

EUROPEISK METOD FÖR ATT BEDÖMA ALGORITMISK DISKRIMINERING



I samarbete med
Interfederala centret för lika möjligheter (Unia), Belgien,
Diskrimineringsombudsmannen (YVV), Finland och
Kommissionen för medborgarskap och jämställdhet (CIG), Portugal

Louise Hooper
Ylja Remmits
Ioanna Papageorgiou

Co-funded
by the European Union



EUROPEAN UNION

COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

Co-funded and implemented
by the Council of Europe

EUROPEISK METOD FÖR ATT BEDÖMA ALGORITMISK DISKRIMINERING

**Louise Hooper
Ylja Remmits
Ioanna Papageorgiou**

**I samarbete med
Interfederala centret för
lika möjligheter (Unia), Belgien,
Diskriminerings-
ombudsmannen (YVV), Finland och
Kommissionen för medborgarskap
och jämställdhet (CIG), Portugal**

Denna publikation har framställts med ekonomiskt stöd från Europeiska unionen och Europarådet. Innehållet är författarnas eget ansvar. De åsikter som uttrycks i publikationen kan under inga omständigheter anses återspegla Europeiska unionens eller Europarådets officiella ståndpunkt.

Det är tillåtet att återge utdrag (upp till 500 ord) förutom för kommersiella ändamål, förutsatt att textens integritet bevaras, att utdraget inte används ur sitt sammanhang, inte ger ofullständig information eller på annat sätt vilseleder läsaren vad gäller textens karaktär, omfattning eller innehåll. Källtexten måste alltid anges enligt följande: "© Europarådet, publiceringsår".

Alla andra förfrågningar om reproduktion eller översättning av hela eller delar av dokumentet ska skickas till: Publications and Visual Identity Division, Council of Europe (F-67075 Strasbourg Cedex eller publishing@coe.int).

Övrig korrespondens som rör detta dokument ska skickas till:
AI Co-ordinator of the Anti-discrimination and Inclusion Department of the Council of Europe, F-67075 Strasbourg Cedex
E-post: anti-discrimination@coe.int

Omslag och layout: Publications and Visual Identity Division Council of Europe
Layout av engelsk version: Luminess
Layout av svensk version: BGRAPHIC, Copenhagen
Översättning: Lingvo House Translation Services Ltd
Juridisk granskning av översättningen: Elina Sagne-Ollikainen
Fotografier: Shutterstock, Unsplash and Better Images of AI

© Europarådet, september 2026

Innehållsförteckning

FÖRKORTNINGAR	5
TERMINOLOGI	6
TACK	8
SAMMANFATTNING	9
INLEDNING	11
Viktiga delar i den nya rättsliga ramen	12
Riskkategorier enligt AI-förordningen	16
Hur används riskkategorier i metodiken?	18
Skyddade egenskaper	19
HUR METODIKEN ANVÄNDS	21
Steg 1	22
Steg 2	22
Steg 3	22
Steg 4	23
STEG 1 – IDENTIFIERING AV AI- ELLER ALGORITMISKA SYSTEM	25
Hur algoritmiska system och AI kan identifieras	26
1.1. Inledande informationsinsamling	27
1.2. Vidare dokumentgranskning	34
1.3. Identifiering av algoritmiska system eller AI-system med hög risk	37
1.4. Fortsatt handläggning av ärendet	40
STEG 2 – INFORMATIONSSAMLING OM DET ALGORITMISKA SYSTEMET ELLER AI-SYSTEMET FÖR ATT IDENTIFIERA DISKRIMINERING	43
2.1. Algoritmiska system	47
2.2. Ytterligare information som är tillgänglig för AI-system med hög risk	58
STEG 3 – METOD FÖR ATT BEDÖMA DEN INFORMATION SOM HAR ERHÅLLITS	75
Introduktion till begrepp	77
Hur steg 3 används	85
3.1. Vilka är de kända eller misstänkta riskerna för diskriminering?	86
3.2. Finns det några kända skyddade egenskaper eller proxyvariabler i indata eller datauppsättningar?	90
3.3. Bedömning av risker för grundläggande rättigheter (riskbedömningar, HRIA, FRIA och DPIA)	93
3.4. Vad är det algoritmiska systemet avsett att göra och hur fungerar det?	99
3.5. Tränings-, test- och valideringsdata	104
3.6. Testning och utvärdering	111
3.7. Bedömning av om bevisningen är tillräcklig för att stöda ett prima facie-fall av diskriminering som involverar ett AI- eller algoritmiskt system	124
3.8. Kan svaranden visa att skillnaden i behandling inte är diskriminerande?	127
STEG 4 – ÅTGÄRDER	135
4.1. Åtgärder I för alla algoritmiska system	137
4.2. Ytterligare åtgärder för AI-system med hög risk	143
4.3. Vad bör göras när ett system eller ett användningsfall medför en risk för grundläggande rättigheter men inte klassificeras som "hög risk" eller "förbjudet"?	145

REFERENSER	149
LITTERATUR	149
LAGSTIFTNING	155
RÄTTSPRAXIS	158
BILAGA A – MALL FÖR BEGÄRAN OM TILLGÅNG TILL PERSONUPPGIFTER (DSAR)	163
BILAGA B – MALL FÖR BEGÄRAN OM EN FÖRKLARING ENLIGT ARTIKEL 86 I AI-FÖRORDNINGEN	166
BILAGA C – EXEMPEL PÅ ANVÄNDNING AV ALGORITMER OCH AI	168
BILAGA D – EXEMPEL PÅ ALGORITMISK DISKRIMINERING ELLER RISK FÖR DISKRIMINERING	170
BILAGA E – PROXYVARIABLER	173
BILAGA F – FÖRTECKNING ÖVER AI- OCH ALGORITMREGISTER I EUROPA	176
BILAGA G – DIREKTIV OM LIKABEHANDLING , SKYDDADE GRUNDER OCH TILLÄMPNINGSOMRÅDE	178
BILAGA H – ANDRA POTENTIELLT RELEVANTA FÖRORDNINGAR OCH RIKTLINJER	180
BILAGA I – AI-SYSTEM MED HÖG RISK	182
BILAGA J – FÖRBJUDNA AI-PRAXIS	184
BILAGA K – DOKUMENTATIONSBLAD FÖR STEG 3	197

Förkortningar

ADR	Alternative dispute resolution (Alternativ tvistlösning)
ADM	Automated decision making (Automatiserat beslutsfattande)
AI	Artificial intelligence (Artificiell intelligens)
CJEU	Court of Justice of the European Union (Europeiska unionens domstol)
CSO	Civil society organisation (Organisation i det civila samhället)
DPA	Data protection authority (Dataskyddsmyndighet)
DPIA	Data protection impact assessment (Konsekvensbedömning av dataskydd)
DSAR	Data subject access request (Begäran om tillgång till personuppgifter)
EB	Equality body (Likabehandlingsorgan)
EDPB	European Data Protection Board (Europeiska dataskyddsstyrelsen)
FRIA	Fundamental rights impact assessment (Konsekvensbedömning av grundläggande rättigheter)
GDPR	General Data Protection Regulation (Allmänna dataskyddsförordningen)
HRIA	Human rights impact assessment (Konsekvensbedömning av de mänskliga rättigheterna)
LED	Law-enforcement directive (Brottsbekämpningsdirektivet)
MSA	Market surveillance authority (Marknadskontrollmyndighet)
RBI	Remote biometric identification (Fjärrbiometrisk identifiering)

Terminologi

I denna metod används följande termer med den betydelse som anges nedan, om inte annat framgår.

AI-system: "ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som kan uppvisa anpassningsförmåga efter driftsättning, och som, för uttryckliga eller underförstådda mål, utifrån de indata det tar emot härleder hur utdata, såsom förutsägelser, innehåll, rekommendationer eller beslut, ska genereras, vilka kan påverka fysiska eller virtuella miljöer" (artikel 3.1 i AI-förordningen).

AI-system med hög risk: eller "högrisk-AI", används för att beteckna ett AI-system enligt den uttömmande definitionen i artikel 6 i AI-förordningen.

Algoritmiskt system: en term som används i denna metod för att omfatta både AI-system och system som inte är AI-system, inklusive system för automatiserat beslutsfattande (ADM-system). Begreppet omfattar även AI-system som inte används i beslutsprocesser.

Anmält organ: detta är ett organ för bedömning av överensstämmelse som utför tredjepartsbedömningar av överensstämmelse, inklusive testning, certifiering och inspektion. Till exempel bedömer det anmälda organet överensstämmelsen hos högrisk-AI-system i enlighet med förfarandena för bedömning av överensstämmelse som anges i EU:s AI-förordning.

Direktiven om likabehandling: avser de EU-direktiv som anges i bilaga G.

Expert: en person med specialiserad kunskap eller teknisk sakkunskap som likabehandlingsorganet konsulterar eller anlitar för att bistå handläggaren i något skede av denna metod. Experten bidrar med specialistkunskap men är inte huvudansvarig handläggare.

Handläggare: en person vid ett likabehandlingsorgan som genomför bedömningar i något av metodens steg inom ramen för arbetet med att utreda diskrimineringsärenden.¹

Likabehandling: i denna metod används termen "likabehandling" i stället för "jämlighet" eller närliggande termer.

Icke-högrisk-AI-system: ett AI-system som inte ingår i den uttömmande definitionen av högrisk-AI enligt AI-förordningen. Observera att detta endast innebär att dessa system inte kategoriseras som högrisk enligt AI-förordningen, inte att dessa AI-system är riskfria.

1. Även om denna metod i första hand riktar sig till likabehandlingsorgan (EB) kan den även vara relevant för nationella människorättsinstitutioner (NHRI), ombudsmannainstitutioner och andra intressenter. När metoden tillämpas av nationella människorättsinstitutioner eller ombudsmannainstitutioner kan termen "handläggare" även avse handläggare vid dessa institutioner.

Indata: all information, alla signaler eller rådata som tillförs ett algoritmiskt system och som ligger till grund för det som utdata-systemet genererar. Exempelvis kan en AI-baserad applikation för cancerscreening som är avsedd att identifiera misstänkta födelsemärken använda ålder, kön och en bild av födelsemärket som indata. I ett system som bygger på en stor språkmodell (LLM), såsom ChatGPT, utgörs indata av texten i den aktuella chattkonversationen och, i förekommande fall, även text från tidigare chattkonversationer.

Klagande: den person eller den grupp av personer som påstår sig ha utsatts för eller som har upplevt (olaglig) diskriminering. För att säkerställa terminologisk överensstämmelse med ordalydelsen i GDPR och AI-förordningen motsvarar termerna "registrerad" respektive "berörd person" begreppet "klagande" när det hänvisas till de relevanta bestämmelserna i dessa förordningar.

Konfusionsmatris: en tabell som används för att utvärdera hur väl ett algoritmiskt system fungerar genom att jämföra systemets förutsägelser med de faktiska resultaten. Raderna i konfusionsmatrisen motsvarar de faktiska resultaten, medan kolumnerna motsvarar systemets förutsägelser. Tabellen visar hur många av förutsägelserna som var korrekta respektive felaktiga.

Kvalificerat ADM: för syftena med denna metodik används kvalificerat ADM för att definiera en underkategori av ADM-system som uppfyller kriterierna som definieras i artikel 22 i GDPR, för "enbart" automatiserat beslutsfattande (förordning (EU) 2016/679). Dessa system utlöser höjda transparens- och informationsskyldigheter enligt GDPR.

Leverantör/användare: termerna används som definieras i artikel 3 i AI-förordningen. Även om de är definierade med hänvisning till AI-system kan de respektive rollerna också tillämpas, analogt, på icke-AI-system.

Marknadskontrollmyndighet: den eller de nationella myndigheter som utför verksamhet och vidtar åtgärder enligt förordning (EU) 2019/1020 om (marknadskontroll och överensstämmelse för produkter).

Outputdata: resultatet av algoritmsystemet efter behandling av inputdata. Till exempel en förutsagd egenskap om en individ, en poäng, rekommendation, beslut eller innehåll som bilder, video eller text.

Svaranden: den organisation eller enhet som utreds av likabehandlingsorganet för olaglig diskriminering efter ett klagomål, ett ex officio-förfarande eller en anmälan från marknadskontrollmyndigheten (MSA). I samband med GDPR och AI-förordningen kan svaranden benämnas "datakontrollanter" och "AI-leverantörer" eller "användare", respektive, som definieras i dessa förordningar.

System för automatiserat beslutsfattande (ADM): termen omfattar ett brett spektrum av automatiserade system som används i samband med eller för beslutsfattande. De kan vara AI-baserade men kan också bygga på andra tekniker, även så enkla som grundläggande om-så-regler. De omfattar allt från helt automatiserade system utan mänsklig medverkan till halvautomatiserade system eller beslutsstödsystem, där både maskinella och mänskliga inslag ingår och som i stor utsträckning används i praktiken.

För terminologi relaterad till EU:s AI-förordning, se [artikel 3 i EU:s AI-förordning](#).

Tack

Publikationen utarbetades inom ramen för projektet "Upholding Equality och Non-Discrimination of Equality Bodies regarding the Use of Artificial Intelligence in Public Administrations", som samfinansieras av Europeiska unionen via Technical Support Instrument och av Europarådet. Projektet genomförs av Europarådet i samarbete med Europeiska kommissionen, Interfederala centret för lika möjligheter (Unia, Belgien), Diskrimineringsombudsmannen (YVV, Finland) och Kommissionen för medborgarskap och jämställdhet (CIG, Portugal).

Publikationen författades av Louise Hooper (internationell konsult, Europarådet), Ylja Remmits och Ioanna Papageorgiou (konsulter för algoritmrevision/internationella konsulter, Europarådet). Europarådet vill framföra sitt tack till projektets mottagande institutioner och till Europeiska kommissionen för deras kontinuerliga engagemang under hela utarbetningsprocessen, och i synnerhet till: Nele Roekens, Sophie Vincent, Thelma Raes och Laetitia Parmentier (Unia, Belgien); Tiina Valonen och Ville Rantala (YVV, Finland); Carla Peixe, Lúcia Pestana och Alexandra Andrade (CIG, Portugal); Massimiliano Santini (Reform and Investment Task Force, Generalsekretariatet, Europeiska kommissionen); samt Menno Ettema, Sara Haapalainen, Ayca Dibekoğlu och Delfine Gaillard (Avdelningen för antidiskriminering och inkludering, Europarådet).

Särskilt tack till Raphaële Xenidis, Kris Shrishak, Brent Mittelstadt, Chris Russell, Ydo Bergstra och Jurriaan Parie för insatser som granskare, och till Laurens Naudts och Plixavra Vogiatzoglou för expertrådgivning. Rapporten har också gynnats av diskussionerna med och bidragen från nationella experter från likabehandlingsorgan och civilsamhällsorganisationer som deltog i ett flernationsseminarium och workshop i Bryssel den 15 april 2026. Alla dessa perspektiv har berikat de analytiska och operativa avsnitten i dessa riktlinjer.



Sammanfattning

Algoritmsystem kan förstärka och ge upphov till diskriminering eftersom de ofta är utformade för att särskilja individer: klassificera, rangordna eller fokusera sig på dem. I kombination med historiska ojämlikheter och mänskliga fördomar som är inbäddade i data, regler och designval hos algoritmsystem kan diskriminerande mönster reproduceras och öka i omfattning ifall de inte proaktivt identifieras och begränsas.

Denna metodik ger steg-för-steg-vägledning till likabehandlingsorgan (EB) och andra människorättsorganisationer i Europa för att bedöma diskrimineringsfall som är kopplade till artificiell intelligens (AI) eller algoritmsystem. Den är avsedd för handläggare, rådgivare och jurister vid dessa organ som utreder klagomål gällande diskriminering där AI eller automatiserat beslutsfattande (ADM) är involverat eller misstänks vara det. Den har som mål att fungera som en referensförteckning över utmaningar och källor att beakta vid varje steg i en bedömning. Metodiken är uppdelad i fyra successiva steg:

Steg 1: identifiering av AI- eller algoritmfall

Steg 2: informationsinsamling om AI- eller algoritmsystemet
för att identifiera diskriminering

Steg 3: metodik för bedömning av mottagen information

Steg 4: åtgärder

Dessa steg bygger på det europeiska regelverket för AI, särskilt EU:s AI-förordning, diskrimineringslagstiftning och hänvisar till annan lagstiftning, såsom dataskydd, exempel på användningsfall och rättspraxis där det är relevant. Stegen särskiljer mellan högrisk-AI-system och andra system i enlighet med EU:s AI-förordning, vilket ger möjligheter att få tillgång till teknisk dokumentation, data och testning av högrisk-system, särskilt av organ som utsetts till "myndigheter som skyddar grundläggande rättigheter" enligt artikel 77 i EU:s AI-förordning.

Metodiken fokuserar inte på nationell lagstiftning, och i vissa länder kan nationell rätt ge starkare möjligheter att få tillgång till information eller utredningsbefogenheter när det gäller algoritmisk diskriminering. Handläggare bör ha detta i åtanke när de använder detta verktyg.

För att underlätta användningen av de fyra stegen innehåller metodiken en omfattande förteckning med bilagor som ger exempel på mallbrev för informationsförfrågningar, stödjer förståelsen av risker för algoritmisk diskriminering (t.ex. efter typ av AI-system eller sektor där systemet används) och hjälper till att bedöma information, t.ex. genom en checklista/registerblad för steg 3.

För att underlätta användningen av denna metodik är en grundläggande förståelse för AI och algoritmisk bias och diskriminering samt relaterad terminologi till hjälp, medan verktyget även ger en terminologilista. Andra publikationer som producerats inom projektet "Upholding Equality and Non-Discrimination by Equality Bodies

regarding the Use of Artificial Intelligence in Public Administrations” kompletterar denna metodik och kan stödja uppbyggnaden av denna grundläggande förståelse.

- ▶ [Rättsligt skydd mot algoritmisk diskriminering i Europa: nuvarande ramverk och kvarstående luckor](#), december 2025 (finns även på [franska](#), [nederländska](#), [portugisiska](#), [finska](#) och [svenska](#)).
 - ▶ [Europeiska policyriktlinjer om AI och algoritm driven diskriminering för lika-behandlingsorgan och andra nationella människorättsstrukturer](#), december 2025 (finns även på [franska](#), [nederländska](#), [portugisiska](#), [finska](#) och [svenska](#)).
-

Inledning

Nya regelverk om AI, såsom EU:s AI-förordning,² EU-direktiven om standarder för likabehandlingsorgan (nedan kallade "standarddirektiven")³ och Europarådets ramkonvention om artificiell intelligens och mänskliga rättigheter, demokrati och rättsstaten (nedan kallad "Europarådets ramkonvention om AI"), ger nya verktyg och befogenheter till likabehandlingsorgan och människorättsorganisationer för att utreda algoritmisk diskriminering.

Dessa verktyg kommer med vissa begränsningar och utmaningar. Till exempel kan vissa diskriminerande algoritmsystem inte klassificeras som AI enligt EU:s AI-förordning.⁴ Dessutom är många områden inom lagstiftningen om AI och diskriminering ännu inte helt klara. Slutligen är detta för många läsare och handläggare ett nytt och relativt komplext rättsområde.

Trots dessa utmaningar syftar denna fyrstegsmetodik till att stödja och ge en struktur för att systematiskt analysera den diskriminerande konsekvensen av ett beslut eller en behandling som potentiellt har fattats av eller med ett algoritmiskt system. Läsare bör inte se varje steg i metodiken som något som måste slutföras. Snarare utgör stegen en referenshandbok som innehåller en mängd verktyg för att identifiera system, samla in och bedöma bevis samt besluta vilken åtgärd som ska vidtas. Genom att följa stegen i metodiken och vid behov söka (teknisk) expertassistans bör det vara möjligt att ha de nödvändiga delarna på plats som krävs för att fastställa antagandet om diskriminering och söka upprättelse i fall av algoritmisk diskriminering.

Likabehandlingsorganen bör närma sig diskrimineringsklagomål som rör AI och ADM på samma sätt som alla andra diskrimineringsfall, men också inse att brist på transparens eller osäkerhet om hur ett algoritmiskt system fungerar kan beaktas när man avgör om det finns en risk för diskriminering eller om ett prima facie-fall har fastställts. Detta beror på en växande insikt om att det finns en ojämlik makt- och informationsbalans mellan företag som skapar AI-system, de som ansvarar för att driftsätta dem och dem som systemen riktas mot.⁵ I sådana fall kan bevisbördan övergå

2. Europaparlamentets och rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens och om ändring av förordningarna (EG) 2008/300, (EU) 2013/167, (EU) 2013/168, (EU) 2018/858, (EU) 2018/1139 och (EU) 2019/2144 samt direktiven 2014/90/EU, (EU) 2016/797 och (EU) 2020/1828 (AI-förordningen).
3. Rådets direktiv (EU) 2024/1499 av den 7 maj 2024 om normer för EB:er på området likabehandling av personer oavsett ras eller etniskt ursprung, likabehandling i arbetslivet oavsett religion eller övertygelse, funktionsnedsättning, ålder eller sexuell läggning, likabehandling av kvinnor och män i fråga om social trygghet samt tillgång till och tillhandahållande av varor och tjänster, och om ändring av direktiven 2000/43/EG och 2004/113/EG, samt Europaparlamentets och rådets direktiv (EU) 2024/1500 av den 14 maj 2024 om normer för EB:er på området likabehandling och lika möjligheter för kvinnor och män i arbetslivet och om ändring av direktiven 2006/54/EG och 2010/41/EU.
4. Vissa exempel på algoritmisk diskriminering bygger på enkla regelbaserade system som inte klassificeras som AI, såsom de nederländska exemplen på diskriminerande algoritmer i skandalen kring barnomsorgsbidrag (Amnesty International 2021) och riskbedömning av bedrägeri som användes av den nederländska myndigheten för utbetalning av stipendier (Rechtbank Overijssel 2024)
5. Detta behandlas utförligt i avsnitt 3.8.

till svaranden för att visa att algoritmsystemet inte diskriminerade eller bidrog till diskriminering på ett otillåtet sätt. För att drabbade personer ska ha effektiv tillgång till rättslig prövning krävs en högre grad av transparens.

Digital Omnibus om AI,⁶ ett lagstiftningspaket som ändrar och förenklar AI-förordningen och annan lagstiftning, antogs av Europaparlamentet. Där det är relevant hänvisas till dessa ändringar.

Viktiga delar i den nya rättsliga ramen

Standarddirektiven

Standarddirektiven, som ska vara införlivade i nationell rätt senast den 19 juni 2026 (artikel 24 i direktiven), fastställer minimikrav för likabehandlingsorganens verksamhet i syfte att öka deras effektivitet och säkerställa deras oberoende. Dessa befogenheter ska läsas tillsammans med de befogenheter som redan följer av EU:s direktiv om likabehandling.

