

RÄTT FÖRUTSÄTTNINGAR FÖR FRAMTIDEN

ARTIFICIELL INTELLIGENS OCH GRUNDLÄGGANDE RÄTTIGHETER

SAMMANFATTNING



3

Viktiga slutsatser och FRA:s yttranden

5

Skydd av grundläggande rättigheter
– räckvidd, konsekvensbedömningar
och ansvarsskyldighet

10

Icke-diskriminering, dataskydd och tillgång
till rättslig prövning: tre övergripande teman

© Europeiska unionens byrå för grundläggande rättigheter, 2021.

Återgivning tillåten med angivande av källan.

För spridning eller användning av foton eller annat material som inte omfattas av upphovsrätten för Europeiska unionens byrå för grundläggande rättigheter måste tillstånd sökas direkt från upphovsrättsinnehavarna.

Varken Europeiska unionens byrå för grundläggande rättigheter eller någon person som agerar på byråns vägnar är ansvarig för hur nedanstående uppgifter används.

Luxemburg: Europeiska unionens publikationsbyrå, 2021.

Print	ISBN 978-92-9461-236-6	doi:10.2811/292766	TK-04-20-657-SV-C
PDF	ISBN 978-92-9461-203-8	doi:10.2811/186836	TK-04-20-657-SV-N

Foton:

Omslag: © HQUALITY/Adobe Stock

Sidan 1: © Mykola Mazuryk/AdobeStock

Sidan 3: © Mimi Potter/AdobeStock

Sidan 4: © Sikov/Adobe Stock

Sidan 7: © Gorodenkoff/Adobe Stock

Sidan 9: © VideoFlow/Adobe Stock

Sidan 11: © Monsitj/Adobe Stock

“AI är fantastiskt, men vi måste lära oss att använda den.”

(Privat företag, Spanien)

“Det finns en risk med att lita alltför mycket på maskiner.”

(Offentlig förvaltning, Frankrike)

Artificiell intelligens (AI) används allt mer i såväl den privata som den offentliga sektorn, vilket påverkar det dagliga livet. Vissa ser AI som slutet på människans kontroll över maskiner. Andra betraktar den som den teknik som kommer att hjälpa mänskligheten att ta itu med några av sina mest akuta utmaningar. Även om ingen av dessa beskrivningar är korrekt är det tydligt att oron över AI:s inverkan på de grundläggande rättigheterna ökar, vilket motiverar att människorättsaktörer granskar hur AI används.

FRA:s rapport *Getting the future right – Artificial intelligence and fundamental rights* (Rätt förutsättningar för framtiden – artificiell intelligens och grundläggande rättigheter) ger en ögonblicksbild av den nuvarande användningen av AI-relaterad teknik i EU och innehåller en analys av dess konsekvenser för de grundläggande rättigheterna. Den är inriktad på användningsfall inom fyra huvudområden – sociala förmåner, prediktivt polisarbete, hälso- och sjukvårdstjänster och riktad reklam.

I denna sammanfattning presenterar vi de viktigaste slutsatserna i rapporten.

Att definiera AI

Det finns ingen allmänt accepterad definition av AI. Begreppet avser inte konkreta tillämpningar, utan snarare den senaste tekniska utvecklingen som omfattar en rad olika tekniker.

FRA har inte tillämpat någon strikt definition av AI på de användningsfall som byrån presenterar i huvudrapporten. För intervjuerna definierades AI brett, med hänvisning till definitionen från kommissionens högnivåexpertgrupp för artificiell intelligens (AI-expertgruppen):

“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. AI-baserade system kan vara helt programvarubaserade och fungera i den virtuella världen (till exempel röstassistenter, bildanalysprogram, sökmotorer, tal- och ansiktsgenkänningsystem), eller inbäddas i hårdvaruenheter (till exempel avancerade robotar, självkörande bilar, drönare eller tillämpningar för sakernas internet).”

Denna **ursprungliga definition** från AI-expertgruppen var föremål för ytterligare diskussioner i gruppen. Se AI-expertgruppen (2019), **A definition of AI: Main capabilities and disciplines**. (En definition av AI: förmågor och vetenskapliga områden).

Vad omfattade studien?

FRA genomförde fältstudier i fem av EU:s medlemsstater: Estland, Finland, Frankrike, Nederländerna och Spanien. Byrån samlade in information från dem som utformar och använder AI-system i viktiga delar av de privata och offentliga sektorerna om hur de hanterar relevanta frågor om grundläggande rättigheter.

Under studien – som bygger på 91 personliga intervjuer – samlades information in om

- AI-teknikens syfte och praktiska tillämpning,
- de bedömningar som görs vid användning av AI och den tillämpliga rättsliga ramen och tillsynsmekanismerna,
- medvetenheten om frågor som rör grundläggande rättigheter och potentiella skyddsåtgärder och
- framtidsplaner.

Därtill intervjuades tio experter på övervakning eller observation av potentiella kränkningar av de grundläggande rättigheterna i samband med användningen av AI, däribland företrädare för det civila samhället och tillsynsorgan samt jurister.

En mer detaljerad beskrivning av forskningsmetoden och de frågor som ställdes under intervjuerna finns i bilaga 1 till huvudrapporten, vilken är tillgänglig på [FRA:s webbplats](#).

“Det viktigaste är att hantera ärenden på ett effektivare sätt. Det handlar om att använda din personal, de personer som handlägger ärenden, så effektivt som möjligt.”

(Offentlig förvaltning, Nederländerna)

“När vi testade systemet tittade vi egentligen inte på de rättsliga aspekterna utan på om systemet är lönsamt.”

(Privat företag, Estland)

FRA:s arbete med AI, stordata och grundläggande rättigheter

FRA:s rapport om AI och grundläggande rättigheter är den viktigaste publikationen från FRA:s **projekt om artificiell intelligens, stordata och grundläggande rättigheter**. Syftet med projektet är att bedöma de positiva och negativa konsekvenserna för de grundläggande rättigheterna av ny teknik, inbegripet artificiell intelligens och stordata.

