheter kan säkerställas.

Gemensamt för dessa och andra rapporter som publicerats av lika-behandlingsorgan är att de lyfter behovet av att analysera hur tillämpning av AI och automatiserat beslutsfattande påverkar grundläggande rättigheter. Tidigare rapporter lyfter fram ett antal risker för diskriminering vid tillämpningen: risker vid bristande kvalitet i data, risker vid bristande träning av AI och risker vid bristande transparens.

21 Institut for menneskerettigheder (2021a; 2021b; 2021c).

Risker vid bristande kvalitet på använd data

Automatiserat beslutsfattande bygger på dataanvändning, och särskilt när det gäller AI behövs stora mängder information. I flera sammanhang uppmuntras offentliga organisationer i Europa att möjliggöra delning och spridning av data.²² Behovet av data har lett till en mångmiljardindustri för att utvinna, sälja och distribuera data, ibland av tveksam kvalitet.²³

Brister i data kan utgöra en risk för diskriminering genom att:

1. nödvändiga data saknas,
2. data som används inte återspeglar faktiska förhållanden utan har påverkats av snedvridning redan vid insamling,
3. data kan återspegla historisk diskriminering och därmed reproducera den. Data kan också tolkas på ett sätt som leder till risk för diskriminering.²⁴

Ett exempel på effekten av snedvridna data är ett stort amerikanskt företags försök att utveckla ett rekryteringsprogram för tekniker som skulle leda till att fler kvinnor anställdes. Som underlag för det självlärande AI-programmet använde man tidigare anställdas meritförteckningar. De tidigare anställda var dock män, vilket resulterade i att AI-programmet skapade ett mönster där män premierades och kvinnor valdes bort. Även när programmerarna tog bort direkta kopplingar till kön identifierade programmet hela tiden nya indirekta kopplingar till kön som ledde till att kvinnor valdes bort.²⁵

22 Myndigheten för digital förvaltning (2019); *Frontiers in Big Data* (2020); Europeiska kommissionen (2020); Myndigheten för digital förvaltning (2019).

23 Rieke, Yu, Robinson, von Hoboken (2016); *New Economics Foundation*. (2018); FOI (2020).

24 Akhil (2019); Reuters (2018).

25 Akhil (2019); Diskrimineringsombudsmannen (2020; 2019a).

Risker med bristande kvalitet vid träning av AI

Flera former av AI utvecklas genom att dataprogrammet går igenom stora mängder data, ser mönster och föreslår tolkningar, rekommendationer och svar. Under lärprocessen fininställs programmet för att fungera tillfredsställande, genom att förslagen till tolkningar och svar bedöms av särskilda handläggare som ger datorprogrammet återkoppling. Vartefter ger datorprogrammet allt mer korrekta svar. Hur denna träning av AI genomförs är därmed avgörande för utfallet i de automatiserade besluten.

Ett exempel på hur data återspeglar historisk diskriminering vid användning av AI finns i den prediktiva polisverksamhet som använts i delar av USA. Polisen använde algoritmer för att bedöma vilka områden som de skulle patrullera i. Den data som AI tränades i byggde på långvarig diskriminering av svarta medborgare, där polisen oftare både utreder och arresterar svarta medborgare. När inmatade data innehåller snedvridning kommer inlärningsalgoritmen att träna en modell som diskriminerar. Om poliserna fortsätter att fokusera på vissa områden kommer en uppdaterad modell diskriminera samma områden och målgrupper, men i ännu högre grad än tidigare.²⁶

Risker med bristande transparens och svarta lådor

Brist på transparens innebär en risk för att diskriminering inte upptäcks.²⁷ Med brist på insyn eller dokumentation av beslutsprocessen minskar människors möjligheter att hävda sina rättigheter genom anmälningar eller överklaganden. Problemet med utebliven transparens gäller inte minst för självlärande AI som kan komma fram till sina svar på sätt som inte ens

26 Institut for menneskerettigheter (2021a).

27 Se exempelvis: Larsson (2019); Diskrimineringsombudsmannen (2019c).

programmerarna förstår. Ytterligare ett problem som lyfts är att sekretess för affärshemligheter kan bidra till bristande transparens.²⁸

EU-kommissionen, Equinet och FRA föreslår obligatorisk granskning av den data som AI använder, vilket kräver att det finns dokumentation som görs tillgänglig för enskilda personer, intresseorganisationer och tillsynsmyndigheter.²⁹

28 Larsson (2018; 2019); Equinet (2020a).

29 Europeiska kommissionen (2021); Se även Institut for menneskerettigheder (2021a; 2021b, 2021c).

Resultat

Här redovisar vi det som framkommit i enkät- och intervjustudien. I de första avsnitten redogör vi för myndigheternas sätt att använda automatiserat beslutsfattande. Därefter kommer vi in på frågan om vilken kunskap myndigheterna har om risker för diskriminering och hinder för lika rättigheter och möjligheter vid tillämpningen av automatiserat beslutsfattande.

Sammantaget bad vi 34 myndigheter att svara på enkäten, och 6 av dessa deltog även i fördjupande intervjuer.³⁰ Kartläggningen omfattar både helt automatiserat beslutsfattande och beslut som tas av handläggare med hjälp av automatiserat beslutsunderlag.

Statliga myndigheters användning av automatiserat beslutsfattande

Få förvaltningsmyndigheter använder automatiserat beslutsfattande, vilket även andra studier har visat.³¹ Riksrevisionens granskning visar dock att det är en stor mängd automatiserade beslut, som berör många individer, som fattas på detta fåtal myndigheter.³²

Det automatiserade beslutsfattandet på myndigheter har i första hand införts för att öka effektiviteten, minska administrativa bördor och förkorta handläggningstider. De intervjuade myndigheterna anser att automatiserat beslutsfattande är mer effektivt, leder till kortare svarstider

30 Av de 34 tillfrågande myndigheterna svarade 33 på vår enkät. I bilaga 2 framgår vilka myndigheter som omfattas av kartläggningen.