I skäl 21 i direktiv (EU) 2024/1500 och skäl 22 i rådets direktiv (EU) 2024/1499 anges följande beträffande tilldelning av ekonomiska resurser:

Det är av avgörande betydelse att uppmärksamma de möjligheter och risker som användningen av automatiserade system, inbegripet artificiell intelligens, medför. Likabehandlingsorganen bör särskilt förses med lämpliga personalresurser och tekniska resurser. Dessa resurser bör bland annat göra det möjligt för likabehandlingsorganen att dels använda automatiserade system i sin verksamhet, dels bedöma sådana system i fråga om deras överensstämmelse medförbuden mot diskriminering. Om likabehandlingsorganet ingår i ett organ med flera uppdrag bör de resurser som krävs för att fullgöra dess uppdrag på likabehandlingsområdet säkerställas.⁷

AI-förordningen – artikel 77: befogenheter för myndigheter som skyddar grundläggande rättigheter och samarbete med marknadskontrollmyndigheter

Genom artikel 77 i AI-förordningen⁸ ges likabehandlingsorgan och nationella människorättsinstitutioner (NHRI) som uppfyller kraven i artikeln nya befogenheter i

6. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).
7. Artikel 26 i ramkonventionen om artificiell intelligens föreskriver också att likabehandlingsorgan ska vara effektiva, opartiska och oberoende samt ha tillräckliga resurser, befogenheter och sakkunskap.
8. Artikel 77.1 i AI-förordningen, inklusive ändringarna genom Digital Omnibus om AI: "Nationella offentliga myndigheter eller organ som utövar tillsyn över eller säkerställer efterlevnaden av skyldigheter enligt unionsrätten till skydd för grundläggande rättigheter, inbegripet rätten till likabehandling, ska ha befogenhet att från den berörda marknadskontrollmyndigheten begära och få tillgång till all information eller dokumentation som har upprättats eller bevarats enligt denna förordning, på ett begripligt språk och i ett maskinläsbart format samt på elektronisk väg, om tillgång till sådan information eller dokumentation är nödvändig för att de effektivt ska kunna fullgöra sina uppdrag inom ramen för sina respektive behörighetsområden. Denna artikel påverkar inte de berörda nationella offentliga myndigheternas eller organens behörighet, uppgifter, befogenheter eller oberoende enligt deras respektive uppdrag."

fråga om AI-system med hög risk, så att de kan uppfylla sitt mandat.⁹ Digital Omnibus om AI syftar till att förtydliga att dessa befogenheter inte påverkar dessa organs befintliga befogenheter eller deras oberoende.¹⁰

Förteckningen över myndigheter som skyddar grundläggande rättigheter borde ha anmälts till Europeiska kommissionen senast den 2 november 2024, och medlemsstaterna ska hålla förteckningen uppdaterad.¹¹ Kommissionen konstaterar att förteckningen över de organ som omfattas av artikel 77 fortlöpande uppdateras i takt med det pågående genomförandet av AI-förordningen.¹² EB:er som medlemsstaterna inte hade anmält före den första tidsfristen för anmälan kan fortfarande läggas till i förteckningen. EB:er kan även tas bort från förteckningen.

Berörda myndigheter har enligt artikel 77.1 i AI-förordningen befogenhet att begära och få tillgång till alla handlingar som har upprättats eller bevarats enligt AI-förordningen, i ett begripligt språk och format, om detta är nödvändigt för att de effektivt ska kunna fullgöra sina uppdrag inom ramen för sina respektive behörighetsområden. Detta innebär att begäran kan avse dokumentation som har upprättats eller bevarats av andra nationella behöriga myndigheter, inbegripet marknadskontrollmyndigheter och anmälda organ, samt av leverantör, användare eller andra aktörer i ett AI-systems livscykel. Digital Omnibus om AI förtydligar vidare att begäran om tillgång till information och dokumentation ska riktas till marknadskontrollmyndigheten, som "ska bevilja tillgång till sådan information och dokumentation, bland annat genom att begära in sådan information och dokumentation från leverantören eller den som sätter in tjänsten, när så är nödvändigt och utan onödigt dröjsmål"¹³ och att berörda myndigheter eller organ "ska samarbeta nära och bistå varandra med det ömsesidiga bistånd som krävs för att fullgöra sina respektive uppdrag".¹⁴ Detta samarbete omfattar även informationsutbyte när det är nödvändigt för en effektiv tillsyn över eller tillämpning av AI-förordningen och annan unionslagstiftning, däribland direktiven om likabehandling (se [bilaga G](#)).

Om dokumentationen inte är tillräcklig för att fastställa huruvida en överträdelse av skyldigheter enligt unionsrätten till skydd för grundläggande rättigheter har skett, får likabehandlingsorganet lämna in en motiverad begäran till

9. Alla likabehandlingsorgan omfattas inte av definitionen i artikel 77, även om många gör det. I denna metod används därför termen "artikel 77-myndighet" för att beteckna ett likabehandlingsorgan eller en nationell människorättsinstitution (NHRI) som uppfyller definitionen i artikel 77 i AI-förordningen.

10. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).

11. AI-förordningen, artikel 77.2.

12. Europeiska kommissionen. Förteckning över myndigheter med ansvar för skyddet av grundläggande rättigheter enligt AI-förordningen. Tillgänglig på: <https://digital-strategy.ec.europa.eu/en/policies/fundamental-rights-protection-authorities-ai-act> (hämtad den 1 juli 2026).

13. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).

14. Ibid.

marknadskontrollmyndigheten om att organisera teknisk provning av AI-systemet med hög risk (artikel 77.3 i AI-förordningen; se [steg 4](#)).

All information och dokumentation som erhålls ska behandlas i enlighet med sekretesskraven i artikel 78. Av artikel 78.2 framgår tydligt att den myndighet som avses i artikel 77 endast får begära in sådana uppgifter som är strikt nödvändiga för bedömningen av den risk som AI-system medför samt för utövandet av sina befogenheter enligt förordningen och unionsrätten. En begäran enligt AI-förordningen bör därför vara tydlig och avgränsad för att undvika att den avslås.

När uppgifter begärs in och erhålls ska likabehandlingsorganet säkerställa att adekvata och effektiva cybersäkerhetsåtgärder finns på plats för att skydda säkerheten och konfidentialiteten för den information och de uppgifter som erhålls. Sådana uppgifter ska raderas så snart de inte längre behövs för det ändamål för vilket de har inhämtats, i enlighet med tillämplig unionsrätt eller nationell rätt.

På motsvarande sätt anges i Europarådets normer, såsom ramkonventionen om artificiell intelligens och [Rekommendation CM/Rec\(2026\)1 om jämlikhet och artificiell intelligens](#), vilka åtgärder medlemsstaterna och andra berörda aktörer bör vidta för att främja jämlikhet samt förebygga och bekämpa diskriminering i alla verksamheter under AI-systemens hela livscykel.

Konfidentiell behandling

Bestämmelser om konfidentialitet finns i artikel 78.1 i AI-förordningen. De gäller för Europeiska kommissionen, marknadskontrollmyndigheter, anmälda organ (artikel 70) samt alla andra fysiska eller juridiska personer som medverkar vid tillämpningen av AI-förordningen. Dessa aktörer ska säkerställa att konfidentialiteten för den information och de uppgifter som erhålls respekteras på ett sätt som skyddar:

- ▶ en fysisk eller juridisk persons immateriella rättigheter samt konfidentiell affärsinformation eller företagshemligheter, inbegripen källkod (artikel 5 i direktiv (EU) 2016/943). För EB:er är artikel 5 i direktiv (EU) 2016/943 särskilt relevant, eftersom den innehåller ett undantag som gör det möjligt att röja missförhållanden, felaktigt handlande eller olaglig verksamhet, förutsatt att svaranden (EB:et eller, i förekommande fall, marknadskontrollmyndigheten) agerade i syfte att skydda allmänintresset (artikel 5b) eller ett berättigat intresse som erkänns i unionsrätten eller nationell rätt (artikel 5d). Skyddet av grundläggande rättigheter samt förebyggande och motverkande av diskriminering omfattas sannolikt av detta undantag;
- ▶ ett effektivt genomförande av förordningen, särskilt vid inspektioner, utredningar eller revisioner;
- ▶ offentliga säkerhetsintressen och nationella säkerhetsintressen;
- ▶ genomförandet av straffrättsliga eller administrativa förfaranden;
- ▶ information som har sekretessklassificerats enligt unionsrätten eller nationell rätt.

Det är för närvarande oklart hur dessa bestämmelser kan åberopas för att påverka EB:er möjlighet att få tillgång till information från berörda parter. Om utlämnande av

information nekas med hänvisning till konfidentialitet kan EB:erna dock vända sig till antingen en behörig myndighet enligt AI-förordningen eller dataskyddsmyndigheten enligt GDPR. Det är sannolikt att den behöriga tillsynsmyndigheten eller domstolen, när företagshemligheter åberopas som skäl för att inte lämna ut information, behöver göra en avvägning mellan de rättigheter och intressen som står mot varandra för att avgöra vilken information som ska lämnas ut (se *CK mot Dun & Bradstreet Austria GmbH och Magistrat der Stadt Wien* (2025), som behandlas i [steg 2.1](#)).

Sektorsövergripande samarbete

För att den nya rättsliga ramen ska få full verkan kommer det också att vara nödvändigt att etablera, utveckla, stärka och formalisera samarbetet med relevanta marknadskontrollmyndigheter, dataskyddsmyndigheter och andra myndigheter enligt artikel 77, om sådant samarbete inte redan finns. Kravet på samarbete formaliseras genom Digital Omnibus om AI genom införandet av artikel 77.1b.¹⁵ För detta ändamål kan det vara nödvändigt eller lämpligt att utarbeta gemensamma samverkansrutiner¹⁶ mellan myndigheter. Exempel och god praxis finns i rutan nedan. De nya direktiven om normer för EB föreskriver att nationella ramar ska inrättas för genomförande av utredningar, så att EB kan bedriva utredningar. Dessa ramar ska även omfatta lämpliga mekanismer för samarbete mellan EB och andra berörda offentliga myndigheter för detta ändamål (artikel 8 i direktiven om normer för EB).

Förhandlingar mellan EB och medlemsstaternas regeringar om att säkerställa en tydlig rättslig grund för informationsutbyte mellan myndigheter enligt artikel 77 – däribland EB och dataskyddsmyndigheter – samt mellan marknadskontrollmyndigheter och myndigheter enligt artikel 77, i enlighet med det samarbete som föreskrivs i artiklarna 73.7 och 79.2 i AI-förordningen, kommer att spela en central roll i detta avseende (Autoriteit Persoonsgegevens 2024, punkt 37). Förekomsten av en sådan rättslig grund är en förutsättning för att kunna ingå bilaterala och multilaterala samarbetsavtal om informationsutbyte. Sådana avtal bör utformas brett så att de omfattar ansvarsområdena för myndigheterna enligt artikel 77 och marknadskontrollmyndigheterna samt möjliggör kompletterande bilaterala överenskommelser i enskilda ärenden när detta behövs (Autoriteit Persoonsgegevens 2024, punkt 37).

Exempel och god praxis

I Storbritannien har Equality and Human Rights Commission samarbetat med Information Commissioner's Office (ICO) inom ramen för Fairness and Innovation Challenge. Samarbetet omfattade testfall inom högre utbildning, finanssektorn, hälso- och sjukvården samt rekrytering i syfte att utveckla nya metoder för att åtgärda statistisk, mänsklig och strukturell snedvridning samt diskriminering

15. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).

16. Se även artikel 26 i Europarådets ramkonvention om AI (2024), "Ett effektivt samarbete mellan institutioner måste underlättas", och artikel 14 i standarddirektiven som kräver lämpliga samarbetsmekanismer.

i AI-system. Initiativet finansierades av Storbritanniens vetenskapsministerium (Department for Science, Innovation and Technology) (Department of Science, Technology and Innovation 2024).

I Norge har likabehandlingsorganet samarbetat med dataskyddsmyndigheten inom ramen för en regulatorisk sandlåda för artificiell intelligens (Norska data-tilsynet 2023).

I Frankrike underlättas samarbetet mellan rättighetsombudet och dataskyddsmyndigheten, Commission Nationale de l'Informatique et des Libertés (CNIL), genom flera mekanismer: de båda institutionerna har ingått ett samförståndsavtal, och rättighetsombudet har dessutom en rådgivande roll gentemot CNIL:s styrelse. Tack vare denna strukturella koppling informeras Diskrimineringsombudsmannen om alla projekt som läggs fram vid CNIL:s styrelsemöten (Europarådet 2025b).

I Storbritannien fastställs i ett samförståndsavtal mellan Information Commissioner's Office (dataskyddsmyndigheten) och Equality and Human Rights Commission (jäinställdhetsmyndigheten) principerna för informationsutbyte och samarbete, de rättsliga grunderna för informationsutbyte, praktiska frågor rörande datakontroll och databehandling

Riskkategorier enligt AI-förordningen

EU:s AI-förordning bygger på ett regelverk som klassificerar AI-system utifrån den risknivå de antas medföra för grundläggande rättigheter, hälsa och säkerhet (AI-förordningen, artikel 1). Detta synsätt utgår från att alla AI-system inte innebär samma grad av risk. Vissa AI-system anses medföra oacceptabla risker och är därför helt förbjudna. Andra bedöms innebära betydande risker och omfattas därför av omfattande krav och tillsyn. Ytterligare andra anses inte kräva någon särskild AI-specifik reglering.

Riskkategorin avgör vilka skyldigheter som gäller – från ett absolut förbud, via omfattande tekniska krav, till enbart grundläggande transparenskrav. För likabehandlingsorgan är denna riskbaserade struktur av direkt betydelse eftersom den avgör vilken information och dokumentation som ska finnas tillgänglig och vilka åtgärder som står till buds enligt AI-förordningen.

Förbjudna AI-metoder

AI-förordningen förbjuder vissa AI-metoder som anses medföra oacceptabla risker för grundläggande rättigheter, säkerhet eller demokratiska värden (artikel 5 i AI-förordningen). Dessa omfattar bland annat:

- ▶ manipulativa tekniker som på ett väsentligt sätt påverkar människors beteende, inklusive generering av sexuellt eller intimt innehåll utan samtycke samt material som utgör sexuella övergrepp mot barn (CSAM);
- ▶ utnyttjande av människors sårbarhet;
- ▶ system för social poängsättning;

- ▶ riskbedömningar av fysiska personer i syfte att förutsäga brottslighet på grundval av profilering;
- ▶ biometrisk fjärridentifiering i realtid på allmän plats (med de begränsade undantag som anges i förordningen);
- ▶ känsloligenkänning i arbetslivet och inom utbildningssektorn;
- ▶ biometrisk kategorisering som används för att dra slutsatser om skyddade egenskaper;
- ▶ urskillningslös insamling av ansiktssbilder från internet eller övervakningskameror.

Se även [bilaga J – Förbjudna användningsområden för AI-system](#).

Högrisk-AI

Högrisk-AI-system omfattar användningsområden som anses medföra betydande risker för hälsa, säkerhet eller grundläggande rättigheter och som används inom de områden som anges i bilaga III till AI-förordningen (artikel 6). Dessa omfattar AI-system som används inom:

- ▶ biometri;
- ▶ förvaltning av kritisk infrastruktur;
- ▶ utbildning och yrkesutbildning;
- ▶ sysselsättning och arbetsledning;
- ▶ tillgång till grundläggande tjänster och förmåner;
- ▶ brottsbekämpning;
- ▶ migration, asyl och gränsförvaltning;
- ▶ rättskipning.

Observera: Ett AI-system som omfattas av bilaga III ska alltid betraktas som ett högrisk-AI-system om det används för profilering av en fysisk person.

Se även [bilaga I – Högrisk-AI-metoder](#)

Om ett AI-system är avsett att utföra någon av följande funktioner ska det inte betraktas som ett högrisk-AI-system (artikel 6.3 i AI-förordningen):

- ▶ en avgränsad processuell uppgift;
- ▶ att förbättra resultatet av en tidigare genomförd mänsklig aktivitet;
- ▶ att identifiera beslutsmonster eller avvikelser från tidigare beslutsmonster;
- ▶ en förberedande uppgift inför en bedömning.

Högrisk-AI-system omfattas av omfattande krav enligt AI-förordningen beträffande dokumentation, transparens och riskhanteringsförfaranden.

AI för allmänna ändamål (GPAI)

Kategorierna högrisk och förbjudna AI-metoder avser specifika användningsområden och bygger på AI-systemets uttryckliga användningsändamål. AI-modeller för allmänna ändamål (General-Purpos AI – GPAI), exempelvis stora språkmodeller (LLM), är däremot modeller som kan utföra ett brett spektrum av olika uppgifter och

integreras i många olika efterföljande system eller tillämpningar (AI-förordningen, artikel 3.63). GPAI har därför, per definition, inget specifikt användningsändamål. En GPAI-modell kan dock ligga till grund för en särskild tillämpning som i sig kan omfattas av kategorin högrisk-AI.

Den dokumentation som tas fram enligt reglerna för GPAI kommer normalt inte att vara tillgänglig för EB:er med stöd av AI-förordningen, eftersom rätten till tillgång enligt artikel 77 är begränsad till tillämpningar som utgör högrisk-AI-system. EB får i sådana fall i stället förlita sig på sina befintliga befogenheter enligt nationell rätt och annan unionsrätt. I de flesta situationer är det dock sannolikt att ett domstolsbeslut krävs för att få tillgång till sådan dokumentation. I denna metodik omfattas GPAI-modeller generellt av begreppet algoritmiska system.

Hur används riskkategorier i metodiken?

Som beskrivits ovan ger AI-förordningen nya möjligheter att få tillgång till information om högrisk-AI-system. Metodiken är därför i första hand uppbyggd kring skillnaden mellan högrisk-AI-system och andra algoritmiska system. När det gäller AI-system som inte klassificeras som högrisk förutsätts EB:er arbeta inom samma informationsramverk som gäller för övriga algoritmiska system. Steg 1 i metodiken hjälper handläggaren att identifiera vilken typ av system som granskas.

AI-förordningen ställer ytterligare regleringskrav på olika aktörer som är involverade i utveckling, distribution och införande av högrisk-AI-system. För dessa system kommer detta att ge betydande och relevant information om riskhanteringsåtgärder, det data som används och de åtgärder för datastyrning som finns på plats, teknisk dokumentation som förklarar viktiga val beträffande utformning och data samt systemets avsedda syfte, liksom information om testning och övervakning samt identifierade risker för diskriminering. Denna dokumentation kommer att utgöra en viktig resurs för likabehandlingsorgan som försöker identifiera eller bedöma potentiella fall av algoritmisk diskriminering.

Sådan information kommer inte att finnas tillgänglig i alla, eller ens de flesta, fall. Dessa krav hjälper dock likabehandlingsorganen att identifiera och tolka bevis beträffande AI och andra algoritmiska system där inga sådana formella krav finns. Detta beror på att de fungerar som referenssteg för att identifiera, minska eller eliminera risken för diskriminering och att dokumentationen ofta ingår i standardmässig programvaru- och teknisk utveckling.

Metodiken skiljer därför genomgående mellan de situationer där ett EB enligt artikel 77 i AI-förordningen har rätt att begära ut dokumentation om ett högrisk-AI-system och de situationer där någon sådan särskild rätt inte finns.

Det är samtidigt viktigt att komma ihåg att nationell lagstiftning i många medlemsstater ger likabehandlingsorganen egna befogenheter inom utrednings- och informationsinhämtning. Även likabehandlingsorgan som inte omfattas av artikel 77 kan därför ofta stödja sig på nationell rätt eller annan unionsrätt för att få tillgång till relevant dokumentation.

Skyddade egenskaper

Genom hela denna metodik används begreppet skyddade egenskaper. Med detta avses de egenskaper som skyddas enligt Europeiska konventionen om de mänskliga rättigheterna och Europeiska unionens direktiv om likabehandling.

Europeiska domstolen för de mänskliga rättigheterna har publicerat en [Vägledning om artikel 14 i Europakonventionen samt artikel 1 i protokoll nr 12 till konventionen](#).

Tabell 1. Skyddade egenskaper enligt europeisk rätt

X i tabellen innebär att egenskapen omfattas av respektive regelverk. En tom ruta innebär att egenskapen inte uttryckligen omfattas.

	Artikel 14 EKMR	EU:s stadga om de grundläggande rättigheterna	EU:s direktiv om likabehandling
Kön	X	X	X (2004/113/ EG, 2000/78/ EC, 2006/54/ EG, 2023/970/ EU, 2010/41/EU)
Raa¹⁷	X	X	X (2000/43/EG)
Färg	X	X	
Språk	X	X	
Religion	X	X (eller tro)	X (2000/78/EG)
Politisk eller annan åsikt	X	X	X (2000/78/EG)
Nationellt eller socialt ursprung	X	X (etniskt eller socialt ursprung)	
Koppling till en nationell minoritet	X	X (medlemskap)	
Fastighet	X	X	
Födelse	X	X	
Annan status	X		
Funktions- nedsättning	Annan status	X	X (2000/78/EG)
Ålder	Annan status	X	X (2000/78/EG)

17. Eftersom alla människor tillhör samma art avvisar Europarådets ministerkommitté teorier som bygger på att det skulle finnas olika "raser". I detta dokument används dock begreppet "ras" för att säkerställa att de personer som generellt och felaktigt uppfattas som "tillhörande en annan ras" inte utesluts från det skydd som lagstiftningen och genomförandet av politiska åtgärder ger.

	Artikel 14 EKMR	EU:s stadga om de grundläggande rättigheterna	EU:s direktiv om likabehandling
Sexuell läggning	Annan status	X	X (2000/78/EG)
Nationalitet	Annan status	X (kvalificerad)	
Genetiska egenskaper	Annan status	X	
Civilstånd	Annan status		
Migrant- eller flyktingstatus	Annan status		



Hur metodiken används

Metodiken bygger på en process av fyra steg (se även figur 1 nedan):

- Steg 1: identifiering av AI- eller algoritmiska ärenden
- Steg 2: informationsinsamling om AI- eller det algoritmiska systemet för att identifiera diskriminering
- Steg 3: metod för bedömning av den information som har samlats in
- Steg 4: åtgärder

Metodiken skiljer, när det är relevant, mellan tre olika utgångspunkter för ett ärende:

- A. ett klagomål (från en enskild person eller en grupp)
- B. ett initiativärende (*ex officio*)
- C. information som lämnats av en marknadskontrollmyndighet (MSA)

Observera: Stegen presenteras som på varandra följande steg för att tydliggöra metodens uppbyggnad. De följer inte på varandra i en rak ordning, utan är i praktiken ofta cirkulära och iterativa.