Rapporten bygger på resultaten i ett antal tidigare dokument:

- **Facial recognition technology: fundamental rights considerations in the context of law enforcement** (Teknik för ansiktsgenkänning: grundläggande rättigheter i samband med brottsbekämpning) (2019): i detta dokument beskrivs och analyseras de utmaningar i fråga om grundläggande rättigheter som uppstår när offentliga myndigheter använder ansiktsgenkänning i realtid för att bekämpa brott. Rapporten innehåller också en kortfattad beskrivning av åtgärder för att förhindra kränkningar av rättigheter.
- **Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights** (Datakvalitet och artificiell intelligens – att minska systematiska och andra fel för att värna om de grundläggande rättigheterna) (2019): i detta dokument betonas vikten av att vara medveten om och undvika dålig datakvalitet.
- **#BigData: Discrimination in data-supported decision making** (Stordata: diskriminering vid dataunderstött beslutsfattande) (2018): i detta fokusdokument diskuterar man hur sådan diskriminering kan uppkomma och föreslår möjliga lösningar.

Som en del av projektet undersöker FRA också möjligheten att studera konkreta exempel på utmaningar som rör grundläggande rättigheter när algoritmer används för beslutsfattande, antingen genom online-experiment eller simuleringsstudier.

Viktiga slutsatser och FRA:s yttranden

Ny teknik har i grunden förändrat vårt sätt att organisera och leva våra liv. Ny datadriven teknik har särskilt främjat utvecklingen av artificiell intelligens (AI) och ökad automatisering av uppgifter som vanligtvis utförs av människor. Covid-19-krisen har lett till en ökad användning av AI och datadelning – vilket har skapat nya möjligheter, men också utmaningar och hot mot mänskliga och grundläggande rättigheter.



Utvecklingen inom AI har fått stor uppmärksamhet av media, civilsamhället, den akademiska världen, människorättsorgan och beslutsfattare. En stor del av denna uppmärksamhet är inriktad på dess potential att stödja ekonomisk tillväxt. Hur olika tekniker kan påverka de grundläggande rättigheterna har ägnats mindre intresse. Hittills har vi inte några omfattande empiriska belegg för det stora spektrum av rättigheter som AI påverkar, eller för de skyddsåtgärder som krävs för att se till att användningen av AI i praktiken är förenlig med de grundläggande rättigheterna.

Den 19 februari 2020 offentliggjorde Europeiska kommissionen vitboken *Om artificiell intelligens – en EU-strategi för spetskompetens och förtroende*. I den beskrivs huvudprinciperna för ett framtida EU-regelverk för AI i Europa. I vitboken konstateras att det är mycket viktigt att ett sådant regelverk bygger på grundläggande europeiska värden, däribland respekten för de mänskliga rättigheterna – artikel 2 i fördraget om Europeiska unionen (EU-fördraget).

FRA:s rapport om AI och grundläggande rättigheter stöder detta mål genom att analysera hur användningen av AI påverkar de grundläggande rättigheterna. På grundval av konkreta AI-användningsfall inom utvalda områden inriktar den sig på situationen ute på fältet när det gäller utmaningar och möjligheter i fråga om grundläggande rättigheter vid användning av AI.

Rapporten bygger på 91 intervjuer med tjänstemän inom offentlig förvaltning och personal i privata företag i utvalda EU-medlemsstater. De tillfrågades om sin användning av AI, sin medvetenhet om frågor som rör grundläggande rättigheter i sammanhanget och om metoder för att bedöma och minska riskerna med användning av AI.

Dessutom genomfördes 10 intervjuer med experter som på olika sätt hanterar de potentiella utmaningarna med AI:s konsekvenser för grundläggande rättigheter. I denna grupp ingick offentliga organ (såsom tillsynsmyndigheter och övervakande myndigheter), icke-statliga organisationer och jurister.



Rättslig ram

Den övergripande ram för grundläggande rättigheter* som gäller för användningen av AI i EU består av EU-stadgan om de grundläggande rättigheterna och Europeiska konventionen om de mänskliga rättigheterna.

Flera andra instrument från Europarådet och internationella instrument som rör de mänskliga rättigheterna är också relevanta i sammanhanget. Dessa inbegriper den allmänna förklaringen om de mänskliga rättigheterna från 1948 och de viktigaste FN-konventionerna om mänskliga rättigheter.**

Dessutom bidrar sektorspecifik sekundärlagstiftning, särskilt EU:s regelverk om dataskydd och EU:s lagstiftning mot diskriminering, till att skydda de grundläggande rättigheterna i samband med AI. Slutligen är även medlemsstaternas nationella lagstiftning tillämplig.

* Mer information finns i FRA (2012), *Bringing rights to life: The fundamental rights landscape of the European Union* (Att konkretisera rättigheter: situationen i fråga om grundläggande rättigheter i Europeiska unionen), Luxemburg, Europeiska unionens publikationsbyrå.

** Dessa konventioner är bland annat den internationella konventionen om medborgerliga och politiska rättigheter från 1966, den internationella konventionen om ekonomiska, sociala och kulturella rättigheter från 1966, den internationella konventionen om avskaffande av alla former av rasdiskriminering från 1965, konventionen om avskaffande av all slags diskriminering av kvinnor från 1979, konventionen mot tortyr från 1984, konventionen om barnets rättigheter från 1989, konventionen om rättigheter för personer med funktionsnedsättning från 2006 och den internationella konventionen till skydd för alla personer mot påtvingade försvinnanden från 2006.

Mer information om den internationella lagstiftningen om allmänt erkända mänskliga rättigheter och verkställighetsmekanismerna finns till exempel i De Schutter, O. (2015), *International Human Rights Law: Cases, Materials, Commentary* (Den internationella lagstiftningen om mänskliga rättigheter: fall, material och kommentarer), Cambridge, Cambridge University Press, 2nd edition.