31 SCB (2019); Myndigheten för digital förvaltning (2019); Riksrevisionen (2020).

32 13 myndigheter hade 2019 totalt 112 miljoner helt automatiserat beslutsprocesser och fattade 137 miljoner automatiserade beslut det året. Riksrevisionen (2020).

och ger en mindre andel fel i besluten jämfört med manuell hantering. Ärendehanteringens upplevs från myndigheternas perspektiv som mer likvärdig, enhetlig och rättssäker. Flera uppger också att detta visar sig i färre överklaganden. Ett par myndigheter har gjort uppföljningar som tyder på att besluten har blivit mer rättssäkra, och att risken för att diskrimineras utifrån kön och ålder har minskat. Av intervjuerna framkommer att myndigheterna först på senare tid har diskuterat frågor om etik, diskriminering och AI och automatiserat beslutsfattande.

Knappt hälften av de tillfrågade myndigheterna använder automatiserat beslutsfattande

Av enkäten framgår att 14 av de 33 myndigheterna använder automatiserat beslutsfattande.³³ 18 svarade att de inte använder sig av automatiserat beslutsfattande och 1 myndighet att man inte vet, se tabell 1. Samtidigt kan definitioner av automatiserat beslutsfattande uppfattas på olika sätt. Exempelvis svarade 1 myndighet att man inte har automatiserat beslutsfattande, samtidigt som samma myndighet senare i enkäten redogjorde för ett tekniskt verktyg som är att likna vid profilering.

Tabell 1. Använder sig myndigheten av automatiserat beslutsfattande gentemot enskilda?

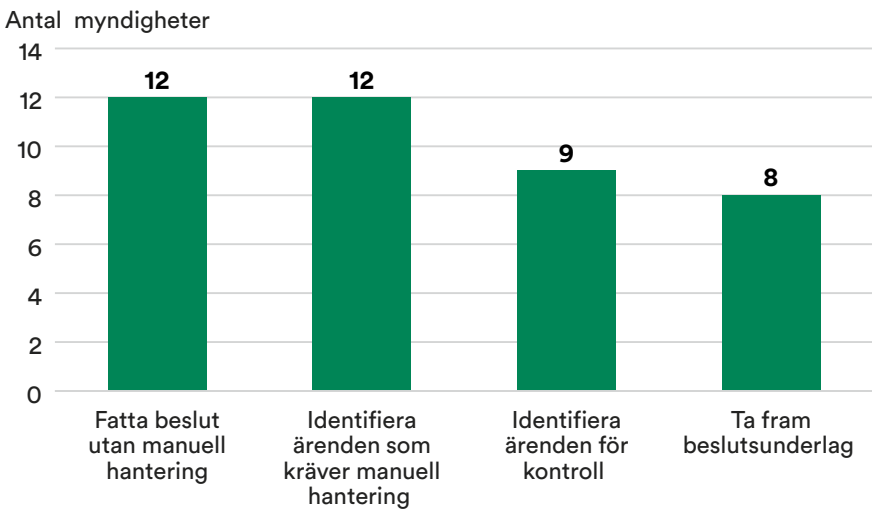
Använder automatiserat beslutsfattande	Använder inte automatiserat beslutsfattande	Vet inte om automatiserat beslutsfattande används	Bortfall	Summa
14	18	1	1	34

33 Se bilaga 1 för information om vilka myndigheter som tillfrågats.

Automatiserat beslutsfattande används i olika processer

En fråga gällde i vilka processer automatiserat beslutsfattande används gentemot enskilda, och enligt svaren är det lika vanligt att myndigheter använder automatiserade processer till att fatta beslut utan manuell hantering som att det används för att identifiera ärenden som kräver manuell hantering (12 myndigheter vardera), se figur 2. Ytterligare 9 myndigheter använder det för att identifiera ärenden för kontroll och åtta myndigheter använder automatiserat beslutsfattande för att ta fram underlag som beslutstöd.

Figur 2: Till vad används de automatiserade processerna?*



* Flera svarsalternativ var möjliga.

Få myndigheter använder AI

På frågan om AI³⁴ används i någon del av det automatiserade beslutsfattandet gentemot enskilda svarade fyra myndigheter att de använder AI. Detta resultat ligger i linje med Riksrevisionens granskning som visar att få myndigheter använder AI.³⁵

De fyra myndigheterna använder AI för att ta fram beslutsunderlag, identifiera ärenden som kräver manuell hantering och identifiera ärenden för kontroll. Vid intervjuerna med samma myndigheter framkom att de har tillämpat AI i den utsträckning de kan utifrån dagens möjligheter.

Endast en myndighet uppger att de använder profilering

Endast en myndighet anger att profilering används i någon del av det automatiserade beslutsfattandet. I enkäten definierade vi profilering som en automatiserad personuppgiftsbehandling som avser att framkalla information om en enskild persons egenskaper, beteende och personliga preferenser. Samtidigt kan vi se att ytterligare en myndighet i enkätsvaren beskriver att de har ett system, som kan tolkas som profilering.

Drygt hälften av myndigheterna informerar om automatiskt beslutsfattande

En fråga gällde om den enskilde blir informerad om att hans ärende behandlas helt eller delvis med hjälp av automatiskt beslutsfattande, och svarar 9 av 14 myndigheter att den enskilde informeras om detta. Dessa myndigheter gör det på olika sätt och vissa myndigheter på mer än ett sätt. Några har informationen på webbplatsen, i e-tjänsten i anslutning

34 Ett system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder – med viss grad av självständighet – för att uppnå särskilda mål.

35 Riksrevisionen (2020).

till ansökan, på beslutsdokument och i informationstext om den berörda processen.