Steg 1

Det första steget består av att identifiera ett diskrimineringsärende där ett algoritmiskt system kan ha använts eller där det finns indikationer på att ett sådant system har varit inblandat.

Viktiga frågor:

- ▶ Är detta ett ärende som rör ett högrisk-AI-system enligt AI-förordningen?
- ▶ Är detta ett ärende som rör ett algoritmiskt system?

Steg 2

Nästa steg är att samla in information om det identifierade systemet som kan bidra till att fastställa om diskriminering förekommit. Steg 2 beskriver vilka typer av information, dokumentation och annan bevisning som kan finnas tillgänglig eller vara åtkomlig för EB:er samt vilka källor denna information kan hämtas från.

Observera: Den information som identifieras i detta steg analyseras mer ingående i steg 3. Därför bör steg 2 och steg 3 alltid ses i nära samband med varandra. Om information saknas och det finns en rättslig möjlighet att begära in den, behandlas detta under steg 4.

Viktig fråga:

- ▶ Vilken information finns tillgänglig och hur kan den nås?

Steg 3

I detta steg genomförs en fullständig analys av den relevanta bevisningen med hjälp av ett antal vägledande frågor. Frågorna är inte uttömmande.

Därefter tillämpas det rättsliga ramverket om likabehandling på den insamlade bevisningen.

Viktiga frågor:

- ▶ Kan ett *prima facie*-fall av diskriminering styrkas?
- ▶ Är det sannolikt att svaranden kan uppfylla sin bevisbörda?
- ▶ Har otillåten diskriminering styrkts när all bevisning vägs samman?

Steg 4

Överväg och tillämpa de åtgärder som står till buds.

Viktig fråga:

- Vilka åtgärder bör begäras eller tillämpas?

I samtliga steg behöver handläggaren dessutom beakta följande:

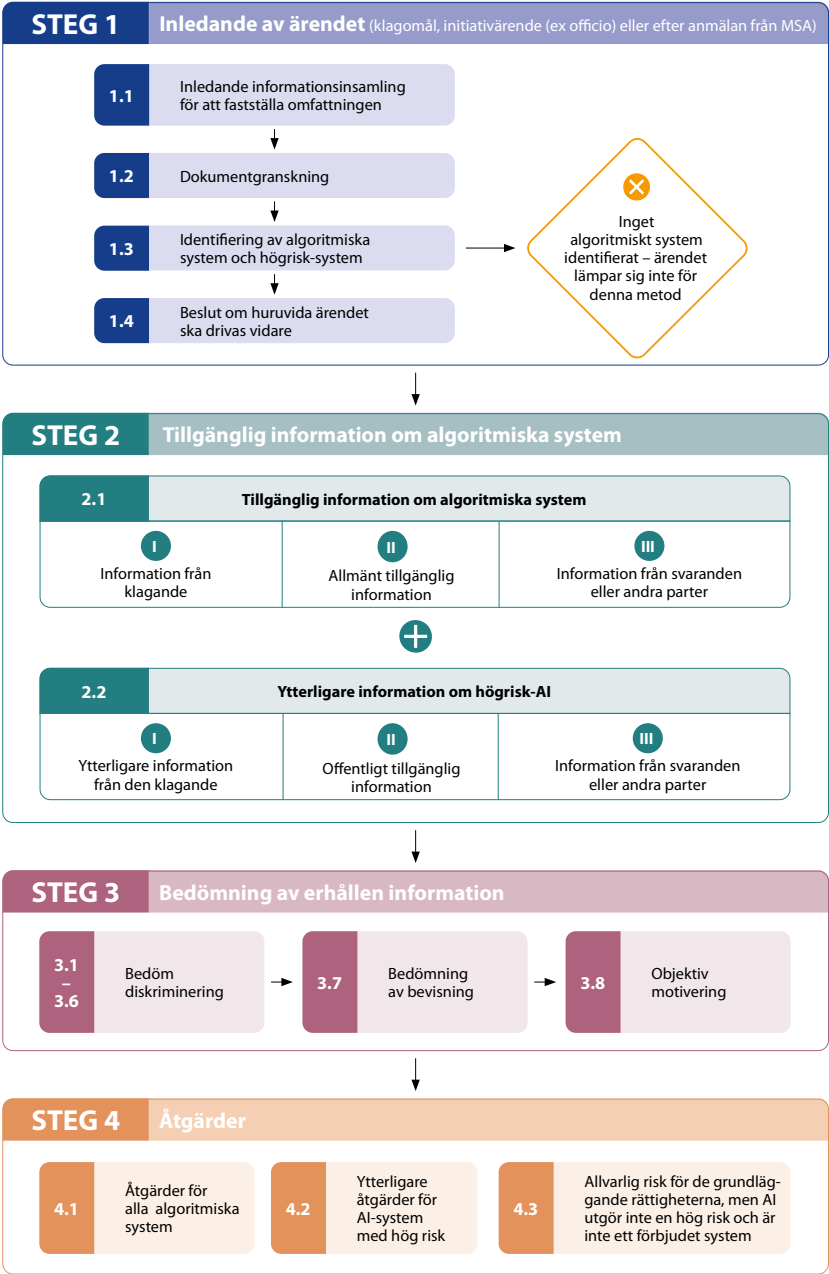
- Om det inte finns något försvarbart påstående om diskriminering, bör ärendet avslutas.
- Om inget algoritmiskt system är inblandat, bör ärendet avslutas inom ramen för denna metodik och i stället handläggas enligt EB:ets ordinarie rutiner för diskrimineringsärenden

Användning av experter

Det kan vara nödvändigt att anlita expertstöd för vissa delar av denna metodik, exempelvis när särskild kompetens inom datavetenskap eller datateknik krävs för att genomföra en teknisk bedömning av en modell eller de data som använts. Metodiken anger vilka typer av frågor som kan behöva besvaras av en expert och vid vilka tillfällen detta bör övervägas. Sedan ändringarna av artikel 77 genom Digital Omnibus om AI¹⁸ antogs kan EB dessutom ha möjlighet att begära stöd från marknadskontrollmyndigheter eller andra offentliga myndigheter och organ med särskild sakkunskap med stöd av bestämmelsen om ömsesidigt bistånd i den föreslagna artikel 77.1 b. tise in reliance on the mutual assistance provision in proposed Article 77(1)(b).

18. Se ovan, under "EU:s AI-förordning, artikel 77: Befogenheter för myndigheter som värnar om de grundläggande rättigheterna".

Figur 1. Metodikens fyra steg





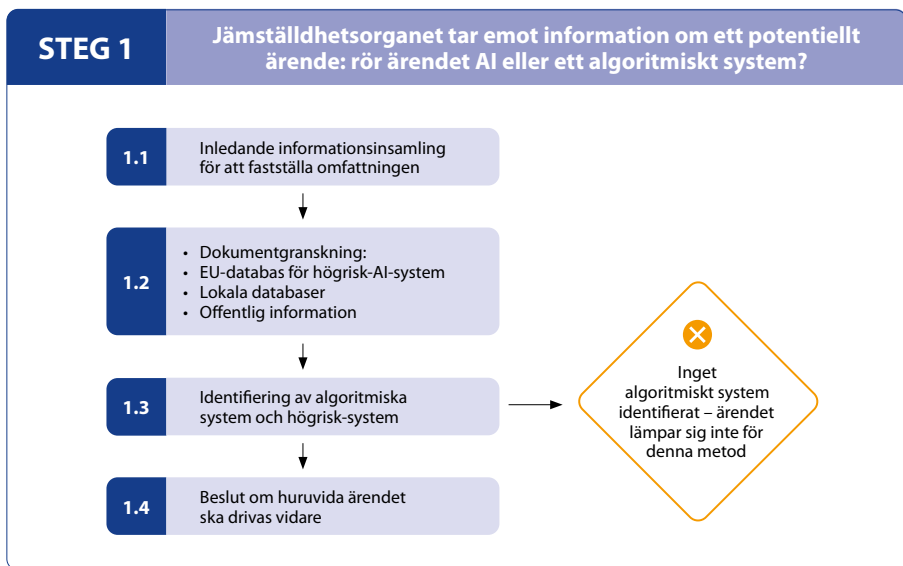
Steg 1

Identifiering av AI- eller algoritmiska system

Syftet med det första steget är enbart att fastställa om ett system för automatiserat beslutsfattande (ADM) eller ett (högrisk-) AI-system är inblandat samt att avgöra om det finns tillräckligt med information för att gå vidare till steg 2. I detta steg genomförs ingen detaljerad utredning.

Steg 1 består av följande delsteg:

Figur 2. Steg 1 – Innebär ärendet användning av AI eller ett algoritmiskt system?



Observera: Stegen presenteras en efter en för att ge metodiken en tydlig struktur. I praktiken är de ofta iterativa och processen kan behöva upprepas när ny information framkommer.

Identifieringen av AI- och ADM-ärenden skiljer sig åt beroende på hur ärendet har uppkommit:

- A. Genom ett klagomål (från en enskild person eller en grupp)
- B. Genom ett initiativärende (*ex officio*)
- C. Genom information från en MSA.

Hur algoritmiska system och AI kan identifieras

Algoritmiska system kan vara så enkla som automatisering av ett begränsat antal regler som definierats av människor. Generellt är algoritmiska system (inklusive AI) svåra att identifiera. Definitionerna har förändrats över tid och tolkas dessutom ofta på olika sätt.

I EU:s AI-förordning anges följande i den tekniska definitionen i artikel 3.1:

AI-system är ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som kan uppvisa anpassningsförmåga efter driftsättning och som, för uttryckliga eller underförstådda mål, utifrån de indata det tar emot härleder hur utdata, såsom prognoser, innehåll, rekommendationer eller beslut, ska genereras och som kan påverka fysiska eller virtuella miljöer.¹⁹

Även när det råder enighet om definitionerna är det vanligt att organisationer inte identifierar sina system som algoritmiska system eller använder andra benämningar.

Vid bedömningen av om ett algoritmiskt system används bör därför flera olika uttryck och beskrivningar beaktas.

Alla följande uttryck kan indikera att ett algoritmiskt system används:

- ▶ automatisering, maskinbaserat, datoriserat, algoritmbaserat;
- ▶ profilering eller skapande av profiler;
- ▶ automatiserat beslutsfattande eller beslutsstödsystem;
- ▶ ett system eller en dator fattar beslut, analyserar, jämför, rekommenderar eller behandlar en persons uppgifter;
- ▶ ett system, en dator eller en kamera ser, lyssnar eller förstår;
- ▶ varje hänvisning till beräkning av poäng, sannolikheter, risknivåer eller uppskattningar av framtida beteende;
- ▶ automatiserad riskbedömning;
- ▶ automatiserad igenkänning;
- ▶ hänvisningar till statistik;
- ▶ hänvisningar till maskininlärning.

19. Detta kompletteras av kommissionens riktlinjer C(2025) 5053 om definitionen av ett system för artificiell intelligens, som fastställts genom förordning (EU) 2024/1689 (AI-förordningen).

Följande uttryck kan särskilt indikera att AI används:

- ▶ maskininläring, djupinläring;
- ▶ att en modell, maskin eller dator har tränats;
- ▶ att mönster har lärts in från data;
- ▶ att systemet bygger på stora språkmodeller (LLM:er);
- ▶ prediktiva modeller.

Exempel: "På grundval av dessa uppgifter beräknar vår tjänsteleverantör sannolikhetsvärden med hjälp av matematisk-statistiska metoder för att bedöma risken för utebliven betalning och för att verifiera din adress."

1.1. Inledande informationsinsamling

A. Klagomål

När ett ärende grundar sig på ett klagomål sker först en inledande registrering av ärendet. I detta skede samlas grundläggande information in för att avgöra om ett algoritmiskt system eller ett AI-system kan vara inblandat. Denna inledande informationsinsamling är avsedd att vara enkel och lätt att integrera i likabehandlingsorganets EB:ets ordinarie rutiner för mottagande av klagomål.

1.1.1. Inledande bedömning av behörighet och tillämpningsområde

- ▶ Hur inkom klagomålet? (telefon, webbformulär, brev eller annat)
- ▶ Den klagandes namn och kontaktuppgifter.
- ▶ En kort beskrivning av problemet.
- ▶ Vilken organisation är svarande? (organisation som är föremål för klagomålet)
 - offentlig verksamhet eller privat aktör?
- ▶ Bedömning av behörighet:
 - skyddad diskrimineringsgrund?
 - sektor?
- ▶ Vilka åtgärder eller vilket syfte eftersträvas?
 - rådgivning?
 - rättslig prövning?
 - åtgärdsrekommendationer?

Bedöm därefter om ärendet ligger inom likabehandlingsorganets behörighet och om klagomålet redan har getts in till någon annan myndighet eller om någon annan aktör redan handlägger ärendet.

1.1.2. Insamling av ytterligare information från den klagande

Den klagande kan ha tillgång till mer information än vad som framgår av det ursprungliga klagomålet. Detta avsnitt innehåller exempel på frågor som kan ställas – exempelvis

via ett webbformulär eller ett telefonsamtal – för att avgöra om ett algoritmiskt system eller ett AI-system har använts.

Både AI-förordningen och GDPR innehåller bestämmelser om transparens som innebär att enskilda personer – registrerade eller berörda personer – ska informeras om användningen av algoritmiska system eller AI-system. Relevant information kan lämnas genom personlig kommunikation eller avtalsvillkor, vanligtvis i integritetspolicyer, användarvillkor eller på organisationens webbplats eller digitala plattformar (se avsnitt [1.2.3 Respondentens webbplats och externa kommunikation](#)).

Det är därför lämpligt att redan i detta skede be den klagande att lämna in relevanta handlingar. Tillsammans med en översiktlig granskning av offentligt tillgänglig information kan detta ge ett värdefullt underlag för nästa steg (se avsnitt [1.2 Vidare dokumentgranskning](#)).

Observera: Dessa bestämmelser omfattar inte alla situationer där AI eller algoritmiska system används. Att information saknas innebär därför inte nödvändigtvis att något sådant system inte har använts – det kan helt enkelt innebära att ytterligare utredning krävs.

AI-förordningen innehåller flera bestämmelser som kräver att människor informeras när AI används.

- ▶ När ett högrisk-AI-system används på arbetsplatsen ska arbetstagarnas företrädare och de arbetstagare som berörs informeras om att de omfattas av användningen av ett högrisk-AI-system (AI-förordningen artikel 26.7).
- ▶ När ett högrisk-AI-system fattar eller bistår vid beslut som rör en fysisk person ska användaren informera den berörda personen om att ett högrisk-AI-system används (AI-förordningen artikel 26.11). Informationen ska även omfatta systemets avsedda användningsområde och vilka typer av beslut det används för. Den som tillhandahåller tjänsten bör också informera fysiska personer om deras rätt till en förklaring enligt denna förordning (skäl 93 i AI-förordningen).
- ▶ När fysiska personer interagerar direkt med AI, exempelvis via en chattbot eller en röstassistent, ska de informeras om detta (AI-förordningen artikel 50.1). Informationen kan exempelvis ges genom ett meddelande såsom: "Hej, jag är en chattbot."/"Du chattar nu med en AI-assistent."/"Jag är X, en AI-assistent som hjälper dig."
- ▶ På samma sätt bör enskilda personer informeras om hur system för känsligenkänning (till exempel ett AI-system som identifierar kunders ilska eller tillfredsställelse utifrån ansikts- eller röstdrag) eller biometriska klassificeringssystem (till exempel ett AI-system som identifierar kön utifrån en persons hår- eller ansiktsform) fungerar (AI-förordningen, artikel 50.3).

GDPR

- ▶ GDPR innehåller också krav på transparens (GDPR, artiklarna 5.1 a, 13 och 14). Den registrerade ska informeras om användningen av algoritmiska system, bland annat den personuppgiftsansvariges identitet, syftet med behandlingen och förekomsten av automatiserat beslutsfattande och profilering (se steg [2.1.1 Uppgifter från den klagande](#)).

Denna information återfinns ofta i användarvillkor eller integritetspolicier. Med anledning av AI-förordningens publicerade EU:s AI-kontor (AI Office) i juli 2026 riktlinjerna om transparenskraven.²⁰

Frågor att ställa till den klagande och/eller beakta vid dokumentgranskningen

Observera: Den klagande vet inte alltid vilken information som är relevant eller ens att den finns. En bra utgångspunkt är därför att fråga: "Vilka webbplatser eller appar använde du?" och "Fick du några dokument eller handlingar?"

Relevanta handlingar kan bland annat omfatta: inter alia, den ansökan som den klagande lämnade in, brev, e-postmeddelanden, fakturor, samtyckesblanketter, integritetspolicier eller integritetsmeddelanden från den svarande organisationen, avtal mellan den klagande och organisationen, inklusive bilagor och senare avtalsändringar som skickats per post, e-post eller via en app eller portal. Om sådan dokumentation finns bör den samlas in och granskas under avsnitt [1.2 – Vidare dokumentgranskning](#), eftersom den kan ge svar på flera av nedanstående frågor.

- ▶ Har den klagande informerats om att ett AI-system används eller har använts?
 - Om ärendet gäller arbetslivet: har den klagande kontrollerat detta med arbetstagarnas representanter eller den fackliga organisationen?
 - Om ja – vilken information lämnades?
- ▶ Har den klagande informerats om att automatiserad behandling av personuppgifter används?²¹
- ▶ Har den klagande ombetts att lämna samtycke till den aktuella personuppgiftsbehandlingen eller att bekräfta att en integritetspolicy har lästs?
 - Om ja – kan texten i samtycket eller integritetspolicyn delas?
 - Om nej – förekom någon annan information om "hur vi behandlar personuppgifter" eller "hur vi använder personuppgifter?"
- ▶ Förekom hänvisningar till automatisering, ett datorsystem som utför uppgifter självständigt eller ett datorsystem som lämnar underlag till ett beslut? Var särskilt uppmärksam på formuleringar som:
 - Ett system eller en dator beslutar, analyserar, jämför, rekommenderar eller behandlar personuppgifter.
 - Ett system, en dator eller en kamera "ser", "lyssnar" eller "förstår".
 - Någon beräkning av poäng, sannolikheter, risknivåer eller uppskattningar av framtida beteenden.
 - Användning av ett datorsystem som stöd i ett arbetsflöde, ibland angivet med ett produkt- eller varumärkesnamn, exempelvis: "Vi använder Clippy/Frank/Maya för att behandla dina uppgifter och stödja vårt arbete."

20. Läs mer här: <https://digital-strategy.ec.europa.eu/en/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems>

21. När personuppgifter används för helt automatiserat beslutsfattande ska användarna ha informerats om att de är föremål för helt automatiserat beslutsfattande (allmän dataskyddsförordning, artikel 13.2 f).

- ▶ Innehåller avtalet mellan den klagande och den svarande organisationen någon hänvisning till: automatisering, algoritmiska processer, namngivna system, statistik, prestationsmätt, betyg eller poängsystem (t.ex. automatiserad avaktivering eller tillförlitlighetspoäng)?
 - Om ja – kan avtalet eller de relevanta avtalsbestämmelserna delas?
- ▶ Var skedde kontakten eller kommunikationen med den svarande organisationen?
- ▶ Om kontakten skedde personligen eller per telefon:
 - Om mötet ägde rum ansikte mot ansikte, använde den andra personen en dator för att fatta beslut eller dra slutsatser?
- ▶ Om kontakten skedde via dator eller mobiltelefon:
 - Fylldes något webbformulär i?
 - Hur lång tid tog det innan ett beslut eller resultat kom (omedelbart, några minuter, några dagar eller längre)?
 - Ombads den klagande lämna samtycke till behandling av personuppgifter, exempelvis genom en kryssruta?
 - Efterfrågades någon uppgift om en skyddad egenskap (se [Skyddade egenskaper](#)) i systemet?

1.1.3. Snabb granskning av offentlig tillgänglig information

EU:s databas för högrisk-AI-system

Från och med 2 augusti 2026 förväntas en EU-databas över högrisk-AI-system²² vara tillgänglig. Databasen ska innehålla högrisk-AI-system innan de släpps ut på marknaden eller tas i bruk. Enligt Digital Omnibus om AI kommer AI-system som tagits i bruk före 2 december 2027 för fristående högrisk-AI-system respektive 2 augusti 2028 för högrisk-AI-system inbyggda i produkter endast att behöva registreras om de genomgår en "väsentlig förändring".²³ De delar av databasen som är offentliga kommer att vara tillgängliga för alla.

Vid klagomål som berör system inom utbildning, sysselsättning, offentlig förvaltning, bank och försäkring, rättskipning, val, eller som avser identifiering av personer, ansikts- eller röstigenkänning, känslöigenkänning, förutsägelse av avsikter eller risker, kan en snabb sökning i databasen – med relevanta sökord såsom den svarande organisationens namn, systemets namn eller den aktuella sektorn – vara en lämplig utgångspunkt.

22. Även om denna databas ännu inte finns och det är osäkert hur databasens utformning och gränssnitt kommer att se ut, tas den med här eftersom den förväntas bli en mycket viktig informationskälla i framtiden.

23. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026). I övriga fall där systemet med hög risk är avsett att användas av offentliga myndigheter måste nödvändiga åtgärder vidtas för att uppfylla kraven i förordningen senast i augusti 2030, AI-förordningen artikel 111.2.

- Om den svarande organisationen finns registrerad i databasen i en roll som är relevant för det aktuella användningsområdet finns det en stark indikation på att ett (högrisk-) AI-system används.²⁴ Fortsätt till [steg 1.3](#).

Observera: Det finns undantag från registreringskyldigheten och det kan även förekomma att en leverantör eller användare inte har uppfyllt sina lagstadgade skyldigheter. Mer information om databasen finns i [steg 2.2.2 – Offentligt tillgänglig information](#).

Jämför med förteckningen över användningsfall

- [Bilaga C](#) innehåller exempel på vanliga användningsområden för algoritmiska system och AI inom olika sektorer. Motsvarar något av dessa exempel den situation, det ärende eller den process som klagomålet gäller? Se även [Bilaga F](#), som innehåller databaser och register över AI-system.