SKYDD AV GRUNDLÄGGANDE RÄTTIGHETER – RÄCKVIDD, KONSEKVENSBEDÖMNINGAR OCH ANSVARSSKYLDIGHET

Att beakta de grundläggande rättigheternas fulla räckvidd i fråga om AI

Användningen av AI-system berör ett brett spektrum av grundläggande rättigheter, oavsett tillämpningsområde. Det handlar bland annat om integritet, dataskydd, icke-diskriminering och tillgång till rättslig prövning.

EU-stadgan om de grundläggande rättigheterna (nedan kallad *stadgan*) blev rättsligt bindande i december 2009 och har samma rättsverkan som EU-fördragen. Den omfattar medborgerliga, politiska, ekonomiska och sociala rättigheter i en enda text. Enligt artikel 51.1 i stadgan ska unionens institutioner, organ och byråer respektera alla rättigheter som fastställs i stadgan. EU:s medlemsstater måste göra detta när de genomför unionsrätten. Detta gäller för AI på samma sätt som för alla andra områden.

Fältarbetet under denna studie visar att många olika system används på AI-området. De analyserade teknikerna innefattar olika grader av automatisering och komplexitet. De varierar också i fråga om omfattning och potentiell inverkan på människor.

FRA:s studier visar att användningen av AI-system berör ett brett spektrum av grundläggande rättigheter, oavsett tillämpningsområde. Det handlar bland annat om integritet, dataskydd, icke-diskriminering och tillgång till rättslig prövning. Trots detta visar intervjuerna att diskussionerna om räckvidden av AI:s inverkan på grundläggande rättigheter ofta begränsas till specifika rättigheter.

Ett bredare spektrum av rättigheter behöver beaktas när AI används, beroende på teknik och användningsområde. Utöver rättigheter som rör integritet och dataskydd, likabehandling och icke-diskriminering samt tillgång till rättslig prövning skulle andra rättigheter kunna beaktas. Det handlar till exempel om människans värdighet, rätten till social trygghet och socialt stöd, rätten till god förvaltning (som främst är relevant för den offentliga sektorn) och konsumentskydd (särskilt viktigt för företagen). Beroende på AI-användningens sammanhang måste andra rättigheter som skyddas av stadgan beaktas.

FRA:s YTTRANDE 1

Vid införandet av ny politik och antagandet av ny lagstiftning om AI måste EU-lagstiftaren och medlemsstaterna, inom ramen för EU-rätten, säkerställa respekt för hela spektrumet av grundläggande rättigheter, i enlighet med stadgan och EU-fördragen. Politiska åtgärder och lagstiftning måste åtföljas av särskilda skyddsåtgärder för grundläggande rättigheter.

I samband med detta bör EU och dess medlemsstater förlita sig på tillförlitliga bevis för AI:s inverkan på de grundläggande rättigheterna för att säkerställa att eventuella begränsningar av vissa grundläggande rättigheter överensstämmer med principerna om nödvändighet och proportionalitet.

Relevanta skyddsåtgärder måste föreskrivas i lag för att ge ett effektivt skydd mot godtyckliga ingrepp i grundläggande rättigheter och för att ge både AI-utvecklare och AI-användare rättslig säkerhet. Frivilliga system för observation och skydd av grundläggande rättigheter vid utveckling och användning av AI kan ytterligare bidra till att minska kränkningarna av rättigheter. I linje med minimikraven på klarhet angående rättsläget – som är grundläggande för upprätthållandet av rättsstatsprincipen och en förutsättning för att säkerställa de grundläggande rättigheterna – måste lagstiftaren iaktta vederbörlig omsorg vid fastställandet av räckvidden för sådana AI-lagar.

Med tanke på den mångfald av teknik som inbegrips i begreppet AI och bristen på kunskap om hela omfattningen av dess potentiella inverkan på de grundläggande rättigheterna kan den rättsliga definitionen av AI-relaterade termer behöva utvärderas regelbundet.

FRA:s YTTRANDE 2

EU-lagstiftaren bör överväga att göra konsekvensbedömningar som täcker hela spektrumet av grundläggande rättigheter obligatoriska. Dessa bör omfatta den privata och den offentliga sektorn och genomföras innan något AI-system används. Konsekvensbedömningarna bör ta hänsyn till AI-teknikens varierande art och omfattning, inbegripet graden av automatisering och komplexitet samt den potentiella skadan. Det bör ställas grundläggande krav på granskning, vilka också kan bidra till att öka medvetenheten om potentiella konsekvenser för de grundläggande rättigheterna.

Konsekvensbedömningarna bör bygga på vedertagen god praxis från andra områden och upprepas med jämna mellanrum under användningen när så är lämpligt. Bedömningarna bör genomföras på ett genomblickbart sätt. Resultaten och rekommendationerna bör i möjligaste mån vara allmänt tillgängliga. För att underlätta genomförandet av konsekvensbedömningarna bör företag och offentliga förvaltningar vara skyldiga att samla in den information som behövs för en grundlig bedömning av de potentiella konsekvenserna för de grundläggande rättigheterna.

EU och medlemsstaterna bör överväga riktade åtgärder för att stödja dem som utvecklar, använder eller planerar att använda AI-system, för att säkerställa att de fullgör sina skyldigheter i fråga om konsekvensbedömningar avseende grundläggande rättigheter. Sådana åtgärder kan inbegripa finansiering, utfärdande av riktlinjer, utbildning och medvetenhetshöjande åtgärder. De bör särskilt, men inte uteslutande, rikta sig till den privata sektorn.

EU och medlemsstaterna bör överväga att använda befintliga verktyg, såsom checklistor eller självutvärderingsverktyg, som har utvecklats på europeisk och internationell nivå. Till dessa hör de verktyg som har utvecklats av EU:s högnivå-expertgrupp för artificiell intelligens.