Kunskap om risker för diskriminering och hinder för lika rättigheter och möjligheter

Det är en svår uppgift att få svar på frågan om vilken kunskap om risker för diskriminering och hinder för lika rättigheter och möjligheter som myndigheterna har avseende automatiserat beslutsfattande. I denna kartläggning har vi försökt att ta reda på hur de arbetar med automatiserat beslutsfattande och vilka strukturer, såsom styrdokument, riskanalyser och kvalitetsuppföljningar, som de har för att säkerställa att individer inte missgynnas. I likhet med Riksrevisionen menar vi att myndigheter behöver strukturer för att säkerställa att automatiserade beslutsprocesserna är både effektiva och rättssäkra.³⁶ I enkäten frågade vi om myndigheterna har styrdokument som ska förhindra att enskilda missgynnas, om de genomför riskanalyser och kvalitetsuppföljningar, och i vilken utsträckning som diskrimineringsgrunderna beaktas i dessa. I intervjuerna bad vi de sex myndigheterna att resonera kring riskerna för diskriminering vid tillämpning av automatiserat beslutsfattande och sätt att förebygga dem.

Våra frågor om risker för diskriminering vid tillämpningen av automatiserat beslutsfattande har i sig gjort myndigheterna mer medvetna om detta. Några av de intervjuade myndigheterna påtalade att de, inför och under intervjuerna, kommit fram till att de inte vet tillräckligt om riskerna för diskriminering vid tillämpning av automatiserat beslutsfattande, och att de har förstått hur de aktivt kan arbeta med dessa frågor. Att känna till att det finns risker för diskriminering är därmed inte samma sak som att veta hur riskerna kan förebyggas i de egna processerna.

36 Riksrevisionen (2020).

Styrdokument, riskanalyser och kvalitetsuppföljning

Myndigheterna har i olika utsträckning strukturer på plats för att säkerställa att individer inte missgynnas. Ingen av myndigheterna har beaktat samtliga diskrimineringsgrunder i sina styrdokument, och även om de flesta gör riskanalyser och kvalitetsuppföljningar är det ovanligt att de beaktar alla diskrimineringsgrunderna. Intervjuszvaren visar att det finns en viss medvetenhet om risker för diskriminering och hinder för lika rättigheter och möjligheter vid tillämpningen av automatiserat beslutsfattande, men att de i nuläget inte har ett systematiskt arbete för att hantera dessa risker.

Få har styrdokument för att förhindra att enskilda missgynnas

Av de fjorton myndigheter som använder automatiserat beslutsfattande har fyra styrdokument som ska säkerställa att enskilda inte missgynnas vid automatiserat beslutsfattande, och samtliga anger att diskrimineringsgrunderna beaktas i dokumenten. Tre av dessa fyra myndigheter har vi intervjuat, och de uppger att deras styrdokument i huvudsak rör diskrimineringsgrunderna kön, etnisk tillhörighet och ålder.

De flesta gör riskanalyser för att förhindra att enskilda missgynnas, men få beaktar alla diskrimineringsgrunderna

Totalt tio myndigheter anger att de gör riskanalyser vid implementering av automatiserat beslutsfattande för att säkerställa att enskilda inte missgynnas. Av dessa tio rapporterar fyra myndigheter att diskrimineringsgrunderna beaktas i riskanalyserna.

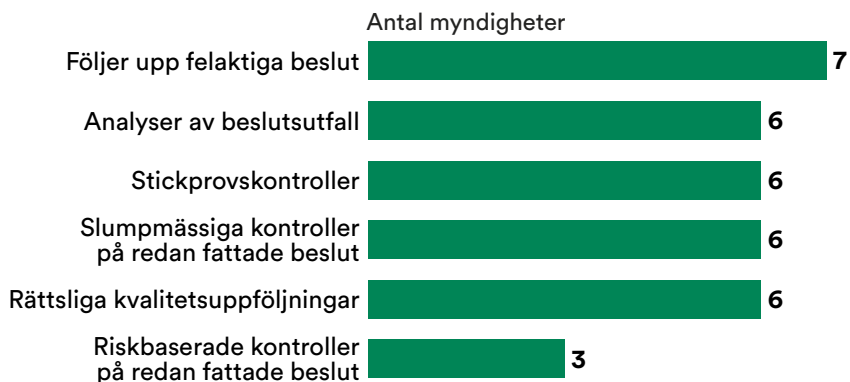
De flesta gör kvalitetsuppföljningar för att enskilda inte ska missgynnas

Av fjorton myndigheter svarar elva att de gör kvalitetsuppföljningar för att säkerställa att enskilda inte missgynnas vid automatiserat beslutsfattande. Av dessa svarar fem myndigheter att de även beaktar diskrimineringsgrunderna i kvalitetsuppföljningarna. Endast en myndighet har identifierat

fall av missgynnande i någon av de processer som omfattas av automatiserat beslutsfattande. Två av dessa fem är intervjuade, och det framgick att de inte beaktar alla diskrimineringsgrunderna.

Figur 3 visar de vilka olika typer av kvalitetsuppföljningar som myndigheterna gör för att säkerställa att enskilda inte missgynnas av beslut som fattas automatiserat. De följer upp felaktiga beslut (sju myndigheter), analyser av beslutsutfall (sex myndigheter), stickprovskontroller (sex myndigheter), slumpmässiga kontroller på redan fattade beslut (sex myndigheter), rättsliga kvalitetsuppföljningar (sex myndigheter) och riskbaserade kontroller på redan fattade beslut (tre myndigheter).

Figur 3: Vilka typer av kvalitetsuppföljningar gör myndigheten för att säkerställa att enskilda inte missgynnas av beslut som fattas automatiserat?*



*Flera svarsalternativ var möjliga.

Alla intervjuade myndigheter gör risk- och konsekvensanalyser samt har övergripande policyer och strategier för automatiserat beslutsfattande. Respondenterna redogör för att myndigheterna i någon mån beaktar diskrimineringsgrunderna kön, ålder och etnisk tillhörighet i detta arbete, men utifrån enskilda tjänstepersoners kunskaper och inte som ett resultat av ett systematiskt arbete.

Vidare beskriver myndigheterna i intervjuerna att diskrimineringslagen inte har vägts in när man utvecklat, implementerat eller testat programvaran för de automatiserade beslutsprocesserna, och därmed inte heller i de färdiga besluten. I stället har myndigheterna beaktat andra lagar och förordningar såsom dataskyddsförordningen GDPR och förvaltningslagen.