Allmän internetsökning

- Genomför en internetsökning efter information om innovation eller automatisering inom den aktuella verksamheten eller hos den svarande organisationen.
 - Till exempel: "Använder den [nationella] polisen AI?"
 - Som ett exempel, för Finland ger en sådan sökning bland annat träffen: [Finland planerar en omfattande utvidgning av polisens befogenheter att samla in biometriska uppgifter – ID Tech](#).
 - Ett annat exempel: "Används AI vid rekrytering i [land]?"
 - För Belgien ger en sådan sökning resultatet att ungefär en tredjedel av de belgiska företagen använder artificiell intelligens i sina rekryteringsprocesser. www.aneuws.com.tr/europe/2023/07/19/in-belgium-one-third-of-companies-use-artificial-intelligence-in-their-recruitment-processes.
 - Var noga med att skilja mellan vanliga sökträffar och svar som genererats av AI.

Observera: Vissa AI-metoder är förbjudna enligt artikel 5 i AI-förordningen. Se [Bilaga J – Förbjudna AI-metoder](#).

Om svaren på dessa frågor talar för att ett AI-system eller ett algoritmiskt system sannolikt har använts bör handläggningen fortsätta till steg 1.3.

Om informationen fortfarande är otillräcklig bör ytterligare (offentlig) information samlas in genom den dokumentgranskning som beskrivs i [steg 1.2](#).

B. Initiativärenden (*ex officio*)

Vissa likabehandlingsorgan har enligt nationell lagstiftning befogenhet att inleda utredningar på eget initiativ *ex officio* eller *actio popularis* (genom nationell lagstiftning).

24. Observera att om ett system inte finns registrerat i denna databas betyder det inte att det inte förekommer något (högrisk-) AI-system; undantag förekommer

I detta skede utgår metodiken från att ingen enskild klagande eller grupp av klagande är involverad. I ett senare skede av utredningen kan likabehandlingsorganet välja att involvera personer som direkt berörs. För denna metodik förutsätts det att det redan finns ett definierat intresseområde, (en viss sektor, ett verksamhetsområde inom sektorn och särskilda organisationer) som redan har fastställts, till exempel prissättningsalgoritmer som används av försäkringsbolag eller automatiserad bedömning av bedrägeririsk vid sociala förmåner.

Information som bör samlas in

- ▶ den situation, det ärende eller den process som ska granskas samt tillgänglig offentlig information;
- ▶ relevanta, (ännu inte avslutade) klagomål hos EB:et eller andra myndigheter; det kan exempelvis röra sig om ärenden som enskilt varit för svaga eller där den klagande återkallat sitt klagomål;
- ▶ indikationer på att AI eller algoritmiska system används, om sådana finns se avsnittet [Hur algoritmiska system och AI kan identifieras ovan](#));
- ▶ uppgifter om den organisation som skulle bli föremål för en utredning, ex officio, (nedan kallad den "svarande") eller, vid fallet av en sektorsgranskning, en förteckning över organisationer som är representativa för den aktuella sektorn.²⁵

Jämför med befintliga användningsfall

- ▶ [Bilaga C](#) innehåller exempel på vanliga användningsområden för algoritmiska system och AI inom olika sektorer. Motsvarar något av dessa exempel den situation, det ärende eller den process som den planerade utredningen gäller?

När ett ärende inleds ex officio kan det redan vara känt att ett AI-system används. I sådana fall är steg 1.2 och steg 1.3 normalt inte nödvändiga och handläggningen kan gå direkt vidare till [steg 1.4](#).

C. Information från en marknadskontrollmyndighet (MSA)

Likabehandlingsorganet kan även få information från en MSA om AI-system som misstänks medföra risk för diskriminering. Detta gäller särskilt i följande situationer:

- ▶ **Allvarliga incidenter (AI-förordningen artikel 73.7)**
 - När en leverantör rapporterar en allvarlig incident till MSA,²⁶ ska denna informera de myndigheter som omfattas av artikel 77, däribland

25. Om en EB till exempel inleder en utredning på eget initiativ av de prissättningsalgoritmer som används av försäkringsbolag, skulle man i detta skede av förfarandet kunna använda en representativ förteckning över försäkringsbolag i stället för en enskild svarande.

26. Med allvarlig incident avses en incident eller ett funktionsfel i ett AI-system som direkt eller indirekt leder till något av följande:
(a) en persons död eller allvarlig skada på en persons hälsa;
(b) en allvarlig och oåterkallelig störning av förvaltningen eller driften av kritisk infrastruktur;
(c) åsidosättande av skyldigheter enligt unionsrätten som syftar till att skydda de grundläggande rättigheterna; (d) allvarlig skada på egendom eller miljön (AI-förordningen, artikel 3.49).

likabehandlingsorganet när det är tillämpligt. Europeiska kommissionen avslutade samrådet om vägledning och mall för incidentrapportering den 7 november 2025.²⁷

► **Produkter som medför en risk (AI-förordningen, artikel 79 jämförd med artikel 3.19 i förordning (EU) 2019/1020)**

- Where the MSA has sufficient reason to believe that an AI system qualifies as a product presenting a risk to the right of non-discrimination, it has a duty to inform and fully co-operate with the relevant Article 77 authorities, including the equality body.

This information may include the following.

- Organisation: public or private body.
- AI system: identification of the system and area of high-risk classification (Annex I or Annex III of the AI Act, including domain and use case).
- Role of the organisation (deployer, provider, other).
- When notification is based on Article 73 of the AI Act (serious incident):
 - the circumstances of the incident or malfunction;
 - the type of harm that occurred or was suspected;
 - the relationship between the AI system and the incident (how the AI caused or contributed to the incident);
 - any available information on the provider's investigation risk assessment of the incident and corrective actions proposed or implemented.
- When notification is based on Article 79 of the AI Act (risk):
 - the reason for considering that the AI system presents a risk to fundamental rights;
 - (optional, if available) relevant information on already conducted evaluations;
 - (optional, if available) corrective actions proposed or implemented.
- In cases involving a high-risk AI system in the areas of law enforcement, migration, asylum and border control management (points 1, 6 and 7 of Annex III), the MSA will be able to access information on the AI system as recorded in the EU database for high-risk AI systems.

When an equality body is informed by an MSA, there is certainty that a high-risk AI system is involved. Steps 1.2 and 1.3. are not relevant in this case: skip to [step 1.4.](#)

27. Läs mer på: AI-förordningen: Kommissionen offentliggör ett utkast till riktlinjer och en rapporteringsmall för allvarliga AI-incidenter och efterfrågar synpunkter från berörda parter.

1.2. Vidare dokumentgranskning

Om resurser finns tillgängliga för dokumentgranskning finns det flera informationskällor som kan användas för att avgöra om automatiserat beslutsfattande (ADM) eller AI sannolikt används i det aktuella ärendet. I detta avsnitt beskrivs dessa informationskällor och vilken typ av information som bör eftersökas.

De steg som beskrivs i detta avsnitt varierar vad gäller hur ingående de är och hur mycket tid som krävs för att genomföra denna dokumentgranskning:

- ▲ Anger moment som bör prioriteras;
- ▼ Anger moment som kan användas för en mer fördjupad dokumentgranskning.

Det bör till exempel vara relativt enkelt att söka i EU:s databas (nedan) när den väl är i drift och förutsatt att den svarande är registrerad i databasen. Detta ger ett snabbt svar på frågan "används ett AI-system?" Om den som använder systemet inte är registrerad krävs ytterligare efterforskningar. Om organisationen inte finns registrerad behöver ytterligare dokumentgranskning genomföras.

Detta avsnitt gäller inte fall där källan är "C. Uppgifter från MSA." I sådana fall har MSA redan fastställt att det rör sig om ett AI-system med hög risk och därför skulle den dokumentstudie som beskrivs i detta avsnitt vara överflödig.

1.2.1. EU-databas för AI-system med hög risk

Sök upp den svarande eller någon känd driftsansvarig eller leverantör i databasen. Finns den svarande upptagen som driftsansvarig i del C av databasen?²⁸

- ▲ Är den svarande driftsansvarig för ett användningsfall som är relevant för klagomålet? Vilken organisation är leverantör?

Relevant information som finns tillgänglig i databasen för en första identifiering:

- ▶ kontaktuppgifter till leverantören och den behöriga företrädaren;
- ▶ namn och kontaktuppgifter till den som tar systemet i bruk (endast offentliga myndigheter, EU-institutioner, organ och byråer, eller personer som agerar på deras vägnar);
- ▶ URL-adressen till AI-systemet som leverantören har lagt in i EU-databasen.

Finns den svarande registrerad i databasen som leverantör i del A eller B av databasen?

- ▲ Är den svarande leverantör för ett användningsfall med hög risk som är relevant för klagomålet?
- ▲ Är den svarande leverantör för ett användningsfall som är avsett för användning inom ett högriskområde, men som enligt egen bedömning omfattas av undantaget i artikel 6.3 i AI-förordningen? Denna bedömning kan begäras av de nationella behöriga myndigheterna.

28. Observera att endast offentliga myndigheter är skyldiga att registrera sig som användare. Övriga organisationer har ingen sådan skyldighet.

Relevant information som finns tillgänglig i databasen för en första identifiering:

- ▶ kontaktuppgifter till leverantören och den behöriga representanten;
- ▶ handelsnamn eller annan referens som möjliggör identifiering och spårning av det AI-systemet (detta kan till exempel vara användbart när ett högrisk-system används som en del av ett annat system);
- ▶ avsett användningsändamål.

Om leverantören/den svarande finns i denna databas innebär det att ett AI-system används. Gå vidare till steg 1.3 och/eller 1.4.

Om inte, sök efter organisationer i databasen som liknar den svarande:

- ▼ Finns det liknande organisationer i databasen som leverantörer eller användare av ett användningsfall som är relevant för klagomålet? Till exempel kan andra liknande försäkringsbolag ange AI-system för prissättning när ärendet gäller (o)rättvis prissättning, eller andra liknande myndigheter i samma land kan ha angett AI-system för bedömning av förmåner när ärendet gäller nekade förmåner.

Enligt artikel 49.4 ska registreringen av högrisk-AI-system inom områdena brottsbekämpning, migration, asyl och gränskontroll ske i en säker, icke-offentlig del av den EU-databas som avses i artikel 71 och bör endast omfatta en specifik uppsättning uppgifter enligt vad som anges i artikel 49.4. Åtkomsten till dessa delar av databasen kommer att vara begränsad.

För mer information om den icke-offentliga delen av databasen och öppenhetsaspekter inom området brottsbekämpning, se begränsningar av tillgången till information under punkt 2.2.2 respektive rutan "Området brottsbekämpning" under punkt 2.2.4.

1.2.2. Andra databaser och inventarier för (högrisk-) AI-system och algoritmiska system

Vissa organisationer och kommuner för en lokal databas, ett register eller ett inventarium över AI-system som används. En förteckning över dessa databaser finns i [bilaga F](#). Alla dessa databaser är inte relevanta. Kontrollera endast de som är relevanta för det nationella sammanhanget.

Sök efter den svarande i databasen:

- ▲ Finns ett algoritmiskt system eller AI, registrerat som används eller utvecklats av den svarande, och som är relevant för den situation, det ärende eller den process som klagomålet avser?

Sök efter liknande organisationer till den svarande:

- ▼ Finns ett algoritmiskt system eller AI registrerat som används eller utvecklats av en jämförbar organisation, som skulle ha varit relevant för den situation,

det ärende eller den process som klagomålet avser om klagomålet gällt denna jämförbara organisation?

Jämför med förteckningen över befintliga användningsfall:

- ▼ **Bilaga C** innehåller en förteckning över vanliga användningsområden för algoritmiska system och AI inom olika sektorer.

Är något av dessa användningsfall relevanta för den situation, det ärende eller den process som klagomålet avser?

1.2.3. Den svarandes webbplats och externa kommunikation

Integritetsmeddelande/integritetspolicy

Sök efter formuleringar som indikerar användning av algoritmiska system, till exempel något av följande:

- ▲ Syften som tyder på automatisering, profilering eller statistiska metoder. Se avsnittet [Hur man identifierar algoritmiska system och AI](#) samt frågorna i [1.1.2 Frågor att ställa till klaganden och/eller beakta vid dokumentgranskning](#).
- ▲ Tredje parter som tar emot personuppgifter (mottagare eller kategorier av mottagare). Kontrollera indikationer på tredje parts-inblandning, såsom scoringföretag eller kreditupplysningsföretag, vilket kan tyda på profilering eller automatiserade beslut.
- ▼ Kontaktuppgifter till personuppgiftsansvarig: är den personuppgiftsansvarige känd för att använda AI-lösningar eller automatisering inom den aktuella sektorn? Om ett AI- eller algoritmiskt system finns är den personuppgiftsansvarige normalt sett systemets användare och den svarande enligt denna metodik.

Övriga typer av kommunikation

Finns det någon hänvisning till ett algoritmiskt system som kan vara relevant för ärendet?

- ▼ Sök på den svarandes webbplats efter artiklar om implementering av AI, innovation och automatisering eller vanliga frågor (FAQ) på webbplatsen/plattformen. Det enklaste sättet att söka en specifik webbplats är att lägga till "site: [den svarandes webbplats]" i sökfrågan. Exempel: innovation AI site: bigbank.com..
- ▼ Kontrollera platsannonser för roller som datavetare, AI eller maskininlärning. Finns det någon hänvisning till ett algoritmiskt system som kan vara relevant för ärendet? En platsannons för en "senior datavetare" vid en bank innehåller följande: "Förebyggande av penningtvätt och finansiering av terrorism är en av de centrala aspekterna för vårt samhälle och vår organisation ... du kommer att ansvara för att optimera vårt nuvarande regelbaserade system för transaktionsövervakning (TM) och på sikt införa mer avancerade analysmetoder såsom maskininlärning i våra TM-modeller." Se [senior data scientist transaction monitoring – Rabobank](#).

Om tillämpligt: granska webbplatsen och kommunikationen från användare av relevanta AI-system med hög risk.

Om användningsfall som är relevanta för ärendet anges av tredje parts-användare i EU-databasen för AI-system med hög risk, granska även leverantörens webbplats.

- ▼ Finns den svarande listad på webbplatsen som en "organisation vi arbetar med" eller "betrodda kunder"?
- ▼ Nämns den svarande på något annat sätt på leverantörens webbplats? Detta kan enklast identifieras genom en internetsökning, till exempel: [svarande] site: [leverantör].com.

1.2.4. Ytterligare genomgång av offentlig information

Vissa av stegen som anges här har redan beskrivits i steg 1.1. För tydligheters skull upprepas de här.

Allmän internetsökning

- ▼ Se internetsökning enligt steg 1.1.3.
- ▼ Förfina sökningen i steg 1.1.3 genom att särskilt kontrollera om icke-statliga organisationer (NGO:er), nationella eller internationella organ eller andra organisationer har rapporterat om användning av AI eller algoritmiska system i en liknande situation, ett liknande ärende eller en liknande process som klagomålet avser, antingen av den svarande eller av liknande organisationer. Relevanta icke-statliga organisationer inkluderar Amnesty, AlgorithmWatch och nationella civilsamhällesorganisationer. I Nederländerna publicerar dataskyddsmyndigheten en rapport om AI- och algoritmrisiker två gånger per år. En sökning i den nederländska nationella kontexten kan vara: "[sektor för klagomål] Autoriteit Persoonsgegevens AI and algorithms risk report".
- ▼ Valfritt: Specificera sökningen och inkludera nationella och internationella AI-strategier.
- ▼ Andra relevanta källor kan vara [OECD:s monitor för AI-incidenter och faror](#) samt [databasen för AI-incidenter](#) eller rapporter från Europarådet såsom "Discrimination, artificial intelligence and algorithmic decision making" (Zuiderveen Borgesius 2018).

1.3. Identifiering av algoritmiska system eller AI-system med hög risk

Baserat på den information som samlats in i steg 1.1 och 1.2 kan en rimlig bedömning göras av huruvida ett algoritmiskt system eller ett AI-system med hög risk används, samt hur säker denna bedömning är.

Påminnelse: denna metodik skiljer endast mellan AI-system med hög risk och algoritmiska system. Algoritmiska system omfattar AI-system som inte är högrisk samt ADM-system. För AI-system med hög risk ställer AI-förordningen omfattande transparenskrav och dokumentationskrav, vilket innebär att ytterligare dokumentation förväntas vara tillgänglig. Om ett AI-system inte är ett högrisk-AI-system är

den information som finns tillgänglig och de steg som ska följas desamma som för enklare algoritmiska system och dessa slås därför samman i en gemensam kategori.

Typ av system

- ▶ **AI-system med hög risk**
- ▶ **Algoritmiskt system**

Grad av säkerhet

- ▶ **Relativt säkert**
- ▶ **Möjligt**

Observera: Att dra slutsatser om huruvida AI eller algoritmiska system används är en svår och nyanserad bedömning i detta skede, där mycket begränsad information finns tillgänglig. De säkerhetsnivåer som anges nedan är vägledande och avsedda enbart som stöd för fortsatt riktning.

EU-databas för AI-system med hög risk

Är systemet registrerat i EU-databasen?

- ▶ JA, i avsnitt A eller C: ett AI-system med hög risk har identifierats i detta ärende. Det är **relativt säkert** att ett **AI-system med hög risk** används i detta ärende.
- ▶ JA, i avsnitt B:²⁹ ett AI-system används, men ingen av den dokumentation som krävs för högrisk-system förväntas vara tillgänglig. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.
- ▶ NEJ, men det finns liknande organisationer (t.ex. en annan bank) som har registrerat system som är relevanta för ärendet (t.ex. system av kreditscoring) i avsnitt A eller C. Det kan därför vara rimligt att anta att ett liknande högrisk-system används av den svarande. I detta skede av metodiken antas det att det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende. Det är **möjligt** att ett **AI-system med hög risk** används i detta ärende.

Information från den svarande (antingen till klaganden, i offentliga dokument eller på begäran)

- ▶ Den svarande har uppgett användning av automatisering, algoritmer eller AI för ändamål som är relevanta för ärendet, antingen direkt till klaganden eller i integritetspolicy/integritetsmeddelande. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.
- ▶ Den svarande har uppgett användning av automatisering, algoritmer eller AI för ändamål som är relevanta för kategorier av högrisk-AI (se bilaga I), antingen direkt till klaganden eller i integritetspolicy/integritetsmeddelande. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende; det är **möjligt** att ett **AI-system med hög risk** används i detta ärende.
- ▶ Den svarande har registrerat ett algoritmiskt system eller AI-system i en offentlig databas (se bilaga F) för ändamål som är relevanta för ärendet. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.

29. Enligt artikel 6.3 i AI-förordningen anses AI-systemet inte utgöra någon hög risk.

- Den svarande har offentligt uppggett användning av algoritmiska system eller AI för ändamål som är relevanta för ärendet i blogginlägg, pressmeddelanden eller andra källor. Observera att hänvisningar till användning av algoritmiska system kan vara vaga eller undvikande (se [steg 1.1](#) för möjliga formuleringar). Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.

Övrig information

- Relevanta och trovärdiga aktörer såsom icke-statliga organisationer och etablerade medieorganisationer har rapporterat om användningen av algoritmiska system eller AI av den svarande för ändamål som är relevanta för ärendet. Ingen bekräftelse från organisationen har kunnat återfinnas. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.
- Andra aktörer har rapporterat användning av algoritmiska system eller AI av den svarande för ändamål som är relevanta för ärendet. Ingen bekräftelse från den svarande har kunnat återfinnas. Det är **möjligt** att ett **algoritmiskt system** används i detta ärende.
- AI och algoritmiska system används ofta för ändamål som är relevanta för ärendet, per [bilaga C](#). Det är **möjligt** att ett **algoritmiskt system** används i detta ärende.
- Liknande organisationer rapporteras använda algoritmiska system för ändamål som är relevanta för ärendet. Det är **möjligt** att ett **algoritmiskt system** används i detta ärende.
- Det finns kontextuella indikatorer som tyder på att algoritmiska system kan vara inblandade i ärendet. Till exempel på grund av beslutsfattandets hastighet, när komplexa processer eller bedömningar som normalt kräver noggrant omdöme hanteras omedelbart eller inom några minuter. Det är **möjligt** att ett **algoritmiskt system** används i detta ärende.
- Det finns kontextuella indikatorer som tyder på att algoritmiska system kan vara inblandade i ärendet. Till exempel på grund av beslutsfattandets hastighet, när komplexa processer eller bedömningar som normalt kräver noggrant omdöme hanteras omedelbart eller inom några minuter och där antingen AI och algoritmiska system ofta används för ändamål som är relevanta för ärendet enligt bilaga C, eller där liknande organisationer rapporteras använda algoritmiska system för ändamål som är relevanta för ärendet. Det är **relativt säkert** att ett **algoritmiskt system** används i detta ärende.

1.4. Fortsatt handläggning av ärendet

Steg 1.1 till 1.3 omfattar bedömningen av huruvida ett algoritmiskt system sannolikt är involverat. För att besluta om ärendet ska drivas vidare bör en ytterligare bedömning göras av sannolikheten för att ett starkt diskrimineringsärende föreligger. Detta bör bedömas mot etablerade kriterier inom likabehandlingsorganet EB:et. När ett algoritmiskt system potentiellt är involverat kan detta kräva att en lägre tröskel för bevisstyrka tillämpas innan beslut fattas om att gå vidare med ärendet.

Manjang v. Uber Eats UK Ltd & Others (2021)

Detta fall rör diskriminering på grund av ras och väcktes med hänvisning till trakasserier, repressalier samt indirekt diskriminering efter införandet av teknik för ansiktsgenkänning. Målsägande, en leveransförelse, mottog ett e-postmeddelande från Uber Eats (den svarande) där han informerades om att han permanent stängts av från plattformen med hänvisning till att han hade "delat sitt Uber Eats-konto som leveranspartner" och att han hade underkänts i en kontroll av ansiktsgenkänning. Han svarade att "er algoritm verkar vara rasistisk och detta behöver åtgärdas eftersom den inte kan känna igen och verifiera mina foton, vilket sannolikt är anledningen till att jag ombuds ta bilder av mig själv flera gånger om dagen". Uber Eats bestred diskrimineringsanspråket och begärde att det skulle avvisas i ett tidigt skede eftersom det, enligt dem, byggde på "sakfel". De lämnade därefter in uppgifter som grund för sitt bestridande om hur tekniken fungerade, och målsägande ändrade därefter sin talan. Domstolen avslog begäran om avvisning av den ändrade talan och konstaterade att det vid tidpunkten för talans utformning rådde oklarhet om varför kontot hade inaktiverats. Några månader senare under processen uppgav svaranden att den ursprungliga informationen om att beslutet var automatiserat var felaktig och att en manuell granskning hade gjorts. Målsägande vidhöll, och domstolen instämde i, att det rådde oklarhet kring orsaken till att bilderna underkänts och kontot inaktiverats. Slutligen nåddes en ekonomisk uppgörelse utanför domstol.