Att använda effektiva konsekvensbedömningar för att förhindra negativa effekter

Förhandsbedömningar av konsekvenserna är främst inriktade på tekniska frågor. De tar sällan upp potentiella effekter på de grundläggande rättigheterna. Detta beror på att det saknas kunskap om hur AI påverkar sådana rättigheter.

Användningen av AI-system berör ett brett spektrum av grundläggande rättigheter, oavsett tillämpningsområde. Enligt artikel 51.1 i stadgan ska EU:s medlemsstater respektera alla rättigheter som fastställs i stadgan när de tillämpar unionsrätten. I linje med befintliga internationella standarder – särskilt FN:s vägledande principer om företag och mänskliga rättigheter – bör företag ha "en process för tillbörlig aktsamhet i fråga om mänskliga rättigheter för att identifiera, förebygga, begränsa och redogöra för hur de hanterar deras inverkan på de mänskliga rättigheterna" (principerna 15 och 17). Detta gäller oberoende av deras storlek och sektor och omfattar företag som arbetar med AI.

Samtidigt som EU fullföljer sina åtaganden när det gäller dessa vägledande principer har EU antagit flera lagstiftningsakter som behandlar sektorsspecifika instrument, särskilt i syfte att fullgöra skyldigheter att visa tillbörlig aktsamhet i fråga om mänskliga rättigheter. Diskussioner pågår för närvarande om förslag till ny sekundärlagstiftning för EU. Sådan lagstiftning skulle innebära att företagen måste genomföra bedömningar av tillbörlig aktsamhet avseende sin verksamhets och sina leveranskedjors potentiella konsekvenser för mänskliga rättigheter och miljön. Sådan lagstiftning skulle sannolikt vara sektorsövergripande och föreskriva sanktioner vid bristande efterlevnad – vilket bör omfatta användning av AI. Se FRA:s nya rapport *Business and Human rights – access to remedy* (Företagen och de mänskliga rättigheterna – tillgången till rättsmedel) där bättre övergripande bestämmelser om omsorg om mänskliga rättigheter i EU-baserade företag efterlyses.

Konsekvensbedömningar är ett viktigt verktyg för både företag och offentliga förvaltningar när det gäller att mildra verksamhetens potentiella negativa effekterna på de grundläggande rättigheterna. Enligt EU-rätten

inom specifika sektorer måste vissa former av konsekvensbedömningar göras, såsom konsekvensbedömningar avseende dataskydd enligt den allmänna dataskyddsförordningen. Många intervjuade personer rapporterade att de konsekvensbedömningar avseende dataskydd som krävs i lagen genomförs. Formerna för dessa varierar dock. Dessutom är förhandsbedömningarna, när de genomförs, främst inriktade på tekniska aspekter. De tar sällan upp potentiella effekter på de grundläggande rättigheterna. Enligt vissa intervjuade personer görs inte konsekvensbedömningar avseende grundläggande rättigheter när ett AI-system inte påverkar, eller inte tycks påverka, de grundläggande rättigheterna negativt.

Studien visar att de intervjuade personernas kunskaper om grundläggande rättigheter – förutom dataskydd och i viss utsträckning icke-diskriminering – är begränsade. De flesta är emellertid medvetna om att användningen av AI påverkar de grundläggande rättigheterna. Vissa intervjuade personer anger att deras system inte påverkar de grundläggande rättigheterna, vilket i viss utsträckning sammanhänger med de uppgifter som AI-systemen används för.

Alla svars personer är medvetna om dataskyddsfrågor. De flesta svars personerna inser också att diskriminering – generellt sett – kan vara ett problem när AI används. Den exakta innebörden och tillämpligheten av rättigheter som rör dataskydd och icke-diskriminering är dock fortfarande oklar för många svars personer.

Forskningsresultaten visar på skillnader mellan den privata och den offentliga sektorn. Intervjuade från den privata sektorn är ofta mindre medvetna om det bredare spektrum av grundläggande rättigheter som kan påverkas. Det finns kunskap om dataskyddsfrågor i den privata sektorn. Andra rättigheter, såsom icke-diskriminering och rättigheter som har med tillgång till rättslig prövning att göra, är dock mindre kända bland företagsrepresentanter som arbetar med AI. Vissa var fullt medvetna om potentiella problem, men andra sade att ansvaret för att bevaka frågor som rör grundläggande rättigheter ligger hos deras kunder.



FRA:s YTTRANDE 3

EU och medlemsstaterna bör se till att det finns effektiva system för ansvarsfördelning så att det är möjligt att övervaka och vid behov effektivt hantera AI-systemens eventuella negativa inverkan på de grundläggande rättigheterna. Utöver konsekvensbedömningar avseende grundläggande rättigheter (se FRA:s yttrande 2) bör de överväga att införa särskilda skyddsåtgärder för att säkerställa att ansvarsfördelningen är ändamålsenlig. Detta skulle kunna innebära ett rättsligt krav på att tillgängliggöra den information som krävs för att göra en bedömning av AI-systemens inverkan på de grundläggande rättigheterna. Detta skulle göra det möjligt för behöriga organ att bedriva extern övervakning och tillsyn av de mänskliga rättigheterna.

EU och medlemsstaterna bör också bättre utnyttja befintliga expertorgan för tillsyn i syfte att skydda de grundläggande rättigheterna vid användning av AI, såsom dataskyddsmyndigheter, likabehandlingsorgan, nationella människorättsinstitutioner, ombudsmän och konsumentskyddsorgan.

Ytterligare resurser bör öronmärkas för att inrätta effektiva system för ansvarsfördelning genom att höja kompetensen hos och diversifiera tillsynsorganens personalstyrkor. Därigenom skulle de kunna hantera komplicerade frågor som hänger samman med utveckling och användning av AI.