Fåtal ser behov av att utveckla arbetet med att minska riskerna för missgynnande

Av enkätsvaren framgår att fem myndigheter ser ett behov av att utveckla sitt arbete mer för att minska riskerna för missgynnande i automatiserat beslutsfattande. En myndighet har svarat att behovet är stort och fyra myndigheter att de i viss mån har behov av detta. En myndighet vet inte och resterande åtta myndigheter har svarat att de inte har något sådant behov.

De intervjuade myndigheterna är tydliga med att de ser ett behov av att öka sina kunskaper och utveckla arbetet mot risker för diskriminering. Flera lyfter behovet av dialog och samarbete med experter på området. De tar även upp att de behöver mer kunskap om vilka risker som finns för diskriminering vid tillämpningen av automatiserat beslutsfattande, och mer kunskap om metoder och verktyg för att motverka dessa risker.

Myndigheterna tar även upp vikten av att ha kompetens inom området för det automatiserade beslutsfattandet för att besluten ska fattas på ett ändamålsenligt sätt. Vidare menar de att medarbetarna behöver behålla sin beslutsfattarkompetens och inte helt förlita sig på det automatiserade beslutsstödet, eftersom de då kan få sämre förmåga att kritiskt granska besluten. Detta skulle i sin tur kunna leda till att felaktiga beslut fattas och att diskriminering inte upptäcks. Det finns en risk för att handläggare ser automatiserade beslut som objektiva och mer korrekta än manuell hanterade beslut. Myndigheterna är medvetna om att det finns risker med att lita för mycket på automatiserade beslut och att det alltid är myndigheten som ansvarar för ett beslut.

Myndigheterna anser att bristen på kompetens inom AI och automatiserat beslutsfattande kan göra det svårare att få insyn i beslutsprocesser och därmed förebygga att enskilda personer missgynnas. Utan rätt kompetens kan de inte förklara bakgrunden till olika automatiserade beslut för sökande och klienter. Flera av myndigheterna betonar att de har svårt att få tillgång till den IT-kompetens som de behöver.

Utmaningar med snedvriden data från tredje part

De intervjuade myndigheterna är medvetna om att den data som de använder kan vara snedvriden och att detta i så fall påverkar hela beslutsprocessen. De använder data från andra myndigheter i sina automatiserade beslutsprocesser, exempelvis behöver Transportstyrelsen information från Polismyndigheten, Arbetsförmedlingen från Skatteverket och Försäkringskassan från hälso- och sjukvården. Data som delas mellan myndigheterna granskas inte, utan den eventuella snedvridningen som finns följer med in i nästa beslutsprocess. Den mottagande myndigheten har antagit att avsändaren har ansvaret för att data är korrekt.

Avslutande iakttagelser

Kartläggningen visar att automatiserat beslutsfattande används vid 14 av de 34 tillfrågade myndigheterna och att de automatiserade besluten berör ett stort antal enskilda individer. Bedömningen är också att de myndigheterna som intervjuats har en viss inblick i risker för diskriminering som kan förekomma, men de talar främst om etik och integritet snarare än risk för diskriminering i relation till automatiserat beslutsfattande. Myndigheterna redogör för hur de aktivt förhåller sig till exempelvis GDPR, men att de inte i någon större utsträckning eller systematiskt beaktar diskrimineringsperspektivet i arbetet med att implementera automatiserat beslutsfattande.

I kartläggningen framkommer också att myndigheter som använder automatiserat beslutsfattande gör riskanalyser och kvalitetsuppföljningar, men att dessa sällan beaktar diskrimineringsgrunderna. Få myndigheter har tydliga styrdokument som ska säkerställa att enskilda individer inte missgynnas vid automatiserat beslutsfattande. Vi kan konstatera att myndigheterna i stort sett saknar ett diskrimineringsperspektiv i de automatiserade beslutsprocesserna. Dessutom är det få myndigheter som ser att de behöver göra mer för att minska risken för missgynnande. De myndigheter som vi har intervjuat understryker dock att de gärna ser ett kunskapsutbyte med DO i dessa frågor.

DO kan konstatera att myndigheterna känner till de tre mest framträdande riskerna för diskriminering vid tillämpning av AI och automatiserat beslutsfattande som andra likabehandlingsorgan har pekat ut:

1. risker vid bristande kvalitet i data
2. risker vid bristande träning av AI
3. risker vid bristande transparens.

Däremot visar inte kartläggningen hur myndigheterna arbetar med att förebygga dessa risker. Utifrån enkätsvar och intervjuer kan DO dock konstatera att myndigheterna behöver mer kunskap om risker för diskriminering vid tillämpning av automatiserat beslutsfattande.

Med utgångspunkt i kartläggningen är det några iakttagelser som vi särskilt vill lyfta:

- Myndigheterna har sällan beaktat diskrimineringsgrunderna i styrdokument, riskanalyser och kvalitetsuppföljningar.
- Ett fåtal av dessa fjorton myndigheter har styrdokument som ska säkerställa att enskilda inte missgynnas. Endast fyra anger att de beaktar diskrimineringsgrunderna i detta arbete.
- Tio av fjorton myndigheter gör riskanalyser för att säkerställa att inte enskilda missgynnas. Endast fyra beaktar diskrimineringsgrunderna i detta arbete.
- De flesta myndigheter gör kvalitetsuppföljningar för att säkerställa att enskilda inte missgynnas. Endast fyra anger att de beaktar diskrimineringsgrunderna i detta arbete.
- Endast nio av fjorton myndigheter informerar individen om att beslut fattas genom automatisering.
- Myndigheterna pekar på risken för att handläggare fäster för stor vikt vid de automatiserade besluten, och mister förmågan att kritiskt granska dessa.
- Få myndigheter svarar att de behöver mer kunskap om och stöd i hur de kan minska risken för missgynnande. De sex intervjuade myndigheterna efterfrågar dock mer kunskap om risker för diskriminering och ökad kompetens i dataprogrammering.

Sammantaget visar kartläggningen att kunskapen om risker för diskriminering och hinder för lika rättigheter och möjligheter inte är tillfredställande. Det innebär att myndigheterna behöver öka sin medvetenhet om förbudet mot diskriminering. Kartläggningen omfattar bara statliga förvaltningsmyndigheter, men automatiserat beslutsfattande används även i andra delar av offentlig sektor och inom näringslivet. Medvetenheten om de risker för diskriminering som kan uppstå vid tillämpningen av AI och automatiserat beslutsfattande kan därmed behöva fördjupas även inom andra delar av samhället.