Detta är ett exempel på ett ärende som finansierats och stöds av ett EB, i detta fall Kommissionen för jämställdhet och mänskliga rättigheter (Storbritannien).

Tabell 2. Säkerhetsgrad för högrisk- eller algoritmiska system och sannolikhet för diskriminering

	Låg sannolikhet för diskriminering	Medelstor risk för diskriminering	Stor risk för diskriminering
Relativt säkert AI med hög risk	Ärende med hög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system	Ärende med hög prioritet – överväg att involvera en (intern) specialist på algoritmiska system	Ärende med hög prioritet – överväg att involvera en (intern) specialist på algoritmiska system
Relativt säkert algoritmiskt system, potentiellt AI med hög risk	Ärende med medelhög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system	Ärende med hög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system	Ärende med hög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system
Relativt säkert algoritmiskt system	Ärende med låg prioritet; vid vidare handläggning bör man överväga att involvera en (intern) specialist på algoritmiska system	Ärende med medelhög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system	Ärende med hög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system
Potentiellt algoritmiskt system	Använd befintliga metoder i stället för den här metoden	Ett fall med låg prioritet enligt denna metod, där sannolikheten för diskriminering är medelhög, men sannolikheten för att det beror på AI-systemet är låg; hänvisa till befintliga metoder istället för denna metod	Ärende med medelhög prioritet – överväg att koppla in en (intern) specialist på algoritmiska system

Misstänkt felaktig klassificering enligt AI-förordningen

Felaktig klassificering som ett system som inte är högrisk

Det finns en möjlighet att leverantörer av AI-system inom de områden som anges i bilaga III undviker regelverket för AI-system med hög risk genom att göra en egen bedömning enligt något av följande.

- Inte ett AI-system. I så fall är AI-förordningen inte tillämplig.
- Inte ett AI-system med hög risk på grund av undantagen (AI-förordningen, artikel 6.3). I så fall ska systemet ändå registreras i avsnitt B i EU-databasen.

Leverantören är då skyldig att dokumentera sin bedömning som ligger till grund för att systemet undantas från reglerna innan det släpps ut på marknaden eller tas i bruk samt registrera systemet i avsnitt B i EU-databasen (AI-förordningen, artikel 49.2). Det finns inget krav på att offentliggöra denna bedömning, men leverantören ska på begäran lämna dokumentationen till de nationella behöriga myndigheterna (MSA eller det anmälda organet).

- Inte ett AI-system med hög risk eftersom ändamålet inte omfattas av beskrivningen av de användningsfall som anges i bilaga III till AI-förordningen (AI-förordningen, artikel 6.3).

Om ett likabehandlingsorgan misstänker att ett system felaktigt har bedömts falla utanför regelverket för AI-system med hög risk bör det underrätta den berörda MSA, oavsett om ett diskrimineringsärende kan styrkas eller inte (se steg 4, MSA förfarande för hantering av AI-system som av leverantören klassificerats som system som inte är högrisk, samt steg 4.3).

Vid tillämpningen av denna metodik bör systemet betraktas som ett algoritmiskt system, eftersom den information som normalt förväntas finnas tillgänglig för AI-system med hög risk sannolikt inte kommer att vara tillgänglig i detta fall.

Frågor att ställa

- Omfattas systemet av de användningsfall som anges i bilaga III?
- Är det sannolikt att systemet är ett AI-system? Se avsnittet [Hur algoritmiska system och AI kan identifieras](#).
- Omfattas det av förteckningen över praktiska exempel? (Se [bilaga C](#) och/eller den förteckning som kommissionen ska ta fram när den finns tillgänglig, artikel 6.4.)
- Kontrollera avsnitt B i databasen. Leverantören av systemet är enligt artikel 49 skyldig att registrera både sig själv och systemet.

Misstanke om förbjudet system

Vissa AI-metoder är förbjudna enligt AI-förordningen (se [bilaga J – Förbjudna AI-metoder](#)).

Om EB:et identifierar ett förbjudet system eller en förbjuden metod:

- eftersom dessa metoder är förbjudna innebär användningen i sig en överträdelse av EU-rätten. Omedelbara åtgärder bör vidtas för att underrätta den berörda MSA:en, oavsett om ett diskrimineringsärende kan styrkas eller inte;
- om EB:et dessutom drar slutsatsen att en metod som för närvarande inte omfattas av denna förteckning ändå kränker grundläggande rättigheter, bör det sammanställa bevisningen och överlämna rapporter till de behöriga myndigheterna i den berörda medlemsstaten samt till kommissionen (se [steg 4.3](#)).

Ärenden som rör förbjudna metoder kan även avse AI-system med hög risk och/eller ha orsakat diskriminering. EB:et behöver därför ta ställning till om ytterligare åtgärder med anledning av diskriminering ska vidtas.



Steg 2

Informationsinhämtning om det algoritmiska systemet eller AI-systemet för att identifiera diskriminering

I steg 2 identifieras vilka typer av information som kan finnas tillgänglig eller vara åtkomlig, både vad gäller systemet och sannolikheten för att diskriminering föreligger. Här ges vägledning om möjliga källor och hur informationen samlas in.

Steg 2 består av följande delsteg.

Figur 3. Steg 2 – Informationsinsamling för att identifiera diskriminering

STEG 2		Tillgänglig information om algoritmiska system	
2.1		Tillgänglig information om algoritmiska system	
2.1.1 Information från klagande ▶ Intervjuer ▶ Dokument från klagande ▶ Begäran om tillgång till personuppgifter (DSAR)	2.1.2 Allmänt tillgänglig information ▶ Extern kommunikation från svarande ▶ Befintlig rättspraxis ▶ Övriga källor	2.1.3 Information från svaranden eller andra parter ▶ Information som begärts från svaranden – Konsekvensbedömning avseende dataskydd (DPIA) – Konsekvensbedömning avseende mänskliga rättigheter eller grundläggande rättigheter (HRIA/FRIA) – Jämlikhets- eller genuskonsekvensbedömning – Dokumentation enligt AI-förordningen ▶ Statistisk information eller uppgifter från undersökningar ▶ Övrig dokumentation om produktsäkerhet	
2.2		Ytterligare information om högrisk-AI	
2.2.1 Ytterligare information från den klagande ▶ Begäran enligt artikel 86: rätt till förklaring	2.2.2 Offentligt tillgänglig information ▶ Databas över AI-system med hög risk enligt AI-förordningen	2.2.3 Ytterligare information från svaranden eller andra parter ▶ Dokumentation enligt AI-förordningen ▶ FRIA ▶ DPIA	2.2.4 Anmälan enligt MSA eller utredning på eget initiativ (<i>ex officio</i>)

Om ett EB inte är en myndighet enligt artikel 77 och/eller systemet inte har identifierats som ett hög risk-system, kan steg 2.2 (till exempel bedömning av konsekvenser för likabehandling samt statistisk information eller enkätdata) ändå användas som vägledande ram för att göra det möjligt för EB:et att begära motsvarande information från användaren eller leverantören med stöd av sina befogenheter att utreda och inhämta information samt att bedöma eventuella svar eller uteblivna svar enligt [steg 3](#). Observera dock att den svarande i dessa situationer inte har någon rättslig skyldighet enligt AI-förordningen att lämna sådan dokumentation (även om en sådan skyldighet kan följa av annan lagstiftning, exempelvis om produktsäkerhet eller konsumentskydd).

Analysen av den bevisning som samlas in enligt steg 2 genomförs i steg 3. I praktiken utförs dessa steg vanligtvis parallellt.

Vid steg 2 har man med en viss grad av säkerhet identifierat att ett algoritmiskt system eller ett AI-system är involverat i det aktuella ärendet och eventuellt även att systemet är klassificerat som ett AI-system med hög risk.³⁰ Nästa steg är att inhämta information om systemet som kan ligga till grund för den efterföljande bedömningen av diskriminering i [steg 3](#).

Det finns tre kategorier av informationskällor.

1. Information som kan lämnas till EB:et av klaganden. Sådan information kan vara tillgänglig för klaganden genom deras egna erfarenheter av det ifrågasatta förfarandet, genom korrespondens med den svarande samt genom utövandet av sina rättigheter till information, framför allt enligt GDPR och AI-förordningen.
2. Annan information kan vara offentligt tillgänglig.
3. Information som EB:et kan begära från den svarande med stöd av sina nationella eller unionsrättsliga befogenheter att utreda och inhämta information, inklusive de befogenheter som följer av EU:s AI-förordning.

Kom ihåg att information också kan lämnas ut eller begäras under korrespondens före en rättslig prövning eller under en domstolsprocess.

Observera: steg 2 gör inte anspråk på att vara uttömmande. Det kan finnas andra typer av information eller andra informationskällor. Nationella rättsliga ramar, däribland lagstiftning om handlingsoffentlighet eller regler om insyn i handlingar, kan ge EB:er ytterligare befogenheter att inhämta information eller utreda förhållanden utöver dem som beskrivs i denna metodik. Tillgången till information kan också stärkas genom de nya utredningsbefogenheter som följer av standarddirektiven (se Inledning och [steg 4](#)).

30. Om det inte har gått att identifiera ett AI- eller algoritmiskt system ska du avbryta detta förfarande och återgå till dina vanliga rutiner.

Kända hinder för att få tillgång till information

Strategier för att inhämta information, såsom begäranden om tillgång till allmänna handlingar, kan möta olika hinder och begränsningar, särskilt invändningar om sekretess och verksamhetsmässiga krav. I Frankrike ledde exempelvis en begäran från civilsamhällesorganisationer om tillgång till information för att få klarhet i ett riskbedömningssystem som användes av den myndighet som administrerar familjeförmåner endast till att källkoden för tidigare modeller (inte den aktuella modellen) samt maskerade förteckningar över använda variabler lämnades ut (Europarådet 2025a; *La Quadrature du Net and Others v. Premier ministre and Ministère de la Culture* 2024; se även *Big Brother Watch* 2025).

I *Kelly v. National University of Ireland (C-104/10)*, där klagande ansåg sig ha blivit utsatt för könsdiskriminering efter att ha nekats tillträde till ett program för yrkesutveckling och begärde uppgifter från utbildningsinstitutionen om övriga sökandes meriter, klargjorde EU-domstolen att den sekundära unionsrätten inte ger någon individuell rätt till tillgång till sådana uppgifter. Domstolen uttalade emellertid att det ankommer på de nationella domstolarna att avgöra om ett utlämnande av dessa uppgifter, med hänsyn till omständigheterna i det enskilda fallet, är nödvändigt för att säkerställa syftet med den relevanta lagstiftningen, nämligen direktiv 97/80/EG.

EU-domstolen kan på motsvarande sätt i framtiden komma att tolka de relevanta bestämmelserna i direktiven om likabehandling så att de säkerställer ett effektivt rättsligt skydd mot diskriminering och underlättar klagandens tillgång till information redan under det skede som föregår att ett *prima facie*-fall har styrkts (Farkas och O'Farrell 2015).

Det kan vara lämpligt att föra ett överskådligt och sökbart register över de källor som har konsulterats. Detta kan bidra till att identifiera mönster av diskriminering och spara värdefull tid för handläggaren. Följande rubriker kan föreslås för litteraturöversikter:

- ▶ källa (inklusive en länk);
- ▶ nyckelord (till exempel "proxy-diskriminering", "minskning av partiskhet");
- ▶ sammanfattning;
- ▶ relevans (till exempel "diskriminering har konstaterats förekomma i systemet för ansiktsigenkänning").

2.1. Algoritmiska system

2.1.1. Information som lämnas av klaganden

Klaganden har ofta mycket relevant information som kan bidra till att identifiera potentiell diskriminering. Sådan information kan vara tillgänglig eller åtkomlig genom klagandens egna erfarenheter av det ifrågasatta förfarandet eller genom korrespondens med den svarande. Den kan även grunda sig på klagandens rätt till information, särskilt enligt GDPR (se rutan nedan).

När ett ärende handläggs inom ramen för en utredning på eget initiativ (*ex officio*) eller har identifierats av en MSA finns det ingen direkt klagande. Vid informationsinhämtningen bör följande övervägas.

- ▶ Involvera klagande från andra, äldre eller ofullständiga klagomål, eller undersök om ett klagomål redan har inlämnats till en annan myndighet eller ett annat organ.
- ▶ Om det är lämpligt, rekommendera ett gemensamt utövande av rätten till tillgång enligt GDPR i ärenden av stor betydelse.³¹

GDPR:s krav på öppenhet och registrerades rättigheter

GDPR innehåller både krav på information i förväg (*ex ante*) och rätt till tillgång i efterhand (*ex post*), vilket möjliggör för registrerade att få information om hur deras personuppgifter behandlas på ett rättvist sätt.

Ex ante

- ▶ Enligt artikel 5.1 och skäl 60 i GDPR ska personuppgifter behandlas på ett lagligt, korrekt och öppet sätt i förhållande till den registrerade. Den registrerade ska informeras om att behandlingen sker och om dess ändamål. Den personuppgiftsansvarige bör även lämna all ytterligare information som krävs för att säkerställa en rättvis och transparent behandling, med hänsyn till de särskilda omständigheter och det sammanhang där personuppgifterna behandlas.

31. Även om GDPR, AI-förordningen och LED erbjuder begränsade möjligheter till kollektiv representation har forskare hävdats att det ändå kan vara givande att utöva individuella rättigheter kollektivt (Mahieu, Asghari och van Eeten 2018). Till exempel skulle flera individer kunna lämna in begäranden om tillgång, och de insikter som detta ger skulle kunna delas med en betrodd mellanhand genom datadonationer. Sådana kollektiva insatser, genom individers insatser, skulle kunna skapa en större kunskapsbas om hur data används av specifika aktörer eller sektorer. På samma sätt skulle det kollektiva – men ändå individuella – utövandet av rätten till tillgång enligt artiklarna 15 och 22 i GDPR, eller rätten till en förklaring enligt artikel 86 i AI-förordningen, av registrerade eller berörda personer, kunna öppna dörrar till en bredare kunskap om hur AI-system fungerar. Enligt de föreslagna ändringarna av rätten till tillgång enligt GDPR i förslaget till förordningen om den digitala omnibusförordningen kan dock effektiviteten hos begäranden om tillgång, både som individuella och kollektiva verktyg för öppenhet, minskas avsevärt. Se mer nedan under [Inlämning och handtering av begäran om registerutdrag \(DSAR\) enligt artikel 15 i GDPR](#).

- ▶ Enligt artiklarna 13 och 14 i GDPR ska registrerade, när personuppgifter behandlas, få information om den personuppgiftsansvariges identitet och kontaktuppgifter, ändamålen med behandlingen, den rättsliga grunden för behandlingen, kategorierna av personuppgifter när uppgifterna har hämtats från tredje part eller offentligt tillgängliga källor samt om förekomsten av automatiserat beslutsfattande, inklusive profilering.³²

Denna *ex ante* är normalt av allmän karaktär och återfinns ofta i integritetsmeddelanden eller -policyer.

Automatiserat beslutsfattande

- ▶ Enligt artiklarna 13.2 f och 14.2 g i GDPR ska registrerade uttryckligen informeras om förekomsten av sådant enbart automatiserat beslutsfattande och sådan profilering som avses i artikel 22 i GDPR. De ska också få meningsfull information om den bakomliggande logiken samt om behandlingens betydelse och de förutsedda konsekvenserna för dem.

Eftersom enbart automatiserat beslutsfattande enligt artikel 22.2 i GDPR endast är tillåtet om beslutet:

- ▶ är nödvändigt för att ingå eller fullgöra ett avtal;
- ▶ är tillåtet enligt unionsrätten eller medlemsstaternas nationella rätt, eller;
- ▶ grundar sig på den registrerades uttryckliga samtycke, bör EB också kunna få tillgång till de relevanta avtalsbestämmelserna, den lagstiftning som tillåter behandlingen eller samtycket samt vad samtycket faktiskt omfattade, exempelvis i en undertecknad handling, ett elektroniskt formulär eller ett e-postmeddelande.

Ex post

- ▶ Enligt artikel 15 i GDPR har registrerade rätt att begära tillgång till sina personuppgifter, inklusive information om kategorierna av personuppgifter, ändamålen med behandlingen, mottagarna eller kategorierna av mottagare av personuppgifterna samt om förekomsten av automatiserat beslutsfattande, inklusive profilering.

Automatiserat beslutsfattande

- ▶ Enligt artikel 15.1 h i GDPR har registrerade, i de fall som avser enbart automatiserat beslutsfattande och profilering enligt artikel 22, även rätt att begära utförlig information om den bakomliggande logiken och om behandlingens betydelse och de förutsedda konsekvenserna.

Europeiska dataskyddsstyrelsen (EDPB) uppmuntrar ansvariga att, som god praxis, lämna meningsfull information om den bakomliggande logiken och förklara betydelsen och de förutsedda konsekvenserna av automatiserat beslutsfattande

32. En väletablerad uppfattning håller på att växa fram att detta även gäller automatiserat beslutsfattande som inte omfattas av artikel 22 i GDPR, eftersom bägge artiklarna kräver att den personuppgiftsansvarige informerar om ändamålen med behandlingen, i kombination med skäl 60 i Europaparlamentets och rådets förordning (EU) 2016/679 (Barros Vale och Zanfir-Fortuna 2022:16).

och profilering även vid icke-kvalificering (Barros Vale och Zanfir-Fortuna 2022:16).

Observera: Nationell lagstiftning kan ge ytterligare rättigheter när det gäller utlämnande av handlingar. I Sverige har exempelvis en person som har sökt ett arbete, en befordran eller en utbildningsmöjlighet men inte blivit utvald rätt att på begäran få skriftlig information från arbetsgivaren om den person som valdes ut, beträffande utbildning, arbetslivserfarenhet och andra meriter, för att kunna jämföra kvalifikationerna (svenska diskrimineringslagen, 2 kap. 4 §).

För länder som inte omfattas av GDPR bör bestämmelserna i den nationella lagstiftningen beaktas.

Viktiga rättsfall: *SCHUFA Holding (2023)* och *CK v. Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien (2025)*

(Siffror inom hakparenteser hänvisar till punkterna i avgörandet.)

Frågan om huruvida ett automatiserat system har använts och vilken information som måste lämnas till den registrerade har aktualiserats inom ramen för GDPR, där särskilda bestämmelser gäller för "beslut som enbart grundas på automatiserad behandling, inbegripen profilering" (artikel 22). Tre villkor måste vara uppfyllda för att artikel 22 ska vara tillämplig.

- ▶ Det måste finnas ett "beslut" (se skäl 71, där det bekräftas att begreppet har ett brett tillämpningsområde och där särskilt automatiskt avslag på en kreditansökan online eller e-rekryteringsförfaranden utan mänsklig medverkan anges som exempel).
- ▶ Beslutet måste "enbart grundas på automatiserad behandling (inbegripen profilering)".
- ▶ Detta beslut, som enbart grundas på automatiserad behandling, måste ha "rättsliga följder eller på liknande sätt i betydande grad påverka" den registrerade.

När dessa villkor är uppfyllda ska organisationer som använder ett system för automatiserat beslutsfattande kunna ge den registrerade en "meningsfull förklaring till hur beslutet fattades". Detta ligger mycket nära formuleringen i artikel 86 i AI-förordningen om "rätt till en förklaring".

SCHUFA Holding

I mål C-634/21 *SCHUFA Holding (Scoring)*, prövade EU-domstolen ett fall där ett privat företag tillhandahöll kreditbedömningar till sina samarbetspartner, exempelvis banker. För att bedöma en persons kreditvärdighet jämfördes personens egenskaper med uppgifter om hur andra personer med jämförbara egenskaper hade agerat. Därefter tilldelades personen ett poängvärde genom matematiska och statistiska metoder för att exempelvis ange sannolikheten för att personen skulle återbetala ett lån [14].

Klaganden, OQ, nekades ett lån av en bank (en tredje part) till följd av negativa uppgifter som SCHUFA hade överfört till banken. OQ begärde information om vilka personuppgifter som hade legat till grund för bedömningen och begärde att vissa uppgifter, som påstods vara felaktiga, skulle raderas [15].

SCHUFA svarade genom att informera OQ om dennes kreditpoäng och redogjorde i allmänna ordalag för de metoder som användes för att beräkna poängen. Med hänvisning till företagshemligheter vägrade företaget dock att lämna ut vilka faktorer som hade beaktats vid beräkningen och hur dessa hade viktats. SCHUFA uppgav slutligen att företaget endast överförde information till banken och att det var banken som fattade det faktiska beslutet. Företaget gjorde således gällande att kreditpoängen inte utgjorde ett "beslut" [16].

Som svar på de tre tolkningsfrågorna uttalade domstolen följande.

- ▶ Beslut: Begreppet beslut är tillräckligt brett för att omfatta resultatet av en beräkning av en persons kreditvärdighet i form av ett sannolikhetsvärde beträffande personens förmåga att fullgöra sina betalningsåtaganden i framtiden [47].
- ▶ "Enbart grundat på automatiserat beslutsfattande, inbegripen profilering:" Det var ostridigt att SCHUFA:s verksamhet uppfyllde definitionen av profilering och att detta villkor därför var uppfyllt.
- ▶ "Rättsliga följder eller på liknande sätt i betydande grad påverkar den registrerade:" Ett otillräckligt sannolikhetsvärde leder i nästan samtliga fall till att banken avslår den ansökta krediten. Eftersom den tredje parten i hög grad grundar sitt beslut på sannolikhetsvärdet var även detta villkor uppfyllt [48].

Domstolen slog därför fast att när "det sannolikhetsvärde som fastställs av ett kreditupplysningsföretag och lämnas till en bank spelar en avgörande roll för beviljandet av kredit, ska fastställandet av detta värde i sig anses utgöra ett beslut som medför rättsliga följder för den registrerade eller på liknande sätt i betydande grad påverkar denne, i den mening som avses i artikel 22.1 i GDPR."

Domstolen betonade också behovet av:

- ▶ att förebygga de risker för diskriminering som kan uppstå till följd av profilering;
- ▶ att använda lämpliga matematiska eller statistiska metoder för profilering;
- ▶ att införa tekniska och organisatoriska åtgärder, särskilt för att säkerställa att felaktiga personuppgifter rättas och att risken för fel minimeras;
- ▶ att säkerställa skyddet av uppgifterna.

Domstolen slog dessutom fast att medlemsstaterna inte får anta bestämmelser som tillåter profilering i strid med kraven i artiklarna 5 (principer för behandling av personuppgifter) och 6 (laglig behandling av personuppgifter) i GDPR och inte heller genom lagstiftning slutgiltigt fastställa hur avvägningen mellan de berörda rättigheterna och intressena ska utfalla [70].