På samma sätt bör lämpliga organ tilldelas tillräckliga resurser, befogenheter och – vilket är mycket viktigt – expertis för att förebygga och bedöma kränkningar av grundläggande rättigheter och effektivt stödja personer vars grundläggande rättigheter påverkas av AI.

Genom att underlätta samarbetet mellan lämpliga organ på nationell och europeisk nivå kan man bidra till utbyte av sakkunskap och erfarenheter. Kontakter med andra aktörer med relevant sakkunskap, till exempel specialiserade organisationer i det civila samhället, kan också vara till hjälp. När medlemsstaterna genomför sådana åtgärder på nationell nivå bör de överväga att använda tillgängliga finansieringsmekanismer inom EU.

Att säkerställa en effektiv tillsyn och övergripande ansvarsskyldighet

Företag och offentliga förvaltningar som utvecklar och använder AI är i kontakt med olika organ som ansvarar för tillsynen av AI-relaterade system inom ramen för sina respektive befogenheter och sektorer. Dessa organ inbegriper dataskyddsmyndigheter. Men de som använder AI är inte alltid säkra på vilka organ som ansvarar för tillsynen av AI-system.

I linje med väletablerade internationella människorättsnormer – till exempel artikel 1 i Europeiska konventionen om skydd för de mänskliga rättigheterna (ECHR) och artikel 51 i stadgan – är staterna skyldiga att säkerställa människors rättigheter och friheter. För att uppfylla detta krav måste staterna bland annat införa effektiva mekanismer för övervakning och kontroll av efterlevnaden. Detta gäller även för AI.

När det gäller övervakningen pekar resultaten på den viktiga roll som specialiserade organ som har inrättats inom specifika sektorer spelar, vilka också ansvarar för tillsynen av AI inom ramen för sina befogenheter. Dessa är till exempel banktillsynsorgan och dataskyddsmyndigheter. En rad olika sådana organ kan vara relevanta för tillsynen av AI med avseende på grundläggande rättigheter. Ansvarsfördelningen för tillsynen av AI är dock fortfarande oklar för många av de personer från den privata och den offentliga sektorn som intervjuades.

Offentliga förvaltningars användning av AI granskas ibland som en del av deras regelbundet återkommande revisioner. Privata företag inom vissa sektorer har också specialiserade tillsynsorgan, till exempel när det gäller hälso- och sjukvård och finansiella tjänster. Dessa granskar också användningen av AI och därmed sammanhängande teknik, till exempel som en del av sina certifieringssystem. Intervjupersoner från den privata sektorn uttryckte en önskan om att organ skulle kunna ge expertråd om möjligheterna med och lagligheten hos potentiella AI-användningar.

I EU finns det en rad väl utvecklade oberoende organ med mandat att skydda och främja grundläggande rättigheter. Det rör sig bland annat om dataskyddsmyndigheter, likabehandlingsorgan, nationella människorättsinstitutioner och ombudsmän. Studien visar att de som använder eller planerar att använda AI

ofta har kontaktat olika organ om sin användning av AI, till exempel konsumentskyddsorgan.

Oftast kontaktade AI-användarna dataskyddsmyndigheter för att få vägledning, synpunkter eller godkännande i fråga om behandlingen av personuppgifter. Intervjuade experter lyfter fram dataskyddsmyndigheternas betydelse för tillsynen av AI-system när det gäller användningen av personuppgifter. De noterar dock också att dataskyddsmyndigheterna har otillräckliga resurser för denna uppgift och saknar särskild sakkunskap om AI-frågor.

Experter, inbegripet de som arbetar för tillsynsorgan som likabehandlingsorgan och dataskyddsmyndigheter, är överens om att befintliga tillsynsorgan behöver större sakkunskap för att kunna utöva en effektiv tillsyn av AI-relaterade frågor. Enligt experterna kan detta vara en utmaning, med tanke på att dessa organs resurser redan är knappa. De framhöll också den viktiga roll som spelas av relevanta organisationer i det civila samhället som är specialiserade på teknik, digitala rättigheter och algoritmer. De kan öka ansvarsskyldigheten vid användningen av AI-system.



IKKE-DISKRIMINERING, DATASKYDD OCH TILLGÅNG TILL RÄTTSLIG PRÖVNING: TRE ÖVERGRIPANDE TEMAN

Studien visar att användningen av AI påverkar olika grundläggande rättigheter. Förutom sammanhangsberoende specifika aspekter som påverkar olika rättigheter i varierande utsträckning ansågs följande frågor om grundläggande rättigheter av flera svarspersoner i studien vara tillämpliga på de flesta användningsfallen: behovet av att säkerställa icke-diskriminerande användning av AI (rätten att inte diskrimineras), kravet på lagenlig hantering av uppgifter (rätt till skydd av personuppgifter) och möjligheten att klaga på AI-baserade beslut och söka gottgörelse (rätt till ett effektivt rättsmedel och till en rättvis rättegång).

De två viktigaste grundläggande rättigheterna som framhölls under intervjuerna var dataskydd och icke-diskriminering. Dessutom togs effektiva sätt att klaga på användningen av AI upp upprepade gånger, kopplat till rätten till en rättvis rättegång och till ett effektivt rättsmedel. Följande tre FRA-yttranden, som speglar dessa resultat, bör läsas tillsammans med de övriga yttrandena, där vi efterlyser ett mer omfattande erkännande av och åtgärder i fråga om hela spektrumet av grundläggande rättigheter som påverkas av AI.

FRA:s YTTRANDE 4

EU:s medlemsstater bör överväga att uppmuntra företag och offentliga förvaltningar att bedöma potentiellt diskriminerande resultat när de använder AI-system.