Referenser

AI Watch (2020). Overview of the use and impact of AI in public services in the EU. Publications office of the European Union. Luxemburg.

Akhil Alfons Kodyan (2019). An overview of ethical issues in using AI systems in hiring with a case study of Amazon's AI based hiring tool.

Andersson Schwarz, Jonas. och Larsson, Stefan, red (2019). Plattformsamhället. Den digitala utvecklingens politik, innovation och reglering. Stockholm: Fores.

Diskriminerings- och jämställdhetsnämnden (2018). Diskriminerings- och jämställdhetsnämnden förbjöd ett kreditinstitut att använda statistiska metoder som diskriminerar creditsökande vid konsumentkrediter. Pressmeddelande 3/2018, 25.4.2018. Helsingfors. https://www.yvtltk.fi/material/attachments/ytalk/tiedotteet/GDnSSoUmd/yvtLTK-tiedote_-3-2018-sv.pdf.

Diskrimineringsombudsmannen (2018). Utveckling av kunskap om automatiserad databehandling med algoritmer och stordata och diskriminering eller risker för diskriminering. LED 2018/387. Stockholm.

Diskrimineringsombudsmannen (2019a). Kunskapsöversikt om användningen och utvecklingen av automatiserad databehandling med algoritmer (artificiell intelligens) och stordata och diskriminering eller risker för diskriminering. Stockholm.

Diskrimineringsombudsmannen (2019b). Promemoria om automatiserad databehandling med algoritmer och risker för diskriminering: inom rekrytering och kreditgivning. Stockholm.

Diskrimineringsombudsmannen (2019c). Tillgång till information som förutsättning för att bedöma förekomsten av diskriminering i samband med automatiserat beslutsfattande. Stockholm.

Diskrimineringsombudsmannen (2020). Uppdragsdirektiv för dialog och kunskapsspridning om AI, automatiserat beslutsfattande och risker för diskriminering. LED 2020/484. Stockholm.

Equinet (2020a). Regulating for an Equal AI: A New Role for Equality Bodies. Meeting the new challenges from increased digitisation and the use of Artificial Intelligence. Bryssel.

Equinet (2020b). The Summary and Framework for Action for Equality Bodies. Bryssel.

Equinet (2020c). A Good Practice Guide. Bryssel.

Equinet (2020d). Regulating for an equal AI: A new role for Equality bodies. Meeting the new challenges to equality and non-discrimination from increased digitation and the use of Artificial Intelligence. AI Taskforce of the Equinet Working Group on Communication. Bryssel.

EU-kommissionen (2021). Proposal for a Regulation of the European parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}, Bryssel.

FADA, Federal Anti-Discrimination Agency (2020). Risks of Discrimination through the Use of Algorithms. Berlin.

FOI, Totalförsvarets forskningsinstitut (2020). Vilse i lasagnen? En upptäcktsfärd i den svenska digitaliseringens mångbottnade problemstruktur. FOI-R--4814—SE.

FRA, European Union Agency for Fundamental Rights, (2020). Getting the future right – Artificial Intelligence and fundamental rights. Bryssel.

Frontiers in Big Data (2020). Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective.

Institut for menneskerettigheder (2021a). Når algoritmer sagsbehandler – Rettigheder og retssikkerhed i offentlige myndigheders brug af profileringsmodeller. Köpenhamn.

Institut for menneskerettigheder (2021b). Når algoritmer sagsbehandler: Gennemsigtighed i det offentliges brug af algoritmiske profileringsmodeller. Notat. Köpenhamn.

Institut for menneskerettigheder (2021c). Når algoritmer forskelsbehandler: Risiko for diskrimination i det offentliges brug af algoritmiske profileringsmodeller. Notat. Köpenhamn.

Kilden kjønnsforskning.no (2018). HVA VET VI OM...? Kunstig intelligens og likestilling. En kartlegging av norsk forskning. Oslo.

Larsson, Stefan (2019). Artificiell intelligens som normativ samhällsutmaning: partiskhet, ansvar och transparens, i Banakar, Dahlstrand & Ryberg-Welander (red.) Festskrift till Håkan Hydén. Lund: Juristförlaget.

Larsson, Stefan (2018). Sju nyanser av transparens: om artificiell intelligens och ansvaret för digitala plattformars samhällspåverkan, i Andersson Schwarz, J. och Larsson, S. (red.) Plattformssamhället. Den digitala utvecklingens politik, innovation och reglering. Stockholm: Fores.

Myndigheten för digital förvaltning (2020). Främja den offentliga förvaltningens förmåga att använda AI. Delrapport i regeringsuppdraget I2019/01416/DF || I2019/01020/DF (delvis). Stockholm.

Myndigheten för digital förvaltning (2019). Uppdrag att öka den offentliga förvaltningens förmåga att tillgängliggöra öppna data, bedriva öppen och datadriven innovation samt använda artificiell intelligens. Slutrapport i regeringsuppdraget I2019/0416/DF samt I2019/01020/DF (delvis). Dnr: 2019–139.

Naarttijärvi, Markus (2017). Rättstatligheten och algoritmiska svarta lådor, ur Jubileumsskrift till Juridiska institutionen 40 år. (red): Edström, Ö, Lindholm, J, Mannelqvist, R. Umeå: Umeå universitet, Juridiska institutionen.

New Economics Foundation (2018). The Rise of the Data Oligarchs. Power and Accountability in the Digital Economy. Part 1: Data Collection. London: NEF.

Nordiskt välfärdscenter (2021b). Ny teknik och digitala lösningar för ökad inkludering: en kunskapsöversikt. <https://nordicwelfare.org/wp-content/uploads/2021/03/ny-teknik-och-digitala-losningar-for-okad-inkludering.pdf>.

Pasquale, Frank (2015). The Black Box Society. The Secrets Algorithms that Controls Money and Information. USA: Harvard University Press.