Dun & Bradstreet

Förklaringar måste göra det möjligt för den registrerade att förstå beslutet och, vid behov, ifrågasätta det samt utöva sina rättigheter enligt dataskyddslagstiftningen. Enbart en redovisning av en komplex matematisk formel, såsom en algoritm, eller

en detaljerad beskrivning av samtliga steg i ett automatiserat beslutsförfarande uppfyller inte detta krav, eftersom sådan information inte är tillräckligt kortfattad och begriplig [58–59].

”Meningsfull information om den bakomliggande logiken” ska beskriva det förfarande och de principer som faktiskt har tillämpats på ett sådant sätt att den registrerade kan förstå vilka av dennes personuppgifter som har använts och på vilket sätt de har använts i det aktuella automatiserade beslutsfattandet. Att behandlingen är tekniskt komplicerad befriar inte den personuppgiftsansvarige från skyldigheten att lämna en förklaring [61].

När det gäller kreditprofilering uttalade EU-domstolen att en domstol kan anse att det är tillräckligt transparent och begripligt att informera den registrerade om i vilken utsträckning en förändring av de personuppgifter som beaktats skulle ha lett till ett annat resultat [62]. Detta behöver dock inte gälla för alla algoritmiska system.

Om den personuppgiftsansvarige anser att denna information inte kan lämnas ut utan att röja uppgifter om tredje man eller företagshemligheter, är den personuppgiftsansvarige skyldig att lämna den påstått skyddade informationen till den behöriga tillsynsmyndigheten eller domstolen. Denna ska därefter göra en avvägning mellan de berörda rättigheterna och intressena för att avgöra omfattningen av den registrerades rätt till tillgång enligt artikel 15 i förordningen.

Dessa båda avgöranden ger vägledning om hur EU-domstolen kan komma att bedöma i vilken utsträckning information ska göras tillgänglig och i vilken form enligt GDPR samt, analogt, vilka krav som kan ställas på de förklaringar som ska lämnas enligt AI-förordningen.

Intervju med den berörda personen

Klagande som framför en anmälan om diskriminering på grund av en eller flera skyddade egenskaper har ofta relevant information om vilka uppgifter de har lämnat till den svarande, antingen nyligen eller tidigare. Det kan också vara värdefullt att samla in annan information om deras skyddade egenskaper och om tillfällen då sådana egenskaper kan ha uppgetts. Var uppmärksam på dina egna skyldigheter enligt dataskyddsreglerna. Överväg frågor som liknar följande.

- ▶ När hade du första och eventuella efterföljande kontakter med denna organisation/myndighet? Om det har funnits tidigare kontakter före den händelse som är föremål för klagomålet kan information om skyddade egenskaper ha lämnats vid ett tidigare tillfälle.
- ▶ Blev du ombedd att fylla i formulär?
- ▶ Om ja, vilka formulär blev du ombedd att fylla i?
- ▶ Beakta förteckningen över skyddade egenskaper. Krävdes det att du lämnade uppgifter om eller kryssade i rutor som avsåg någon av dessa?
- ▶ Beakta förteckningen över kända proxyvariabler i bilaga E. Blev du ombedd att lämna uppgifter om någon av dessa?

Dokument från klaganden

Relevant information kan finnas i dokumentation som är tillgänglig för eller innehas av klaganden. Detta kan omfatta:

- ▶ dokument som klaganden själv lämnat in (ansökningsformulär, korrespondens, begäran om tillgång till uppgifter eller andra förfrågningar);
- ▶ dokument som den svarande har tillhandahållit klaganden (eventuella avtal mellan den svarande och klaganden tillsammans med bilagor, information om personuppgiftsbehandling eller integritetspolicyer samt intern kommunikation, inklusive avtalsuppdateringar).

Inlämning och hantering av begäran om registerutdrag (DSAR) enligt artikel 15 i GDPR

Registrerade har rätt att begära tillgång till sina personuppgifter från den personuppgiftsansvarige (vanligen den som tillhandahåller systemet) enligt artikel 15 i GDPR. Korrekt hanterade svar på en DSAR kan innehålla mycket relevant information (se steg 3 för detaljer), anpassad till individens specifika ärende, inklusive:

- ▶ kategorierna av behandlade personuppgifter samt en digital kopia av dessa uppgifter;
- ▶ ändamålen med behandlingen;
- ▶ mottagarna eller kategorierna av mottagare av personuppgifterna;
- ▶ förekomsten av automatiserat beslutsfattande och/eller profilering samt meningsfull information om den bakomliggande logiken och de förväntade konsekvenserna av sådan behandling för den registrerade.

Om en klagande ännu inte har lämnat in en DSAR rekommenderas att uppmuntra detta innan man går vidare med klagomålet eller utredningen. Det finns inga formella krav på hur en DSAR ska lämnas in, men det är lämpligt att göra begäran skriftligen och dokumentera den (spara kopia av e-post eller brev, spara bevis på post eller leverans, ta skärmdumpar vid användning av webbformulär, portal eller kundtjänstkanal). För en mall, se [bilaga A](#).

Underlåtenhet att följa GDPR:s krav beträffande DSAR är utbredd. Det är därför vanligt att möta uteblivna svar, försenade svar eller bristfälliga eller ofullständiga uppgifter efter en begäran. Klagande som utövar sina rättigheter som registrerade bör vara beredda på eventuell fortsatt kommunikation, inklusive påminnelser och begäran om kompletterande information från organisationen. I ärenden av stor betydelse kan samarbete med den relevanta nationella dataskyddsmyndigheten (DPA) övervägas (se [Sektorövergripande samarbete](#)).

Den föreslagna Digital Omnibus-förordningen innebär ändringar av den registrerades rätt till tillgång enligt GDPR. Det föreslagna tillägget till artikel 12.5 i GDPR³³

33. Skäl 35, artikel 3, förslag till Europaparlamentets och rådets förordning om ändring av förordningarna (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725 och (EU) 2023/2854 samt direktiven 2002/58/EG, (EU) 2022/2555 och (EU) 2022/2557 vad gäller förenkling av det digitala regelverket, och om upphävande av förordningarna (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868 och direktiv (EU) 2019/1024 (det digitala omnibuspaketet). Tillgänglig på: [EUR-Lex - 52025PC0837](#).

har kritiserats (Skiotyté och Sadauskaité 2026). Om det antas skulle det ge personuppgiftsansvariga möjlighet att avslå eller ta ut avgift för DSAR som anses vara överdrivet omfattande eller som anses ha ett "icke-dataskyddsrelaterat" syfte. Det bör dock inte påverka giltiga begäranden som görs för att utöva den registrerades rättigheter enligt dataskyddslagstiftningen.

2.1.2. Offentligt tillgänglig information

Externa kommunikationer från den svarande

Relevanta uppgifter kan kommuniceras av den svarande på deras webbplats eller plattform, inklusive i FAQ-sidor, allmänna villkor eller integritetspolicier, i enlighet med tillämpliga transparenskrav enligt GDPR (se rutan om GDPR:s transparens-skyldigheter och rättigheter ovan).

Lägg märke till hänvisningar till systemets ändamål, ändamålen med personuppgiftsbehandling, vilka typer av uppgifter som behandlas, inklusive eventuella hänvisningar till känsliga personuppgifter – även kallade särskilda kategorier av personuppgifter – samt eventuella uppgifter om förekomsten av (kvalificerat) automatiserat beslutsfattande (ADM) eller profilering, dess logik och konsekvenser. Detta kan indikera användning av algoritmer eller skyddade egenskaper (för mer information se [steg 3.2](#) och [3.3](#)).

Offentliga myndigheter i vissa länder (t.ex. Finland) omfattas av informationsskyldigheter och måste publicera information om användningen av ADM-system på sina webbplatser samt informera registrerade om sådan användning vid beslutsfattande. Det finns dock i stort sett ingen systematisk nationell eller transnationell förteckning, trots lokala initiativ att skapa databaser över AI- och ADM-tillämpningar (som i Belgien och Finland). Detta innebär att EB:er möter svårigheter i att utöva sitt uppdrag att utreda, övervaka, identifiera och ifrågasätta fall av algoritmisk diskriminering samt stödja utsatta (Council of Europe 2025a).

Gällande rättspraxis

Granska rättspraxis som rör antingen liknande AI- eller algoritmiska system eller automatiserade beslutsprocesser. Detta kan omfatta nationell rättspraxis, europeisk rättspraxis (inklusive rättspraxis från Europeiska domstolen för de mänskliga rättigheterna och EU-domstolen) samt utländsk nationell rättspraxis, eftersom algoritmiska system och plattformar ofta fungerar enligt liknande logik och återkommande mönster i olika länder.

Eftersom EU-rättspraxis om AI och algoritmreglering fortfarande är begränsad bör även GDPR-relaterad rättspraxis beaktas, särskilt avgöranden om automatiserat beslutsfattande, då dessa kan ge vägledande stöd. Överväg följande källor:

- ▶ Databaser såsom [Case law – EUR-Lex](#) eller [Europeiska unionens byrå för grundläggande rättigheters rättsdatabas för EU-domstolens avgöranden](#) eller andra domstolsavgöranden, med sökord som "artificiell intelligens", "automatiserat beslutsfattande", "profilering" och "kreditbedömning" för att granska befintliga

domar eller följa pågående nationella begäranden av förhandsavgöranden av EU-domstolen gällande relevanta aspekter som rör AI-förordningen eller GDPR.

- För databaser särskilt inriktade på AI eller algoritmiska system kan även [Case Law Database – Tech Litigation](#) användas, som samlar rättsfall om automatiserade system, inklusive domar, beslut och yttranden från nationella och internationella domstolar samt dataskyddsmyndigheter. Även [DAIL – Database of AI Litigation – Ethical Tech Initiative](#), som sammanställer amerikanska rättsfall om AI och algoritmiska system, kan ge värdefulla insikter om framväxande rättsliga trender kring algoritmisk diskriminering, systembrister och risker, kopplade till specifika tekniker, inklusive relevanta uppgifter om USA-baserade företag verksamma inom EU, antingen som leverantörer av liknande AI-system eller som personuppgiftsansvariga (t.ex. Clearview AI).
- Vissa civilsamhällesorganisationer börjar sammanställa databaser specifikt inriktade på fall av algoritmisk diskriminering ([AlgorithmWatch rapporteringsformulär](#)). Dessa databaser blir mer användbara om EB:er bidrar genom att rapportera sina egna rättsfall eller utredningsresultat. Detta måste ske i enlighet med dataskydds- och sekretessregler.
- [GDPR Enforcement Tracker – förteckning över GDPR-böter](#) och [Europeiska dataskyddsstyrelsens sanktionsbeslut](#): ger en översikt över böter och sanktioner som dataskyddsmyndigheter inom EU har beslutat om enligt GDPR, inklusive land, myndighet, datum, bötesbelopp, personuppgiftsansvarig/ biträde, sektor och typ.
- [Automatiserat beslutsfattande enligt GDPR](#): rapport från Future of Privacy Forum: en sammanställning av rättspraxis om ADM enligt GDPR från april 2022.
- Nationella rättsdatabaser.

Övriga källor

Det finns ett brett spektrum av datakällor som kan vara användbara för att identifiera, utreda eller styrka ett fall av algoritmisk diskriminering:

- Landsrapporter från Europeiska kommissionen mot rasism och intolerans (ECRI) 7:e övervakningscykel kommer att beakta AI.
- Rapporter från civilsamhällesorganisationer (CSO:er), som omfattar antingen liknande AI- eller algoritmiska system eller liknande beslutsprocesser.
- Slutsatser från FN:s övervakningsorgan, inklusive Kommittén för avskaffande av diskriminering av kvinnor (CEDAW), Kommittén för avskaffande av rasdiskriminering (CERD), Kommittén för barnets rättigheter, Kommittén mot tortyr (CAT), särskilda rapportörer, Världshälsoorganisationen (WHO) samt Kommittén för ekonomiska, sociala och kulturella rättigheter.
- Slutsatser från andra internationella organ, såsom OSSE:s kontor för demokratiska institutioner och mänskliga rättigheter (OSSE/ODIHR) samt Organisationen för ekonomiskt samarbete och utveckling (OECD). I detta avseende kan OECD:s register över AI-incidenter och risker ([OECD AI Incidents Monitor, an evidence base for trustworthy AI – OECD.AI](#)) vara särskilt relevant.

- ▶ Slutsatser från nationella specialiserade organ, inklusive ombudsmän, likabehandlingsorgan och icke-statliga organisationer.
- ▶ Administrativa register, såsom centrala och lokala folkbokföringsregister.
- ▶ Klagomålsdata såsom statistik från polis, åklagare samt domstols- och ärenderegister.
- ▶ Akademisk och ad hoc-forskning, inklusive resultat från EU:s forskningsinstitutioner såsom Europeiska unionens byrå för grundläggande rättigheter (FRA) och Europeiska institutet för jämställdhet (EIGE). En sökning per domän i [MIT AI Risk Repository](#) kan också vara informativ.
- ▶ Jämförande information om rättspraxis, policyer och praxis mellan medlemsstaterna ("europeisk konsensus"-data).

2.1.3. Information som tillhandahålls av den svarande eller andra parter

Information som begärs från den svarande med stöd av nationella eller andra unionsrättsliga befogenheter

Likabehandlingsorgan kan begära information från svarande och andra parter, till exempel den som tillhandahåller eller använder systemet (om detta är en annan organisation än den svarande), genom att utöva sina utrednings- och informationsinhämtningsbefogenheter enligt relevant lagstiftning, inklusive nationella lagar som genomför standarddirektiven,³⁴ lagstiftning om tillgång till allmänna handlingar samt internationella instrument såsom Europarådets ramkonvention om artificiell intelligens³⁵ eller genom samarbetskanaler med nationella dataskyddsmyndigheter.

Information kan lämnas ut under utredningar, förhandlingar och pre-processuella kontakter eller inom ramen för rättsliga förfaranden.

Undersökningar: standarddirektiven (artikel 8)

Artikel 8 i direktiv 2024/1499 och 2024/1500 kan stärka likabehandlingsorganens (EB:er) rätt till tillgång till information, särskilt:

- ▶ enligt artikel 8(1) i standarddirektiven ska medlemsstaterna ge likabehandlingsorgan befogenhet att genomföra en undersökning för att fastställa om ett brott mot principen om likabehandling har ägt rum;

34. Enligt artikel 16 i förordningen om likabehandlingsorgan har dessa organ en rad befogenheter när det gäller insamling av uppgifter, bland annat att samla in uppgifter för att utarbeta rapporter och genomföra oberoende undersökningar (artikel 16.2) samt rätten att få tillgång till statistiska uppgifter när detta är nödvändigt för att göra en övergripande bedömning av situationen vad gäller diskriminering i medlemsstaten och för att utarbeta rapporter (artikel 16.3).

35. Europarådets ramkonvention om AI (2024) föreskriver att parterna ska anta eller bibehålla åtgärder för att säkerställa adekvata krav på öppenhet och tillsyn (artikel 8) samt för att säkerställa ansvarsskyldighet och ansvar för negativa konsekvenser för de mänskliga rättigheterna (artikel 9). I artikel 14 krävs särskilt att parterna garanterar tillgängliga och effektiva åtgärder, inbegripet att dokumentera och tillhandahålla information om AI-system med betydande inverkan på de mänskliga rättigheterna till behöriga organ och, i förekommande fall, till berörda personer, så att enskilda kan bestrida AI-drivna beslut och lämna in klagomål till behöriga myndigheter.

- ▶ enligt artikel 8(2) i standarddirektiven ska medlemsstaterna inrätta ett ramverk för genomförandet av undersökningar som gör det möjligt för likabehandlingsorgan att utföra faktainsamling, inklusive:
 - effektiva rättigheter att få tillgång till information och handlingar som är nödvändiga
 - för att fastställa om diskriminering har förekommit;
 - lämpliga mekanismer för att likabehandlingslikabehandlingsorgan ska kunna samarbeta med relevanta offentliga organ i detta syfte.

Det är viktigt att notera att artikel 8.3 medger att medlemsstaterna kan överlåta dessa befogenheter till ett annat behörigt organ. Detta kan till exempel vara en MSA. I så fall är det MSA som utövar befogenheterna enligt artikel 8.1 och 8.2. När ett sådant behörigt organ har slutfört sin undersökning ska det, på begäran, tillhandahålla likabehandlingslikabehandlingsorganet information om resultatet av undersökningen.

Om MSA har anförtratts dessa befogenheter är det viktigt att notera att detta enligt AI-förordningen kräver samarbete med likabehandlingslikabehandlingsorganet (artikel 79.2).

- ▶ undersökningar kan vara särskilt relevanta beträffande system som klassas som icke högrisk, där det har varit svårt att få fram dokumentation, eller där det finns misstanke om gruppdiskriminering men otillräcklig bevisning i enskilda fall.

Observera: inga undersökningar får inledas eller fortsätta om domstolsförfaranden i samma ärende pågår. Undersökningar styrs sannolikt av nationella rättsbestämmelser.

Konsekvensbedömning av dataskydd (DPIA)

All behandling av personuppgifter som sannolikt medför en hög risk för den registrerades rättigheter på grund av behandlingens art, omfattning, sammanhang eller syften kräver att den personuppgiftsansvarige utför en konsekvensbedömning av dataskydd (DPIA) i enlighet med artikel 35.1 i GDPR. Följaktligen är DPIA:er mycket relevanta i fall som rör automatiserat beslutsfattande och profilering genom algoritmsystem (artikel 35.3 i GDPR). Dock, bortsett från en snäv grupp av användare av högrisk-AI-system som har publiceringsskyldigheter enligt AI-förordningen (se även DPIA under steg 2.2.3 nedan), finns det ingen rättslig skyldighet enligt vare sig GDPR eller AI-förordningen för användaren att publicera eller lämna ut DPIA:n till ett likabehandlingsorgan. All kontroll över konsekvensbedömningen är beroende av dataskyddsmyndigheternas egna initiativ – samarbetskanaler med dataskyddsmyndigheter kan vara relevanta i detta avseende. Detta bör inte avskräcka EB:et från att göra en sådan förfrågan. Vissa organisationer kan gå med på att lämna ut DPIA:n om den finns – antingen enligt artikel 35.3 i GDPR eller frivilligt – eller kan till och med välja att publicera en sammanfattning av den. En DPIA kan begäras under korrespondensen inför rättsprocessen eller om en rättsprocess inleds.

Konsekvensbedömningar avseende mänskliga rättigheter och grundläggande rättigheter (HRIA/FRIA)

En uttrycklig skyldighet att utföra någon typ av FRIA förutses i AI-förordningen endast för vissa användare och leverantörer av högrisk-AI-system (se även bedömning av grundläggande rättigheter under steg 2.2.3 nedan). För andra algoritm- eller AI-system finns ingen sådan skyldighet enligt AI-förordningen. Detta hindrar dock inte leverantörer från att frivilligt anta sådana instrument (European Center for Not-for-Profit Law och Danish Institute for Human Rights 2025; FRA 2025), baserat på deras övergripande skyldighet att respektera och skydda grundläggande rättigheter som fastställts på EU-nivå och nationell nivå, och som ett sätt att stärka AI-operatörers ansvarsskyldighet i detta avseende (Mantelero 2024).

På samma sätt är en HRIA en allmän term för processen och verktyget att identifiera, utvärdera och åtgärda potentiella och faktiska negativa konsekvenser av produkter eller verksamheter på mänskliga rättigheter, och det finns olika sätt att genomföra denna bedömning. HUDERIA-metodiken och modellen som utvecklats av Europarådet är en typ av bedömning av risker för mänskliga rättigheter (Europarådet 2025c, 2026).

På samma sätt kan en HRIA eller FRIA ha genomförts frivilligt av utvecklare för icke-högrisk-AI-system – inklusive enligt deras egna tillbörlig noggrannhetsplikt- eller människorättsåtaganden eller i enlighet med ramverk som FN:s vägledande principer för företag och mänskliga rättigheter. Nationella eller regionala lagar, policyer eller praxis kan också kräva att bedömningar gällande konsekvenserna för mänskliga rättigheter eller grundläggande rättigheter genomförs och publiceras.

I alla dessa fall kan likabehandlingsorganen begära utlämnande av dessa konsekvensbedömningar som en del av sina informationsinsamlingsbefogenheter.

Jämställdhets- eller könsrelaterad konsekvensbedömning

Likabehandlingsorgan kan begära konsekvensbedömningar ifall dessa krävs enligt nationell rätt. Det kan saknas en rättslig skyldighet att genomföra eller lämna ut dem.

Dokumentation enligt AI-förordningens ramverk

Där ett system inte klassificeras som högrisk kan likabehandlingsorganen ändå begära jämförbar information från användaren eller leverantören med stöd av sina utrednings- eller informationsinsamlingsbefogenheter. Även om svaranden inte är rättsligt skyldig att ha utarbetat och/eller lämnat ut sådan dokumentation för icke-högrisk- eller icke-AI-system kan likabehandlingsorganen EB:er vägledas av AI-förordningens ramverk när de strukturerar sina förfrågningar och utvärdera svarandens reaktion – eller uteblivna reaktion – i enlighet med detta i steg 3 och steg 4.

Undersökningar och statistisk information

Likabehandlingsorganen kan genomföra eller få tillgång till undersökningar, inklusive urvalsundersökningar som kartlägger erfarenheterna hos grupper som utsatts för diskriminering, och opinionsundersökningar som mäter förekomsten av fördomar (standarddirektiven, artikel 16). De kan också få tillgång till relevanta officiella

statistikkällor, såsom de som sammanställs av nationella statistikmyndigheter i samband med EU:s antidiskrimineringsdirektiv, samt data som samlats in genom folkräkningar och andra stora källor som Labour Force Survey (LFS) och Survey on Income and Living Conditions (SILC).

Standarddirektiven

- ▶ Enligt artikel 16.2 i direktiven 2024/1499 och 2024/1500 får likabehandlingsorganen genomföra oberoende undersökningar om diskriminering.
- ▶ Enligt artikel 16.3 i direktiven 2024/1499 och 2024/1500 måste likabehandlingsorganen kunna få tillgång till statistik som rör de rättigheter och skyldigheter som följer av EU:s antidiskrimineringsdirektiv när det är nödvändigt för att göra en övergripande bedömning av situationen när det gäller diskriminering i medlemsstaten, och för att upprätta periodiska rapporter om deras medlemsstat.