Europeiska kommissionen och medlemsstaterna bör överväga att tillhandahålla finansiering för riktad forskning om potentiellt diskriminerande effekter av användningen av AI och algoritmer. Sådan forskning skulle gynnas av en anpassning av etablerade metoder inom den samhällsvetenskapliga forskningen som används för att identifiera potentiell diskriminering på olika områden – från rekrytering till kundprofilering.

Med utgångspunkt i resultaten av sådan forskning bör vägledning och verktyg utvecklas för att hjälpa dem som använder AI att upptäcka eventuella diskriminerande resultat.

Särskilda skyddsåtgärder för att säkerställa icke-diskriminering vid användning av AI

De intervjuade personerna nämnde sällan att de gjorde detaljerade bedömningar av potentiell diskriminering när de använde AI. Detta tyder på att det saknas fördjupade bedömningar av sådan diskriminering vid automatiserat beslutsfattande.

Skyldigheten att respektera principen om icke-diskriminering fastställs i artikel 2 i EU-fördraget, artikel 10 i EUF-fördraget (som ålägger unionen att bekämpa diskriminering på ett antal olika grunder) och artiklarna 20 och 21 i stadgan (likhet inför lagen och icke-diskriminering på en rad grunder). Mer specifika och detaljerade bestämmelser i flera olika EU-direktiv med varierande tillämpningsområden bygger också på denna princip.

Automatisering och användning av AI kan avsevärt effektivisera tjänster och utöka omfattningen av uppgifter som människor inte skulle kunna utföra. Det är dock nödvändigt att se till att tjänster och beslut som bygger på AI inte är diskriminerande. Med tanke på detta betonade Europeiska kommissionen nyligen behovet av ytterligare lagstiftning för att säkerställa icke-diskriminering vid användning av AI i **EU:s handlingsplan mot rasism 2020–2025**.

De flesta intervjuade är i princip medvetna om att diskriminering kan förekomma. Ändå tog de sällan upp denna fråga själva. Endast ett fåtal anser att deras system faktiskt skulle kunna diskriminera.

De intervjuade personerna nämnde också sällan detaljerade bedömningar av potentiell diskriminering, vilket innebär att det saknas fördjupade bedömningar av potentiell diskriminering.

De intervjuade personerna nämnde också sällan detaljerade bedömningar av potentiell diskriminering, vilket innebär att det saknas fördjupade bedömningar av potentiell diskriminering.

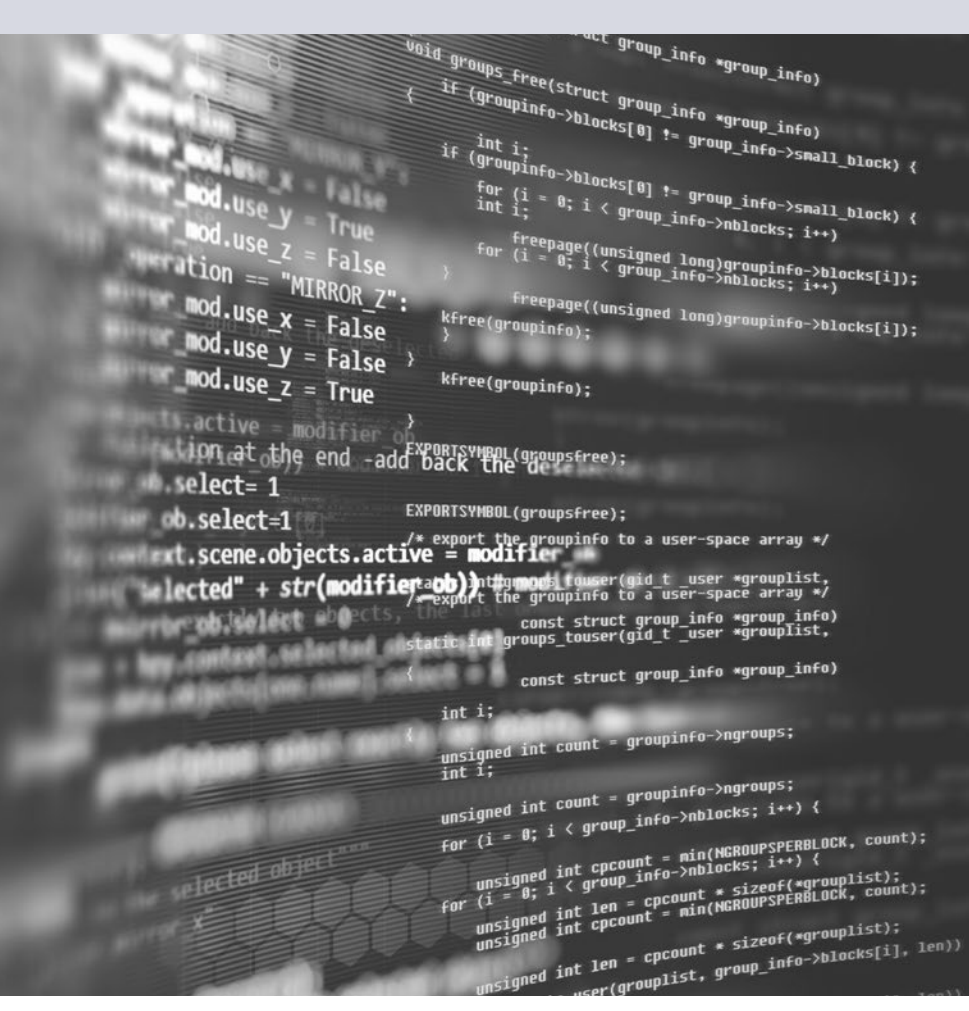
En allmän uppfattning är att man kan garantera att ett AI-system inte diskriminerar genom att utelämna information om skyddade egenskaper såsom kön, ålder eller etniskt ursprung. Detta är dock inte nödvändigtvis fallet. Information som kan indikera skyddade egenskaper (proxyvariabler), vilken ofta finns i datamängder, kan leda till diskriminering.