Reuters (2018). Amazon scraps secret AI recruiting tool that showed bias against women.

Riksrevisionen (2020). Automatiserat beslutsfattande i statsförvaltningen – effektivt men kontroll och uppföljning brister. RiR 2020:22. Stockholm.

Rieke, Aaron, Yu, Harlan, Robinson, David, Joris von Hoboken, J (2016). Data Brokers in an Open Society. An Upturn Report, prepared for the Open Society.

SCB, Statistikmyndigheten (2019). Artificiell intelligens i Sverige. Solna.

Zuiderveen Borgesius, F (2018). Discrimination, artificial intelligence and algorithmic decision-making. Council of Europe, Directorate General of Democracy. <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>

Bilaga 1. Metod och urval

Rapporten bygger på en sammanställning av tidigare studier på området, en enkät till ett urval av statliga myndigheter och intervjuer med tjänstepersoner på ett urval av dessa myndigheter.

Tidigare studier

I arbetet med kartläggningen har vi sammanställt resultaten från studier och rapporter som handlar om AI och automatiserat beslutsfattande kopplat till frågeställningar om diskriminering, mänskliga rättigheter och etik och juridik. Vi har tagit del av vetenskapligt granskade artiklar i olika vetenskapliga tidskrifter, men även studier som tagits fram av myndigheter och organisationer – så kallad grå litteratur. Vi har dels sökt med ett brett fokus på automatiserat beslutsfattande, AI och diskriminering, dels med ett smalare fokus på myndigheter och risker för diskriminering. Studierna identifierades genom sökningar på Google och Google Scholar samt i databaserna Libris, Researchgate, DIVA, DOAJ, SwePub, Human Resources Abstracts, PsycInfo och SocIndex. Sökningar har gjorts på svenska och engelska. Sökord har varit artificiell intelligens, automatiserat beslutsfattande och varianter på detta begrepp samt digitalisering och begrepp såsom digitisering, robotisering, automatisering och deras engelska motsvarigheter. Dessa begrepp har i olika söksträngar kopplats till begrepp såsom diskriminering, etik, mänskliga rättigheter, snedvridning, rättvisa och hållbarhet. I den smala sökningen kopplades sökorden till nya begrepp såsom myndigheter och offentlig administration (på engelska agencies, government, authorities, public authorities och administration). Sökningarna begränsades från 1995 och framåt.

Sammantaget har vi identifierat cirka 50 relevanta artiklar och studier.¹ Utöver sökningar i de nämnda databaserna har vi tagit del av rapporter från aktörer som verkar inom diskrimineringsområdet såsom andra likabehandlingsorgan, Equinet och FRA.

Urval av myndigheter för enkät

Vid urvalet av myndigheter för enkäten använde vi SCB:s allmänna myndighetsregister.² För att inkluderas skulle en förvaltningsmyndighet ha i uppdrag att fatta beslut gentemot enskilda, och i beslutsprocessen kunna antas använda automatiskt beslutsfattande som skulle kunna leda till diskriminering. Urvalet baserades på tidigare studier och en genomläsning av de statliga förvaltningsmyndigheternas hemsidor, årsredovisningar och regleringsbrev.

Granskningarna resulterade i ett urval av 34 statliga myndigheter, från de 249 statliga förvaltningsmyndigheter som ingår i SCB-registret. De utvalda myndigheterna varierar i storlek och typ av verksamhet. Bland dem fanns stora organisationer såsom Arbetsförmedlingen, Skatteverket och Försäkringskassan, medelstora såsom Centrala studiestödsnämnden och Lantmäteriet samt relativt små myndigheter såsom Statens tjänstepensions- och grupplivnämnd.

Enkätens utformning och genomförande

Syftet med enkätundersökningen var att få en bild av i vilken grad statliga myndigheter använder automatiserat beslutsfattande och en bild av myndigheternas kunskaper om risker för diskriminering eller hinder för lika rättigheter och möjligheter vid automatiserat beslutsfattande.

-
- 1 Vi har fått rekommendationer på studier från personer som vi talat med vid myndigheter, akademien och intresseorganisationer.
 - 2 SCB (2020). Universitet, högskolor och länsstyrelser är inte inkluderade i undersökningen.

Enkäten omfattade tjugo frågor som kvalitetssäkrades av projektets referensgrupp och av Stefan Larsson, docent i teknik och social förändring vid Lunds universitet, som forskar om rättsliga och etiska aspekter av AI.³

En förfrågan om att delta i enkäten skickades till registratorer vid 34 myndigheter med information om syfte och förutsättningar. Information om enkäten spreds också i det AI-nätverk för myndigheter som DO är del av.⁴

Myndigheterna fick tre veckors svarstid. Totalt besvarade 33 myndigheter enkäten, vilket innebär en svarsfrekvens på 97 procent.

Urval av myndigheter för intervjuer

Syftet med intervjuerna var att få en fördjupad bild av hur myndigheterna använder automatiserat beslutsfattande och vilken kunskap dessa myndigheter har om förhållanden som kan leda till diskriminering eller utgöra hinder för lika rättigheter och möjligheter kopplat till automatiserat beslutsfattande.

Vi genomförde intervjuer med Arbetsförmedlingen, Försäkringskassan, Transportstyrelsen, Skatteverket och Kriminalvården. Det framgår av tidigare studier att Arbetsförmedlingen, Försäkringskassan, Transportstyrelsen och Skatteverket har utvecklat och använder automatiserat beslutsfattande gentemot enskilda i stor omfattning. Kriminalvården intervjuades på grund av sitt svar på enkäten, som indikerade att de har intressanta erfarenheter av automatiserat beslutsfattande.