Dokumentation enligt annan produktsäkerhetslagstiftning

För icke-högriskprodukter som säljs av ett företag till en konsument kan det finnas krav enligt den allmänna produktsäkerhetsförordningen³⁶ (GPSR) och MSA-förordningen,³⁷ som inkluderar skyldigheter och dokumentationskrav som liknar dem i AI-förordningen. Detta ligger utanför omfattningen av denna metodik eftersom befogenheterna tillhör MSA; dock nämns det eftersom GPSR betraktas som ett säkerhetsnät för icke-högrisk-AI (skäl 166 i AI-förordningen) och dess krav kan leda till att relevant dokumentation är tillgänglig antingen direkt för den enskilde eller för likabehandlingsorganet i samarbete med MSA. Till exempel krävs det av tillverkare som omfattas av förordningen att deras produkt åtföljs av tydliga instruktioner och säkerhetsinstruktioner för konsumenter (artikel 9.7) och att upprätta teknisk dokumentation och genomföra riskbedömningar som ska göras tillgängliga för MSA (artiklarna 9.2–3 och 15).

2.2. Ytterligare information som är tillgänglig för högrisk-AI-system

2.2.1. Ytterligare information som är tillgänglig för klaganden

Att göra och erhålla svar på en begäran om förklaring enligt artikel 86 i AI-förordningen

Denna rätt kan åberopas utöver DSAR-begäran (se [Inlämning och hantering av begäran om registerutdra \(DSAR\) enligt artikel 15 i GDPR](#)). Varje person som är föremål för ett beslut som väsentligt och negativt påverkar personen och som fattas på grundval av output från ett högrisk-AI-system, har rätt att erhålla information om AI-systemets roll i beslutsprocessen och de huvudsakliga elementen i beslutet. Dessa ska vara tydliga

36. Förordning (EU) 2023/988.

37. Förordning (EU) 2019/1020.

och meningsfulla förklaringar. Rätt hanterade svar kan innehålla mycket relevant information, inklusive inputdata och de kriterier som påverkar beslutet (se steg 3 för detaljer) för den efterföljande identifieringen av ett prima facie-fall av diskriminering. Om en klagande ännu inte har lämnat in en sådan begäran rekommenderas det att uppmuntra personen att göra det innan man går vidare med något klagomål eller utredning. För en mall, se [bilaga B](#).

Denna rätt gäller endast i den utsträckning som rätten inte annars tillhandahålls enligt unionslagstiftning. Till exempel, om algoritmsystemet är ett helt automatiserat system som behandlar den drabbade personens personuppgifter, kan det vara så att rätten till förklaring tillhandahålls enligt GDPR i artiklarna 13–15 (förordning (EU) 2016/679). När svaranden avvisar förklaringsbegäran med denna motivering skulle en uppföljningsbegäran enligt GDPR (annan relevant unionslagstiftning) vara nästa steg (om den inte redan har utövats). För en mall, se [bilaga A](#).

2.2.2. Offentligt tillgänglig information

EU-databas

Se även [1.2.1 EU-databas](#).

I steg 1 genomförde du en preliminär kontroll av EU-databasen. Om svaranden är listad bör du redan ha identifierat om de är listade i databasen som:

- ▶ en leverantör, listad i avsnitt A eller B i databasen:
 - för ett högriskanvändningsfall som är relevant för klagandens situation;
 - för ett användningsfall som är avsett att användas inom ett högriskområde, men som faller under undantaget i artikel 6.3 i AI-förordningen;
- ▶ en användare, listad i del C i databasen:
 - en användare för ett användningsfall som är relevant för klagomålets situation; vilken organisation är leverantören?

Nedan följer de huvudsakliga intressanta uppgifterna i databasen (AI-förordningen, bilaga VIII).

Avsnitt A: högrisk-AI-system

- ▶ Kontaktuppgifter till leverantören och den auktoriserade representanten.
- ▶ Varumärkesnamn eller annan referens som möjliggör identifiering och spårbarhet av AI-systemet (detta kan vara användbart till exempel när ett högrisk-system används som en del av ett annat system).
- ▶ Avsett syfte.
- ▶ Information som används av systemet (data, indata) och dess funktionslogik (gäller inte system inom brottsbekämpning, migration, asyl och gränskontroll).
- ▶ Status för AI-systemet (till exempel om det fortfarande finns på marknaden?).
- ▶ Typ, nummer och utgångsdatum för det certifikat som utfärdats av det anmälda organet samt namn eller identifikationsnummer för det organet och en skannad kopia av certifikatet (gäller inte system inom brottsbekämpning, migration, asyl och gränskontroll).

- Medlemsstater där systemet är eller har varit marknadsfört (användbart för samarbete med andra myndigheter).
- En kopia av EU-försäkran om överensstämmelse (AI-förordningen, artikel 47).
- Elektroniska bruksanvisningar (gäller inte system inom brottsbekämpning eller migration, asyl och gränskontroll).

Avsnitt B: AI-system för vilket leverantören har dragit slutsatsen att det inte är högrisk

- Kontaktuppgifter till leverantören och den auktoriserade representanten.
- Varumärkesnamn eller annan referens som möjliggör identifiering och spårbarhet av AI-systemet.
- Avsett syfte.
- Relevant villkor enligt artikel 6.3 och motivering till varför systemet inte anses vara högrisk (gäller inte brottsbekämpning, migration, asyl eller gränskontroll).
- Status för AI-systemet (finns det fortfarande på marknaden?).
- Medlemsstater där systemet är eller har varit marknadsfört.

Avsnitt C: användare av högrisk-system (AI-förordningen, artikel 49.3)

- Namn och kontaktuppgifter till användaren (endast offentliga myndigheter, EU-institutioner, organ och byråer – eller personer som agerar å deras vägnar).
- URL till AI-systemets post i EU-databasen (som registrerats av leverantören).
- Sammanfattning av bedömningen av konsekvenserna på grundläggande rättigheter (FRIA) (gäller inte brottsbekämpning, migration, asyl eller gränskontroll).
- Sammanfattning av dataskyddskonsekvensbedömningen (DPIA).

Inte alla högrisk-AI-system är listade i den offentliga delen av EU-databasen.

Tabellen nedan visar i vilka användningsfall information ska tillhandahållas i databasen.

Tabell 3. Information som är tillgänglig i EU:s och nationella högriskdatabaser

Offentlig sektor (artikel 49.1–49.3), (med undantag för privata aktörer, såvida de inte agerar på uppdrag av en offentlig myndighet)	Avsnitt som inte är offentligt (artikel 49.4)	Nationellt genomförande/ register (artikel 49.5)
Alla uppgifter som beskrivs nedan är obligatoriska och finns i databasens offentliga del.	Endast ett begränsat antal uppgifter är obligatoriska och dessa finns endast i en säker, icke-offentlig del av databasen.	Registreringskraven gäller endast på nationell nivå.
Utbildning och yrkesutbildning	Brottsbekämpning	Kritisk infrastruktur
Sysselsättning, arbetsledning och möjligheter till egenföretagande	Migration, asyl och gränskontroll	

Viktiga tjänster och förmåner	Högrisk-AI-system under testning	
Rättskipning och demokratiska processer	Biometri som används inom brottsbekämpning, migrations- och asylfrågor samt gränskontroll	

Begränsningar av tillgång till information

Enligt artikel 49.4 d i AI-förordningen ska information om högrisk-AI-system inom rättsväsende, migration, asyl och gränsförvaltning (punkterna 1, 6 och 7 i bilaga III) som registreras i EU-databasen föras in i en icke-offentlig del och endast vara tillgänglig för Europeiska kommissionen och berörda MSA:er. För dessa ändamål är det endast dataskyddsmyndigheten eller Europeiska datatillsynsmannen (EDPS, för kommissionen) som är relevanta MSA:er (artiklarna 74.8 och 74.9) och som har tillgång till informationen i databasen.

Detta innebär inte att det underliggande material som uppgifterna bygger på också är otillgängligt. Det innebär dock att ytterligare motivering krävs för att få tillgång till handlingarna och att sådan tillgång kan nekas. Likabehandlingsorgan (EB) som är artikel 77-myndigheter har en rätt att begära och få tillgång till all information eller dokumentation som skapats eller upprätthålls enligt förordningen (se nedan, steg 2.2.3). En sådan begäran måste vara tillräckligt väl motiverad för att visa att den är nödvändig för att effektivt kunna fullgöra deras uppdrag inom ramen för deras rättskipningsområde.

Förordningen klargör också tydligt att marknadskontrollmyndighete "inte på något sätt ska påverka domstolarnas oberoende eller på annat sätt inkräkta på deras verksamhet när de agerar i sin dömande funktion" (artikel 74.8). Detta innebär att i fråga om information som ska göras tillgänglig för EB eller enskilda genom rättsliga förfaranden, gäller inte begränsningarna i artikel 74.8. Tillgång till känsligt material, även genom sådana förfaranden, regleras sannolikt av andra sekretesskrav samt hänsyn till allmän och nationell säkerhet (se [Konfidentiell behandling](#)).

Om rättsliga grunder för informationsdelning mellan likabehandlingslikabehandlingsorgan, MSA:er och andra artikel 77-myndigheter har inrättats (se [Sektorsövergripande samarbete](#)), skulle det vara möjligt för likabehandlingslikabehandlingsorgan att få tillgång till information som erhållits av MSA:er enligt artiklarna 71.4 och 49.4 (Council of Europe 2025b).

2.2.3. Ytterligare information som tillhandahålls av den svarande eller andra parter

AI-förordningens dokumentations- och rapporteringsskyldigheter

Se även [EU:s AI-förordningen, artikel 77: Befogenheter för myndigheter som skyddar grundläggande rättigheter](#).

AI-förordningen skapar skyldigheter för leverantörer och användare av högrisk-AI-system att ta fram och upprätthålla olika typer av dokumentation under systemets hela livscykel. Likabehandlingsorgan som är artikel 77-myndigheter har rätt att begära och få tillgång till all dokumentation som skapats eller upprätthålls enligt AI-förordningen, i tillgängligt språk och format, när det är nödvändigt för att effektivt fullgöra deras uppdrag inom ramen för deras behörighet.³⁸ (EU:s AI-förordning, artikel 77)

Denna dokumentation anges i tabellen nedan över olika typer av teknisk dokumentation.

Det rekommenderas att begära all denna information, eftersom alla dessa dokument tillsammans ger en tydlig bild av systemets användning, konstruktion, data och risker.

I steg 3 ges detaljerade förklaringar vad denna information kommer att visa och hur du kan använda den.

Dokumentationen kommer sannolikt att vara organiserad utifrån universella och sektorsspecifika dokumentationsstandarder som branschen har utvecklat under de senaste åren, inklusive hänvisningar till "datablad" och "modellkort" i bilaga IV. För likabehandlingsorgan skulle det vara till hjälp att känna till dessa två typer av dokumentation (Mittelstadt 2024). Den struktur och det innehåll som krävs för dessa typer av dokumentation kommer att regleras av standarder från Joint Technical Committee-21 (JTC-21) (se nedan, [Harmoniserade standarder](#)).

Se även [Samarbete mellan sektorer](#).

Nedan beskrivs en lista över krav som ska uppfyllas för högrisk-AI-system enligt AI-förordningen. Även om inte alla krav uttryckligen kräver dokumentation, kan man utgå från att det åtminstone kommer att finnas intern dokumentation om vart och ett av dessa krav.

38. Likabehandlingsorgan som inte omfattas av artikel 77 har ingen ytterligare direkt rättslig grund enligt AI-förordningen för att få tillgång till relevant dokumentation för högrisk-AI-system och kommer att behöva förlita sig på sina befintliga befogenheter. Det är fortfarande tveksamt om de kommer att kunna skaffa sig tillräcklig insikt i hur högrisk-AI-system fungerar för att kunna fullgöra sina uppdrag på ett effektivt sätt.

Tabell 4. Krav på AI-system med hög risk enligt EU:s AI-förordning

Krav (Artikel)	Beskrivning av skyldighet	Roll
Riskhanterings-system (artikel 9)	Ett riskhanteringssystem som identifierar och analyserar rimligt förutsebara risker för hälsa, säkerhet eller grundläggande rättigheter, i enlighet med riskhanteringsrättsmedlen.	Leverantör
Data och dataförvaltning (artikel 10)	Datamängder för träning, validering och testning ska omfattas av datastyrning. Detta innefattar bland annat designval, datainsamling och datakälla, relevant förberedelse, formulerade antaganden, granskning av systematiska fel samt lämpliga åtgärder för att förebygga partiskhet.	Leverantör
Teknisk dokumentation för ett AI-system med hög risk (artikel 11, bilaga IV)	Allmän beskrivning av AI-systemet samt detaljerad beskrivning av vissa delar och processer. En mer detaljerad beskrivning ges i tabellen nedan.	Leverantör
Loggning (artikel 12)	Systemen bör registrera händelser som är relevanta för att identifiera risker för hälsa, säkerhet eller grundläggande rättigheter och underlätta övervakning efter marknadsintroduktionen. Detta innefattar registrering av användningsperiod, referensdatabas, indata samt identifiering av fysiska personer som deltar i verifikationen av resultaten.	Leverantör
Transparens (artikel 13) och mänsklig kontroll (artikel 14)	Systemens drift bör vara tillräckligt transparent. Detta innefattar systemets egenskaper, kapacitet och begränsningar, tillsammans med dess avsedda ändamål och prestandamått. Dessutom bör mänsklig tillsyn och de resurser som krävs för systemets funktion ingå.	Leverantör till användare
Kvalitetsstyrnings-system (artikel 17)	Dokumentation av policyer, rutiner och instruktioner, inklusive efterlevnadsstrategi, tekniker för utformning och kvalitetskontroll samt granskning före, under och efter utvecklingen.	Leverantör

Krav (Artikel)	Beskrivning av skyldighet	Roll
Automatiskt genererade loggar (artikel 19)	Detta omfattar även tekniska specifikationer och system för datahantering. Slutligen ska dokumentation beträffande hantering av allvarliga incidenter, kommunikation med myndigheter, dokumentförvaring, resurshantering och ansvarsfördelning finnas.	Leverantör
Automatiskt genererade loggar (artikel 26)	Leverantörer måste bevara automatiskt genererade loggar i minst sex månader. Finansiella institutioner ska bevara loggarna i enlighet med tillämplig lagstiftning om finansiella tjänster.	Användare
Konsekvensbedömning av grundläggande rättigheter (artikel 27)	Driftsansvariga måste bevara automatiskt genererade loggar i minst sex månader. Finansiella institut bevarar loggarna i enlighet med tillämplig lagstiftning om finansiella tjänster.	Användare
EU-försäkran om överensstämmelse (artikel 47, bilaga V)	Leverantören ska upprätta en (maskinläsbar) skriftlig och undertecknad EU-försäkran om överensstämmelse för varje högrisk-AI-system, som ska vara tillgänglig för nationella myndigheter i 10 år efter det att systemet har släppts ut på marknaden, och i vilken det anges att systemet uppfyller kraven i avsnitt 2 i AI-förordningen.	Leverantör
CE-märkning (artikel 48)	Systemen omfattas av de allmänna principerna i artikel 30 i förordning (EG) nr 765/2008 och ska förses med CE-märkning.	Leverantör
Uppgifter om övervakning efter utsläppande på marknaden (artikel 72)	Leverantörerna måste inrätta ett system för övervakning efter marknadsintroduktionen, som samlar in, dokumenterar och analyserar relevanta uppgifter om systemets prestanda, för att möjliggöra en bedömning av efterlevnaden.	Leverantör
Rapportering av incidenter (artikel 73)	Leverantörerna är skyldiga att rapportera allvarliga händelser till marknadsövervakningsmyndigheterna i respektive medlemsstat.	Leverantör till MSA. MSA ska involvera EB:er.

Den tekniska dokumentation som leverantörer ska upprätta enligt artikel 11 i AI-förordningen, tillsammans med bilaga IV till AI-förordningen, kommer att utgöra en särskilt värdefull informationskälla för likabehandlingsorgan (EB).

Följande tabell fokuserar på denna tekniska dokumentation och visar vilka obligatoriska delar som kan vara relevanta för likabehandlingsorgan vid tillämpningen av denna metod.

Tabell 5. Olika typer av teknisk dokumentation enligt artikel 11 och bilaga IV
(bearbetat med tillstånd från Mittelstadt 2024)

Typ av dokumentation	Förklaring
1. Allmän beskrivning av AI-systemet	Detta omfattar: ▶ avsett användningsändamål; ▶ interaktioner med systemet; ▶ versionshantering och reflektion; ▶ utseende i programvara och hårdvara, eller i produkter där den ingår som komponent; ▶ hårdvara som den är avsedd att användas på; ▶ användargränssnitt och instruktioner för den som driftsätter systemet.
2. Konstruktions-specifikationer	Detta omfattar: ▶ metoder och steg för utvecklingen av systemet; ▶ konstruktionsspecifikationer; logik, algoritm och viktiga val, inklusive antaganden och optimeringsval; ▶ relevansen hos de valda parametrarna; ▶ förväntat resultat och dess kvalitet; ▶ möjliga avvägningar när det gäller efterlevnad.
3. System-arkitektur	Beskrivning av mjukvarukomponenternas struktur, deras samverkan och integration samt de beräkningsresurser som används för utveckling, träning, testning och validering.
4. Datakrav	Datakraven omfattar datablad som beskriver de träningsmetoder, tekniker och datamängder som använts, med information om deras ursprung, omfattning och egenskaper, insamling, märkningsförfaranden och rensningsmetoder.
5. Mänsklig tillsyn	Bedömning av de åtgärder för mänsklig tillsyn som införts i enlighet med artikel 14, tillsammans med de tekniska åtgärder som krävs för att underlätta tolkningen av AI-systemet för de användare som tar det i bruk.
6. Prestanda-mått	Validerings- och testförfaranden som använts, tillsammans med information om testdatans huvudsakliga egenskaper, mått som använts för att mäta noggrannhet, robusthet och efterlevnad, samt potentiellt diskriminerande effekter. Innehåller även alla testloggar och testrapporter, inklusive förutbestämda ändringar.

Typ av dokumentation	Förklaring
7. Funktioner	Det krävs detaljerad information om övervakning, funktion och styrning av AI-systemet, särskilt dess kapacitet och begränsningar. Detta inkluderar noggrannhetsgrader för specifika grupper.
8. Riskhantering	Riskhanteringssystem enligt artikel 9.
9. Övervakning efter marknadsintroduktion	En detaljerad beskrivning av det system som finns på plats för att utvärdera systemets prestanda efter marknadsintroduktionen i enlighet med artikel 72.
10. Övrig EU-lagstiftning	Tillämpade harmoniserade standarder, kopia av EU-försäkran om överensstämmelse (artikel 47).

Harmoniserade standarder

Harmoniserade standarder är tekniska standarder som för närvarande håller på att utarbetas av de europeiska standardiseringsorganisationerna (CEN/CENELEC). Om dessa följs ger de leverantörer av AI-system med hög risk en "antagande om överensstämmelse" med AI-förordningen. I praktiken kommer de sannolikt att omfatta ett brett spektrum av informationselement, inklusive sådana som behövs för att stödja förståelsen av hur AI-systemet fungerar, dess egenskaper, kapacitet, styrkor, begränsningar och prestanda, i enlighet med den rättsliga regleringen (Soler Garrido et al. 2024).

Följande standarder som för närvarande håller på att utvecklas har bedömts vara särskilt viktiga för att fastställa kraven på den tekniska dokumentation som ska vara tillgänglig för likabehandlingsorgan.³⁹

- **JT021036** "Artificial intelligence – Concepts, measures and requirements for managing bias in AI systems": Syftet med denna standard är att definiera "begrepp, metoder och krav för bedömning och hantering av bias i AI-system. Detta omfattar bias som är oönskad enligt AI-leverantörens och AI-användarens specifikation av AI-systemet inom ramen för AI-förordningen."
- **JT021008** "AI trustworthiness framework", del 1 (loggning), del 2 (noggrannhet och robusthet) och del 3 (transparens och mänsklig tillsyn): syftet med denna standard är att definiera terminologi, begrepp, krav och vägledning beträffande loggning, noggrannhet, robusthet, transparens och mänsklig tillsyn.
- **JT021024** "AI risk management": syftet med denna standard är att fastställa krav och ge vägledning för riskhantering av AI-system under hela AI-systemets

39. För en förteckning och beskrivning av samtliga standarder som för närvarande utvecklas, se CEN/CENELEC, "CEN/CLC/JTC 21 Work Programme", tillgänglig på: [CEN – CEN/CLC/JTC 21](#), hämtad den 25 juni 2026. Även andra standarder, däribland redan befintliga internationella standarder (såsom ISO/IEC-standarder), kan vara relevanta vid hantering av bias och transparens i AI-system. Internationella standarder är vanligtvis inte kostnadsfria och omfattas av immaterialrättsligt skydd, vilket kan försvåra likabehandlingsorganens tillgång till dem.

livscykel. Dokumentet kräver att leverantörer fastställer objektiva kriterier för riskacceptans, men anger inte vilka risknivåer som är godtagbara.

Dessa standarder kommer sannolikt inte att fastställa tröskelvärden eller förklara när olik behandling av personer eller grupper av personer utgör olaglig diskriminering. Enligt Mittelstadt (2024) bör EB:er vara medvetna om att standarderna kommer att ange förväntningar på hur bias kan mätas och begränsas i AI-system, men de kommer inte att avgöra om sådana bias är olagliga eller oetiska.

Standarder, bedömning av överensstämmelse och certifikat

AI-system med hög risk som överensstämmer med harmoniserade standarder ska presumeras uppfylla kraven i avsnitt 2 i AI-förordningen (krav för AI-system med hög risk). Det är dock viktigt att notera att efterlevnad av transparenskrav eller andra skyldigheter inte innebär att användningen av AI-systemet eller dess resultat är laglig enligt AI-förordningen eller annan lagstiftning och påverkar inte andra transparenskyldigheter för användare av AI-system enligt unionsrätten eller nationell rätt (skäl 137). Leverantörer ska ange vilka harmoniserade standarder som har tillämpats på AI-systemet. Om harmoniserade standarder inte har använts ska leverantören förklara hur kraven i AI-förordningen (kapitel III, avsnitt 2) har uppfyllts på annat sätt, exempelvis genom användning av andra relevanta standarder och tekniska specifikationer. Detta bör göra det möjligt för likabehandlingsorgan att avgöra om en relevant standard beträffande bias, tillförlitlighet eller harmoniserad riskhantering har följts.