I vissa fall kan AI-system också användas för att testa och upptäcka diskriminerande beteende, som kan vara kodat i datamängder. Mycket få intervjuade personer nämnde dock möjligheten att samla in sådan information om missgynnade grupper för att upptäcka eventuell diskriminering. I avsaknad av en djupanalys av potentiell diskriminering i den faktiska användningen av AI-system finns det heller nästan ingen diskussion och analys av de potentiella positiva effekterna av att använda algoritmer för att fatta rättvisare beslut. Dessutom nämnde ingen av de intervjuade personer som arbetar med AI att användningen av AI för att upptäcka eventuell diskriminering ger positiva resultat, i den meningen att diskriminering kan upptäckas bättre när data analyseras med avseende på eventuell partiskhet.

Eftersom det fortfarande är svårt att upptäcka potentiell diskriminering genom användning av AI och algoritmer och de intervjuade personerna endast kort berörde denna fråga krävs det olika åtgärder för att ta itu med detta. Några sådana är att ställa krav på hänsyn till diskrimineringsfrågor vid bedömning av användningen av AI och att investera i ytterligare studier av potentiell diskriminering som använder en rad olika metoder.

Det kan till exempel handla om test av diskriminering. Detta skulle kunna bygga på liknande etablerade metoder för att testa partiskhet i vardagslivet, till exempel när det gäller jobbansökningar, där sökandens namn ändras för att (indirekt) visa etnicitet. När det gäller AI-tillämpningar skulle sådana tester kunna omfatta att skapa falska profiler för nätverktyg, som endast skiljer sig åt när det gäller skyddade egenskaper. På så sätt kan resultaten kontrolleras med avseende på potentiell diskriminering. Avancerade statistiska analyser skulle också kunna utnyttjas i forskningen för att upptäcka skillnader i datamängder för skyddade grupper som kan användas som grund för att undersöka potentiell diskriminering.

Slutligen framhölls det i vissa intervjuer i studien att resultaten av komplexa maskininlärningsalgoritmer ofta är mycket svåra att förstå och förklara. Därför kan ytterligare forskning som syftar till att bättre förstå och förklara sådana resultat (så kallad förklarande AI) också bidra till att förbättra möjligheterna att upptäcka diskriminering när AI används.



FRA:s YTTRANDE 5

För att värna om de grundläggande rättigheterna bör Europeiska dataskyddsstyrelsen (EDPB) och Europeiska datatillsynsmannen (EDPS) överväga att ge ytterligare vägledning om och stöd till ett effektivt genomförande av de bestämmelser i den allmänna dataskyddsförordningen som är direkt tillämpliga på användningen av AI, särskilt när det gäller vad som räknas som personuppgifter och användningen av dem i AI, inbegripet i AI-datamängder för utbildning.

Det råder stor osäkerhet om innebörden av automatiserat beslutsfattande och rätten till mänsklig granskning i samband med användningen av AI och automatiserat beslutsfattande. EDPB och EDPS bör därför också överväga att ytterligare klargöra begreppen "automatiserat beslutsfattande" och "mänsklig granskning", när de nämns i EU-lagstiftningen.

Dessutom bör nationella dataskyddsorgan ge praktisk vägledning om hur dataskyddsbestämmelserna ska tillämpas på användningen av AI. Sådan vägledning skulle kunna omfatta rekommendationer och checklistor, baserade på konkreta fall av användning av AI, för att stödja efterlevnaden av dataskyddsbestämmelserna.

ting att säga till om när beslut automatiseras. Kunskapen om denna rättighet är betydligt större bland dem som arbetar med AI – majoriteten av de intervjuade tog upp denna fråga. Många av de intervjuade, däribland experter, hävdade dock att det behövs större klarhet om räckvidden och innebörden av rättsliga bestämmelser om automatiserat beslutsfattande.

När det gäller sociala förmåner nämnde de intervjuade personerna endast ett exempel på helt automatiserade, regelbaserade beslut. Alla andra tillämpningar som nämndes granskas av människor. Intervjuade personer inom den offentliga förvaltningen betonade vikten av att alla beslut granskas av människor. De beskrev dock sällan vad en sådan granskning faktiskt innebär och hur annan information användes vid granskningen av resultaten från AI-system.

De intervjuade personerna är inte överens om huruvida den befintliga lagstiftningen är tillräcklig eller inte, men många efterlyste en mer konkret tolkning av de befintliga dataskyddsreglerna när det gäller automatiserat beslutsfattande, i enlighet med artikel 22 i den allmänna dataskyddsförordningen.

Mer vägledning om dataskydd

Det behövs större klarhet om räckvidden och innebörden av rättsliga bestämmelser om automatiserat beslutsfattande.

Dataskydd är avgörande för utvecklingen och användningen av AI. I artikel 8.1 i stadgan och artikel 16.1 i EUF-fördraget föreskrivs att var och en har rätt till skydd av sina personuppgifter. Den allmänna dataskyddsförordningen och dataskyddsdirektivet för brottsbekämpande myndigheter (direktiv (EU) 2016/680) vidareutvecklar denna rättighet och innehåller många bestämmelser som är tillämpliga på användningen av AI.

De intervjuade personerna uppgav att de flesta AI-system som de använder hanterar personuppgifter, vilket innebär att dataskyddet påverkas på många olika sätt. I ett fåtal tillämpningar används dock inte personuppgifter, eller endast anonymiserade uppgifter, enligt de intervjuade personerna och därmed är inte dataskyddslagstiftningen tillämplig. Om personuppgifter används är alla principer och bestämmelser om dataskydd tillämpliga.

Denna rapport belyser en viktig fråga med anknytning till dataskydd, som också är relevant för andra grundläggande rättigheter i samband med automatiserat beslutsfattande. Enligt en Eurobarometerundersökning vet endast 40 procent av européerna att de har någon-

Faktisk tillgång till rättslig prövning i mål som rör AI-baserade beslut

För att på ett effektivt sätt kunna bestrida beslut som grundar sig på användning av AI måste man veta om att AI används och hur och var man kan klaga. Organisationer som använder AI måste kunna förklara hur deras AI-system och beslutsfattande baserat på AI fungerar.