3 Enkäten konstruerades i enkätverktyget LimeSurvey för att kunna besvaras elektroniskt.

4 AI-nätverket för myndigheter: <https://tillvaxtverket.se/om-tillvaxtverket/samverkan/ai-natverket.html>.

Därutöver intervjuade vi även myndigheten DIGG eftersom de stödjer användandet av AI i offentliga myndigheter.⁵

De intervjuade var verksamhetschefer och jurister men få IT-specialister

Intervjupersonerna hade olika yrkesroller men de flesta var verksamhetschefer, verksamhetsutvecklare och jurister. Majoriteten av juristerna hade inriktning mot personsuppgiftsskydd, jämställdhet, icke-diskriminering och liknande. Ett fåtal hade titlar och uppgifter som kunde associeras med IT eller data. De olika verksamhetscheferna hade olika ansvar. En del var högre chefer som ansvarade för ett affärsområde eller en tjänst där automatiserat beslutsfattande användes.

Intervjuernas utformning och genomförande

Vi utformade två intervjuguider, en för de utvalda myndigheterna och en för DIGG och Vinnova. Inför intervjuerna fick deltagarna information om syftet och de frågeområden som skulle tas upp. Intervjuerna varade i en till en och halv timme och genomfördes digitalt. Oftast deltog flera personer från varje myndighet, och totalt deltog 18 personer i intervjustudien (se bilaga 4). En fördel med flera deltagare är att vi fick ta del av fler perspektiv och att intervjupersonerna uppmuntrade varandra till förtydliganden och kompletterade varandras svar.

5 Myndigheten för digital förvaltning (2020).

Bilaga 2: Enkätfrågor

1. Vilken myndighet arbetar du på? (Denna fråga är obligatorisk.)

- Arbetsförmedlingen
- Arbetsmiljöverket
- Bolagsverket
- Boverket
- Centrala studiestödsnämnden
- E-hälsomyndigheten
- Ekobrottsmyndigheten
- Försvarsmakten
- Försäkringskassan
- Integritetsskyddsmyndigheten
- Konsumentverket
- Kriminalvården
- Kronofogdemyndigheten
- Lantmäteriet
- Migrationsverket
- Naturvårdsverket
- Pensionsmyndigheten
- Polismyndigheten
- Skatteverket
- Socialstyrelsen
- Statens institutionsstyrelse
- Statens jordbruksverk
- Statens servicecenter

- Statens tjänstepensions- och grupplivsverk
- Statens tjänstepensionsverk
- Svenska ESF-rådet
- Säkerhetspolisen
- Tillväxtverket
- Totalförsvarets plikt- och prövningsnämnd
- Transportstyrelsen
- Tullverket
- Universitets- och högskolerådet
- Universitetskanslersämbetet
- Åklagarmyndigheten

2. Använder sig myndigheten av automatiserat beslutsfattande gentemot enskilda? (Denna fråga är obligatorisk.)

Hjälp text: Automatiskt beslutsfattande omfattar algoritmiska och automatiserade beslutsprocesser såväl med som utan AI. I begreppet inkluderar vi både helt automatiserat beslutsfattande och automatiserade processer som används som beslutstöd för handläggare.

- Ja
- Nej
- Vet ej

Frågor till myndigheter som använder automatiserat beslutsfattande

3. Om Ja på fråga 2: Ange i vilka processer automatiserat beslutsfattande används gentemot enskilda? (Denna fråga är obligatorisk.)

Hjälptext: Exempelvis processer som avser att ta fram beslut om bidrag/försäkringar (exempelvis tillfällig föräldrapenning), beslut om inkomstskatt, stöd, annan ekonomisk ersättning och arbetsmarknadsstödjande åtgärder.

4. Till vad används de automatiserade processerna? Fler svarsalternativ är möjliga. (Denna fråga är obligatorisk.)

- Fatta beslut utan manuell hantering
- Identifiera ärenden för kontroll
- Identifiera ärenden som kräver manuell hantering
- Ta fram beslutsunderlag
- Om annat, ange vad:

5. Används AI i någon del av det automatiserade beslutsfattandet gentemot enskilda? Om du svarar ja på frågan, ange till vad det används i rutan till höger. (Denna fråga är obligatorisk.)

Hjälptext: Artificiell intelligens avser system som uppvisar intelligent beteende genom att analysera sin miljö och vidta åtgärder – med viss grad av självständighet – för att uppnå särskilda mål.

Det kan exempelvis röra sig om att fatta beslut utan manuell hantering, identifiera ärenden för kontroll, identifiera ärenden som kräver manuell hantering eller ta fram beslutsunderlag.

- Ja
- Nej
- Vet ej

6. Används profilering i någon del av myndighetens automatiserade beslutsfattande? (Denna fråga är obligatorisk.)

Hjälptext: Profilering är en automatiserad personuppgiftsbehandling som avser att framkalla information om en enskild persons egenskaper, beteende och personliga preferenser. Automatiserat beslutsfattande kan ske med eller utan profilering.

- Ja
- Nej
- Vet ej

Riskanalys och kvalitetsuppföljning

7. Har myndigheten styrdokument som ska säkerställa att enskilda inte missgynnas vid automatiserat beslutsfattande? (Denna fråga är obligatorisk.)

Hjälptext: Ett missgynnande innebär att ett beslut får en negativ följd för en individ, det vill säga en skada, en nackdel eller ett obehag. Det kan handla om att en person kommer i ett sämre läge eller går miste om en förbättring, en förmån eller service. Det spelar ingen roll om det finns en avsikt att missgynna eller inte; det är effekten eller resultatet som avgör om det är ett missgynnande.

- Ja
- Nej
- Vet ej

- 8.** Om Ja på fråga 7: Beaktas diskrimineringsgrunderna i styrdokumentet? (Denna fråga är obligatorisk.)

Hjälptext: I diskrimineringslagen anges följande diskrimineringsgrunder: kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning och ålder.

- Ja
- Nej
- Vet ej

- 9.** Gör myndigheten riskanalyser vid implementering av automatiserat beslutsfattande för att säkerställa att enskilda inte missgynnas? (Denna fråga är obligatorisk.)

Hjälptext: Riskanalys handlar i denna enkät om att identifiera och värdera risken för att enskilda missgynnas vid automatiserat beslutsfattande. Vi avser riskanalyser som genomförs innan systemet är i drift.

- Ja
- Nej
- Vet ej

- 10.** Om Ja på fråga 9: Vilka typer av riskanalyser görs av det automatiserade beslutsfattandet? (Denna fråga är obligatorisk.)