Tillgång till rättslig prövning är både ett medel och ett mål, och avgörande för enskilda personer som vill utnyttja andra processuella och materiella rättigheter. Ett antal grundläggande mänskliga rättigheter berörs, bland andra rätten till en rättvis rättegång och till ett effektivt rättsmedel enligt artiklarna 6 och 13 i Europeiska konventionen om de mänskliga rättigheterna och artikel 47 i EU-stadgan om de grundläggande rättigheterna. Kravet på tillgång till rättslig prövning innebär således att staterna är skyldiga att säkerställa varje individs rätt att vända sig till domstol – eller, under vissa omständigheter, till ett alternativt tvistlösningsorgan – för att utverka en åtgärd om det konstateras att den enskildes rättigheter har kränkts.

I enlighet med dessa standarder måste ett offer för en kränkning av de mänskliga rättigheterna till följd av att en offentlig eller privat enhet har utvecklat eller använt ett AI-system ges tillgång till rättsmedel inför en nationell myndighet. I linje med relevant rättspraxis enligt artikel 47 i stadgan och artikel 13 i Europeiska konventionen om de mänskliga rättigheterna måste rättsmedlet vara effektivt ”i såväl faktiskt som rättsligt hänseende”.

Forskningsresultaten visar att förutsättningar för att rättsmedlet ska vara effektivt i faktiskt hänseende i fall som rör AI-system och deras inverkan på de grundläggande rättigheterna är att alla är medvetna om när AI används och vet hur och var man kan klaga. Organisationer som använder AI måste se till att allmänheten får information om deras AI-system och om de beslut som fattas på grundval av dem.

Resultaten visar att det kan vara svårt att förklara hur AI-system fungerar och hur beslut fattas på lekmannaspråk. Immateriella rättigheter kan förhindra tillhandahållandet av detaljerad information om hur en algoritm fungerar. Dessutom är vissa AI-system komplexa. Detta gör det svårt att ge meningsfull information om hur ett system fungerar och om beslutsfattandet.

En del av de intervjuade företagen hanterar detta problem genom att helt och hållet undvika att använda komplicerade metoder för visst beslutsfattande, eftersom de inte skulle kunna förklara besluten. Alternativt hanterar de det genom att använda enklare metoder för dataanalys för samma problem i syfte att få en viss förståelse av de viktigaste faktorer som påverkar vissa resultat. Några av de intervjuade inom den privata sektorn pekade på de ansträngningar som har gjorts för att successivt öka deras förståelse för AI-teknik.

FRA:s YTTRANDE 6

EU-lagstiftaren och medlemsstaterna bör säkerställa faktisk tillgång till rättslig prövning för enskilda personer i ärenden som rör AI-baserade beslut.

För att se till att tillgängliga rättsmedel kan användas i praktiken skulle EU-lagstiftaren och medlemsstaterna kunna överväga att införa en rättslig skyldighet för offentliga förvaltningar och privata företag som använder AI-system att ge dem som söker gottgörelse information om hur deras AI-system fungerar. Detta inbegriper information om hur dessa AI-system kommer fram till automatiserade beslut. Denna skyldighet skulle bidra till att skapa lika villkor i fall där enskilda personer söker rättvisa. Den skulle också bidra till att effektivisera den externa övervakningen av AI-system och övervakningen av konsekvenserna för de mänskliga rättigheterna (se FRA:s yttrande 3).

Med tanke på svårigheten att förklara komplexa AI-system bör EU tillsammans med medlemsstaterna överväga att utarbeta riktlinjer i syfte att stödja ansträngningarna för att skapa öppenhet på detta område. De bör därvid utnyttja sakkunskapen hos nationella människorättsorgan och organisationer i det civila samhället som är verksamma på detta område.



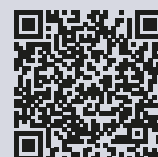
I denna sammanfattning presenteras de viktigaste resultaten från FRA:s rapport *Getting the future right – Artificial intelligence and fundamental rights* (Rätt förutsättningar för framtiden – artificiell intelligens och grundläggande rättigheter). Huvudrapporten är tillgänglig på [FRA:s webbplats](#).



PROMOTING AND PROTECTING YOUR FUNDAMENTAL RIGHTS ACROSS THE EU

FRA:s rapport om artificiell intelligens och grundläggande rättigheter innehåller konkreta exempel på hur företag och offentliga förvaltningar i EU använder eller försöker använda AI. Den är inriktad på fyra huvudområden – sociala förmåner, prognostiserat polisarbete, hälso- och sjukvårdstjänster och riktad reklam. Rapporten innehåller en diskussion om potentiella konsekvenser för de grundläggande rättigheterna och en analys av hur dessa rättigheter beaktas när AI-tillämpningar används eller utvecklas.

I den här sammanfattningen presenteras de viktigaste resultaten i rapporten. De kan ligga till grund för EU:s och medlemsstaternas politiska ansträngningar att reglera användningen av AI-verktyg i överensstämmelse med mänskliga och grundläggande rättigheter.



FRA – EUROPEISKA UNIONENS BYRÅ FÖR GRUNDLÄGGANDE RÄTTIGHETER
Schwarzenbergplatz 11 – 1040 Wien – Österrike
Tfn +43 158030-0 – Fax +43 158030-699

fra.europa.eu

[facebook.com/fundamentalrights](https://www.facebook.com/fundamentalrights)
twitter.com/EURightsAgency
[linkedin.com/company/eu-fundamental-rights-agency](https://www.linkedin.com/company/eu-fundamental-rights-agency)



Europeiska unionens
publikationsbyrå