Hjälptext: Det kan exempelvis handla om en övergripande och samlad riskanalys av de processer som omfattas av automatiserat beslutsfattande. Det kan också handla om en riskanalys för respektive process som omfattas av automatiserat beslutsfattande.

- 11.** Beaktas risker för missgynnande utifrån diskrimineringsgrunderna i riskanalyserna av det automatiserade beslutsfattandet? (Denna fråga är obligatorisk.)

Hjälptext: I diskrimineringslagen anges följande diskrimineringsgrunder: kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning och ålder.

- Ja
- Nej
- Vet ej

- 12.** Gör myndigheten kvalitetsuppföljningar för att säkerställa att enskilda inte missgynnas vid automatiserat beslutsfattande? (Denna fråga är obligatorisk.)

Hjälptext: Kvalitetsuppföljning avser åtgärder som utförs för att säkerställa och kontrollera att det automatiserade beslutssystemet fattar korrekta beslut. Kvalitetsuppföljning avser åtgärder som utförs när ett automatiserat beslutssystem är i drift.

- Ja
- Nej
- Vet ej

- 13.** Om Ja på fråga 12: Vilka typer av kvalitetsuppföljningar gör myndigheten för att säkerställa att enskilda inte missgynnas av beslut som fattas automatiserat? Flera svarsalternativ är möjliga. (Denna fråga är obligatorisk.)

- Analyser av beslutsutfall
- Följer upp felaktiga beslut
- Stickprovskontroller

- Riskbaserade kontroller på redan fattade beslut
- Slumpmässiga kontroller på redan fattade beslut
- Rättsliga kvalitetsuppföljningar
- Om annat, ange vad:

- 14.** Beaktas risker för missgynnande utifrån diskrimineringsgrunderna i kvalitetsuppföljningar av automatiserade beslutsfattandet? (Denna fråga är obligatorisk.)

Hjälptext: I diskrimineringslagen anges följande diskrimineringsgrunder: kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning och ålder.

- Ja
- Nej
- Vet ej

- 15.** Har myndigheten identifierat fall av missgynnande i någon av de processer som omfattas av automatiserat beslutsfattande? Om svaret på frågan är "Ja", ange i hur många fall. (Denna fråga är obligatorisk.)

Hjälptext: Exempelvis beslut om bidrag, inkomstskatt, stöd, annan ekonomisk ersättning och arbetsmarknadsstödjande åtgärder.

- Ja
- Nej
- Vet ej

- 16.** Om Ja på fråga 15: Hur hanterar ni automatiserade beslut där ni har identifierat att någon kan ha missgynnats? (Denna fråga är obligatorisk.)

17. Ser ni behov av att utveckla ert arbete med automatiserat beslutsfattande när det gäller att minska riskerna för missgynnande? (Denna fråga är obligatorisk.)

- Ja, ett stort behov
- Ja, i viss mån
- Nej
- Vet ej

Information till enskilda

18. Blir den enskilde informerad om att hens ärende behandlas helt eller delvis med hjälp av automatiskt beslutsfattande? (Denna fråga är obligatorisk.)

- Ja
- Nej
- Vet ej

19. Om Ja på fråga 18: Hur informeras enskilda om att myndigheten har processer med automatiserat beslutsfattande? Flera svarsalternativ är möjliga. (Denna fråga är obligatorisk.)

Ange var informationen finns:

- Beslutsdokument
- E-tjänst i anslutning till ansökan/anmälan
- Informationsbroschyr
- Hemsida vid informationstext om den berörda processen
- Hemsida, men inte i direkt anslutning till den berörda processen
- Pappersblankett
- Annat. Ange vad:

Avslutande reflektioner

- Är det något mer som ni vill tillägga?

Bilaga 3: Intervjuguide

- Berätta kortfattat om din/er bakgrund och din/er roll på myndigheten.
- Berätta övergripande om hur ni som myndighet använder automatiserat beslutsfattande och hur tillämpningen har utvecklats över tid.
- Beskriv vad som har varit eller är era utmaningar när det gäller automatiserat beslutsfattande och frågor om missgynnande och etik. Vad är likheter och skillnader jämfört med manuellt beslutsfattande?
- Hur arbetar ni för att säkerställa att enskilda inte missgynnas i ert automatiska beslutsfattande? Vad har ni lärt er? Hur har ni lärt er?
- Är det någon/några processer ni upplever som särskilt problematiska när det gäller att säkerställa att ingen enskild missgynnas? Varför? Beskriv på vilket sätt.
- Ser ni ytterligare behov av att utveckla ert arbete med automatiserat beslutsfattande när det gäller att minska riskerna för missgynnande?
- Behövs något för att ni (och andra myndigheter) ska kunna utveckla ert arbete med att förebygga riskerna för missgynnande av enskilda vid tillämpningen av automatiserat beslutsfattande?
- Anser ni att ni har de resurser och det kunnande ni behöver inom myndigheten för att kunna undvika att missgynna enskilda? Vad är det för resurser och kunnande? Vet ni var ni kan få tillgång till sådana resurser och kunnande?
- Vilka råd skulle ni ge andra myndigheter om hur de kan förebygga riskerna för missgynnande av enskilda vid tillämpningen av automatiserat beslutsfattande?
- Är det något mer som ni vill ta upp som vi inte har frågat om? Är det någon person eller myndighet som ni rekommenderar oss att tala med?

Bilaga 4. Intervjuförteckning

Nr	Myndighet	Datum	Antal respondenter
1	Transportstyrelsen	8 oktober 2021	5
2	Skatteverket	11 oktober 2021	3
3	Försäkringskassan	11 oktober 2021	6
4	Kriminalvården	18 oktober 2021	2
5	DIGG	19 oktober 2021	1
6	Arbetsförmedlingen	20 oktober 2021	1



DO

www.do.se

Box 4057

169 04 Solna

Telefon 08-120 20 700

[facebook.com/Diskrimineringsombudsmannen](https://www.facebook.com/Diskrimineringsombudsmannen)

ISBN 978-91-88175-27-4