Betydelsen av ett antagande om överensstämmelse

Ett antagande om överensstämmelse hindrar inte utredning eller andra åtgärder med anledning av diskriminering eller andra risker för de grundläggande rättigheterna. För det första erkänner AI-förordningen själv att ett AI-system som uppfyller förordningens krav ändå kan utgöra en risk, och artikel 82 innehåller ett förfarande för att hantera en sådan situation. På motsvarande sätt föreskrivs i artikel 7.3 i den allmänna produktsäkerhetsförordningen (EU) 2023/988 (som kan vara tillämplig på AI-system som inte är AI-system med hög risk) att ett antagande om överensstämmelse med de allmänna säkerhetskraven i förordningen inte hindrar MSA:en från att vidta alla lämpliga åtgärder enligt förordningen om det finns bevis för att produkten, trots antagandet om överensstämmelse är farlig.

Observera: Enligt marknadskontrollförordningen (EU) 2019/1020 ska MSA:en, när en ekonomisk aktör lägger fram provningsrapporter eller certifikat som styrker att produkten överensstämmer med unionens harmoniseringslagstiftning, ta "vederbörlig hänsyn" till rapporten eller certifikatet.

Konsekvensbedömning av grundläggande rättigheter

Kravet på att genomföra en konsekvensbedömning av grundläggande rättigheter (Fundamental Rights Impact Assessment, FRIA) gäller endast vissa användare av AI-system med hög risk, i enlighet med artikel 27 och bilaga III punkt 5 b och c i AI-förordningen. Detta omfattar:

- a. användare av AI-system med hög risk som är offentliga organ eller privata enheter som tillhandahåller offentliga tjänster;
- b. användare av AI-system med hög risk som används för att bedöma fysiska personers kreditvärdighet eller fastställa deras kreditpoäng (med undantag för AI-system som används för att upptäcka finansiella bedrägerier), samt AI-system som används för att bedöma risker eller fastställa premier inom liv- och sjukförsäkring.

I Digital Omnibus om AI erkänns att det finns en överlappning mellan FRIA och konsekvensbedömningen av dataskydd (Data Protection Impact Assessment, DPIA). Det kan därför vara nödvändigt att begära in DPIA:n om denna hänvisas till i FRIA:n.⁴⁰ i stället för att vara införlivad i den.

När FRIA:n har slutförts är användaren enligt artikel 27.3 och 27.5 i AI-förordningen skyldig att underrätta den behöriga MSA:en om resultaten genom att lämna in en ifylld (standardiserad) mall. Dessutom ska en sammanfattning av FRIA:ns slutsatser lämnas in av användaren som en del av registreringsprocessen i EU-databasen, för de användare som omfattas av registreringsskyldigheten enligt artikel 49.3 i AI-förordningen. Användare som omfattas av den andra kategorin (kredit- och försäkringssektorn) omfattas inte av denna skyldighet. Likabehandlingsorganen kan därför förvänta sig att finna en sammanfattning i avsnitt C i EU-databasen, enligt bilaga VIII, endast för de obligatoriska FRIA:er som genomförts av offentliga organ eller privata enheter som omfattas av den första kategorin.

Observera: Tillämpningen av båda dessa skyldigheter är begränsad när offentliga myndigheter använder AI-system med hög risk inom brottsbekämpning, migration, asyl samt gränsförvaltning och gränskontroll. Enligt artikel 49.4 i AI-förordningen ska någon sammanfattning av FRIA:ns slutsatser inte lämnas till EU-databasen – inte ens till dess säkra, icke-offentliga del. Därutöver får sådana användare enligt artikel 27.3 i AI-förordningen, i undantagsfall – bland annat av hänsyn till den allmänna säkerheten – undantas från skyldigheten att underrätta MSA:erna om resultaten av FRIA:n, med stöd av undantaget från förfarandet för bedömning av överensstämmelse enligt artikel 46.1 i AI-förordningen (se även rutan nedan om brottsbekämpningsområdet). Likabehandlingsorganens möjligheter att få tillgång till information i sådana fall är därför begränsade enligt AI-förordningen.

Användare av andra kategorier av AI-system med hög risk – till exempel ett rekryteringsföretag som använder AI för att granska eller välja ut kandidater – är inte rättsligt skyldiga att genomföra en FRIA. Detta innebär dock inte nödvändigtvis att någon FRIA inte genomförs i sådana fall. Vissa organisationer kan göra det enligt nationell lagstiftning, i enlighet med ramverk såsom FN:s vägledande principer för företag och mänskliga rättigheter, eller för att uppfylla sina egna åtaganden om tillbörlig aktsamhet eller mänskliga rättigheter. I dessa fall kan likabehandlingsorgan begära utlämnande av information inom ramen för sina befogenheter.

40. Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).

Leverantörer av högrisk-AI-system omfattas inte av kravet att genomföra en konsekvensbedömning av grundläggande rättigheter av den specifika typ som avses i artikel 27 i AI-förordningen. Att genomföra en FRIA som ett allmänt verktyg för att identifiera, utvärdera och hantera potentiella eller faktiska negativa konsekvenser på grundläggande rättigheter, med utgångspunkt i ramverket för konsekvensbedömning av mänskliga rättigheter (HRIA), ingår dock i överensstämmelsebedömningen och leverantörens skyldighet att upprätthålla ett riskhanterings- och kvalitetsledningssystem, enligt vad som anges i artiklarna 9.2, 17.1 g och 43 i AI-förordningen (Mantelero 2024). Denna typ av konsekvensbedömning skulle vanligtvis följa en standardbaserad metod och kan likaså begäras, tillsammans med övrig dokumentation som förs enligt AI-förordningen, av likabehandlingsorgan som utövar sina befogenheter enligt artikel 77 i AI-förordningen (se ovan, under 2.1.3).

Utöver de rättsliga kraven i EU:s AI-förordning kan nationell eller regional lagstiftning, policyer eller praxis kräva att konsekvensbedömningar av mänskliga rättigheter eller grundläggande rättigheter genomförs och i vissa fall offentliggörs. I sådana fall kan likabehandlingsorgan begära utlämnande av information inom ramen för sina befogenheter.

Konsekvensbedömning av dataskydd (DPIA)

All behandling av personuppgifter som sannolikt leder till hög risk för de registrerades rättigheter på grund av behandlingens art, omfattning, sammanhang eller ändamål kräver att den personuppgiftsansvarige genomför en konsekvensbedömning av dataskydd (DPIA) enligt artikel 35.1 i GDPR. DPIA:er är därför per definition mycket relevanta i ärenden som rör AI-system med hög risk där personuppgifter behandlas och där automatiserat beslutsfattande och profilering förekommer (jfr artikel 35.3 i GDPR).

Kravet att lämna in sammanfattningar av DPIA:er till EU-databasen (avsnitt C) gäller dock endast vissa begränsade kategorier av användare av AI-system med hög risk som enligt artikel 49.3 i AI-förordningen är skyldiga att registrera sin användning av AI i EU-databasen. Detta omfattar offentliga myndigheter, unionsinstitutioner, organ, kontor och byråer – eller personer som agerar på deras vägnar – som använder AI-system med hög risk. För övriga användare och övriga AI-system med hög risk finns det inget direkt rättsligt krav enligt vare sig GDPR eller AI-förordningen att offentliggöra eller lämna ut DPIA:n till ett likabehandlingsorgan. Tillsynen över konsekvensbedömningen ligger i stället hos dataskyddsmyndigheterna, och samarbetskanaler med dessa kan därför vara relevanta. Detta bör dock inte avhålla likabehandlingsorganen från att begära in DPIA:n inom ramen för sina befogenheter över informationsinhämtning enligt nationell rätt. Vissa organisationer kan gå med på att lämna ut en DPIA om den finns tillgänglig eller välja att offentliggöra en sammanfattning av den. DPIA:n kan också begäras under korrespondens före en rättslig process eller inom ramen för en domstolsprocess.

Förslaget till Digital Omnibus-förordning⁴¹ kräver att Europeiska dataskyddsstyrelsen (EDPB) utarbetar gemensamma EU-omfattande förteckningar över behandlingsaktiviteter som kräver respektive inte kräver en DPIA, tillsammans med en gemensam

metodik och mall.⁴¹ kräver att Europeiska dataskyddsstyrelsen (EDPB) utarbetar gemensamma EU-omfattande förteckningar över behandlingsaktiviteter som kräver respektive inte kräver en DPIA, tillsammans med en gemensam metodik och mall.⁴² Detta förväntas ersätta det nuvarande fragmenterade systemet med nationella förteckningar och harmonisera dessa på EU-nivå.

2.2.4. Underrättelse från MSA:en eller utredning på eget initiativ (*ex officio*)

Likabehandlingsorganen kan förvänta sig att få information från MSA om AI-system som utgör en risk för rätten till likabehandling eller som är inblandade i allvarliga incidenter som innebär intrång i dessa rättigheter. Närmare information om vilken typ av uppgifter som kan förväntas lämnas finns i [1.1 C: Underrättelse från MSA](#).

Likabehandlingsorganen kan få tillgång till den dokumentation som krävs enligt AI-förordningen via MSA:en, om denna redan har inhämtat dokumentationen från leverantören. Detta kan påskynda EB:ets tillgång till informationen.

Respondenter kan inte hindra MSA från att dela dokumentationen med EB med hänvisning till skydd för företagshemligheter, särskilt när EB:et begärdokumentationen genom MSA:en (Europarådet 2025b). För ett motsvarande exempel på hur detta har tillämpats i praktiken, se CK v. Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien, punkt 76.

När ett ärende handläggs som en utredning på eget initiativ (*ex officio*) eller efter en underrättelse från en MSA bör följande övervägas:

- ▶ begära att MSA:en lämnar ut de handlingar som är nödvändiga för EB:ets utredning;
- ▶ sammankalla ett rundabordssamtal eller en workshop med relevanta icke-statliga organisationer (NGO:er) eller andra berörda organisationer och använda diskussionen för att samla in relevant information;
- ▶ öka medvetenheten och genomföra samråd för att identifiera berörda personer eller grupper, bland annat genom enkäter eller andra deltagandebaserade metoder för informationsinsamling.

41. Artikel 3 i förslaget till Europaparlamentets och rådets förordning om ändring av förordningarna (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725 och (EU) 2023/2854 samt direktiven 2002/58/EG, (EU) 2022/2555 och (EU) 2022/2557 vad gäller förenkling av den digitala lagstiftningsramen, och om upphävande av förordningarna (EU) 2018/1807, (EU) 2019/1150 och (EU) 2022/868 samt direktiv (EU) 2019/1024 (Digital Omnibus), [EUR-Lex – 52025PC0837 – EN – EUR-Lex](#), hämtad den 25 juni 2026.

42. Europeiska dataskyddsstyrelsen genomförde i maj 2026 ett samråd om en mall för konsekvensbedömning av dataskydd (DPIA), som finns tillgänglig här: www.edpb.europa.eu/our-work-tools/documents/public-consultations/2026/edpb-dpia-template_en.

Brottsbekämpningsområdet

Möjligheterna att samla in information är mer begränsade inom brottsbekämpningsområdet, vilket gör det svårare att identifiera användningen av AI eller algoritmiska system och eventuell diskriminering som följer av sådan användning.⁴³

AI-förordningen

- ▶ Enligt artikel 49.4 i AI-förordningen ska registreringen av AI-system med hög risk inom brottsbekämpning, migration, asyl samt gränsförvaltning och gränskontroll inte bara ske i en säker, icke-offentlig del av den EU-databas som avses i artikel 71 i AI-förordningen, utan registreringen ska också endast innehålla ett begränsat urval av uppgifter. Detta innebär exempelvis att information om indata och systemets funktionslogik inte ska ingå. Denna icke-offentliga del av databasen är dessutom endast tillgänglig för de MSA:er som utsetts för AI-system inom brottsbekämpning och gränskontroll i enlighet med artikel 74.8 och är i princip inte tillgänglig för EB:er. För ytterligare information, se [Begränsningar av tillgång till information under 2.2.2](#).
- ▶ Enligt artikel 27.1 i AI-förordningen är brottsbekämpande myndigheter, i egenskap av offentliga myndigheter, skyldiga att genomföra en konsekvensbedömning av grundläggande rättigheter (FRIA) innan ett AI-system med hög risk tas i bruk. Två undantag gäller dock. För det första ska sådana användare enligt artikel 49.4 c i AI-förordningen inte lämna in någon sammanfattning av resultaten från FRIA:n eller konsekvensbedömningen av dataskydd (DPIA) till EU-databasen – inte ens till dess säkra, icke-offentliga del. För det andra får de, i undantagsfall, bland annat av hänsyn till den allmänna säkerheten, efter godkännande från MSA:en släppa ut ett AI-system med hög risk på marknaden eller ta det i bruk utan att uppfylla förfarandena för bedömning av överensstämmelse, eller – i brådskande fall – helt utan föregående godkännande (artiklarna 46.1 och 46.2 i AI-förordningen). Detta innebär enligt artikel 27.3 i AI-förordningen även ett undantag från skyldigheten att underrätta MSA:en om resultaten av FRIA:n.

Observera: Artiklarna 27 och 46 i AI-förordningen innehåller inte något undantag från skyldigheten att genomföra en FRIA som sådan. Undantaget i artikel 46.1 är uttryckligen begränsat till den "begränsade tidsperiod under den tid under vilken de nödvändiga förfarandena för bedömning av överensstämmelse genomförs". Enligt artikel 46.2 ska dessutom tillstånd i brådskande fall sökas under eller efter användningen, utan onödigt dröjsmål.

Det är viktigt att notera att om tillstånd därefter vägras:

- ▶ användningen av AI-systemet med hög risk ska omedelbart upphöra;
- ▶ samtliga resultat och utdata från sådan användning ska omedelbart kasseras.

De möjliga vägarna för EB att få tillgång till information som rör både DPIA och FRIA behandlas nedan.

43. För en jämförelse mellan rättigheter beträffande algoritmiska data i GDPR och LED, se González Fuster (2020).

Brottsbekämpningsdirektivet

Därutöver innehåller brottsbekämpningsdirektivet (LED)⁴⁴ flera undantag från och möjliga begränsningar av enskildas rätt till information genom medlemsstaternas nationella genomförandebestämmelser, om detta är nödvändigt för att inte äventyra förebyggande, upptäckt, utredning eller lagföring av brott eller verkställighet av straffrättsliga påföljder samt för att skydda den allmänna säkerheten och den nationella säkerheten (artiklarna 13.3 och 15 i LED).

Praxis skiljer sig därför åt mellan medlemsstaterna när det gäller vilken typ av information som ska lämnas ut och hur den kan göras tillgänglig. Hänvisning till nationell lagstiftning är därför av avgörande betydelse på detta område.

Automatiserat beslutsfattande

- ▶ Artikel 11 i LED förbjuder vissa former av automatiserad behandling, däribland profilering, och kräver att sådan behandling uttryckligen ska vara tillåten enligt unionsrätten eller nationell rätt samt omfattas av lämpliga skyddsåtgärder för den registrerades rättigheter och friheter. Den lägsta skyddsnivå som krävs är rätten att få mänsklig medverkan från den personuppgiftsansvariges sida. Det är därför nödvändigt att granska de nationella reglerna om automatiserad behandling inom brottsbekämpningen för att fastställa vilken information om sådan behandling som ska vara tillgänglig. För olika genomföranden i medlemsstaterna, se Vogiatzoglou och Marquiene (2022); se även skäl 38.
- ▶ Artiklarna 13 och 14 i LED innehåller inte några uttryckliga informationsskyldigheter beträffande användning av automatiserat beslutsfattande, inklusive profilering.

Utlämnande av information

- ▶ Enligt artikel 13.1 i LED är personuppgiftsansvariga skyldiga att lämna viss minimiinformation, däribland uppgifter om den personuppgiftsansvariges identitet och kontaktuppgifter, den registrerades rätt till tillgång till sina personuppgifter samt de avsedda ändamålen med personuppgiftsbehandlingen.

Denna information avser inte en viss registrerad utan ett visst behandlingsförfarande och alla registrerade som potentiellt berörs av det. Den görs normalt tillgänglig på den behöriga myndighetens webbplats. Det har rapporterats att det inte alltid är enkelt eller ens möjligt att hitta relevant information på webbplatsen för den centrala myndighet som ansvarar för de brottsbekämpande myndigheterna, den nationella polismyndigheten eller det behöriga ministeriet, se Vogiatzoglou och Marquiene (2022).

- ▶ Artikel 13.2 i LED utvidgar omfattningen av den information som ska lämnas till de registrerade. Den omfattar bland annat den rättsliga grunden för behandlingen och, i förekommande fall, de kategorier av mottagare som personuppgifterna

44. Hänvisningen till LED i fråga om informationskällor ska inte anses vara uttömmande; andra rättsakter inom det bredare området för frihet, säkerhet och rättvisa, såsom rådets beslut om Schengen Information System II (SIS II), Visa Information System (VIS) och European Asylum Dactyloscopy Database (EURODAC), kan också innehålla materiella och processuella bestämmelser om de registrerades rättigheter, inbegripet rätten till information och tillgång.

kommer att lämnas ut till samt, när det är nödvändigt, ytterligare information, särskilt när personuppgifterna har samlats in utan den registrerades kännedom. Enligt artikel 13.3 får sådan information enligt nationell rätt senareläggas, begränsas eller helt utelämnas om en sådan åtgärd är nödvändig och proportionerlig i ett demokratiskt samhälle.

Om informationen är tillräckligt specifik kan den göras tillgänglig på myndighetens webbplats tillsammans med den obligatoriska minimala informationen.

Rätt till tillgång (rätt till registerutdrag)

- ▶ Enligt artikel 14 i brottsbekämpningsdirektivet (LED) har registrerade rätt att få tillgång till mer information om personuppgiftsbehandlingen än den allmänna information som görs tillgänglig för allmänheten. Detta omfattar bland annat uppgifter om ändamålen med behandlingen, vilka kategorier av uppgifter som behandlas samt mottagarna eller kategorierna av mottagare till vilka personuppgifterna lämnas ut, liksom information om de personuppgifter som behandlas och deras ursprung.

Till skillnad från GDPR innehåller artikeln ingen uttrycklig rätt att informeras om förekomsten av automatiserat beslutsfattande eller den bakomliggande logiken, eller om de möjliga konsekvenserna för den registrerade. Brottsbekämpande myndigheter kan dock enligt nationell lagstiftning vara skyldiga att lämna information om hur en individ har profilerats.

- ▶ Enligt skäl 43 i LED behöver den information som lämnas inte innehålla en faktisk kopia av de behandlade uppgifterna, såsom enligt GDPR. I stället är en fullständig sammanfattning tillräcklig. Innehållet i svaret – när tillgång inte nekas på formella grunder – består ofta av standardiserade svar som är kortfattade och saknar substantiell information, men som ändå anses uppfylla de formella kraven (Erdogan 2024).
- ▶ Enligt artikel 15 i LED får rätten till tillgång skjutas upp, begränsas eller helt uteslutas enligt nationell lagstiftning, om en sådan åtgärd är nödvändig och proportionerlig i ett demokratiskt samhälle för att bland annat skydda den allmänna säkerheten eller den nationella säkerheten eller för att inte hindra officiella eller rättsliga utredningar, förundersökningar eller förfaranden.

Begäran om tillgång

- ▶ Artiklarna 14 och 17.3 i LED föreskriver två olika förfaranden för att utöva rätten till tillgång. "Direkt tillgång" enligt artikel 14 i LED, där den registrerade vänder sig direkt till den brottsbekämpande myndigheten, det vill säga den personuppgiftsansvarige, med frågor om sina personuppgifter. "Indirekt tillgång", ett undantagsförfarande där rätten till tillgång utövas genom den nationella dataskyddsmyndigheten (DPA), som fungerar som mellanhand.

Nationell lagstiftning som genomför systemet med indirekt tillgång, såsom i Belgien där indirekt tillgång är huvudregeln, har kritiserats för sina strukturella och processuella brister (Dimitrova och De Hert 2024), vilka begränsar möjligheterna till effektiv tillsyn av personuppgiftsbehandling och försvårar förmedlingen av relevant information till de berörda registrerade.

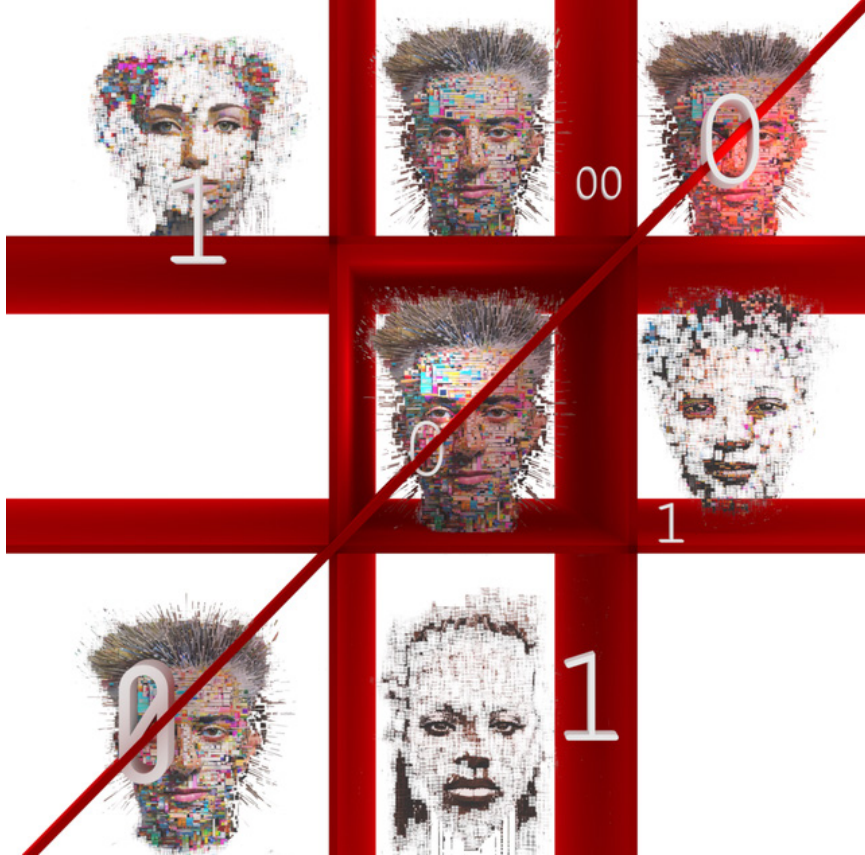
Observera: När information samlas in av en behörig myndighet för brottsbekämpande ändamål omfattas den av brottsbekämpningsdirektivet (LED) – exempelvis uppgifter om en persons tidigare brottshistorik. Om informationen däremot hämtas från en annan offentlig källa eller samlas in för andra ändamål än brottsbekämpning (såsom civilrättsliga, arbetsrättsliga eller utbildningsrelaterade ändamål), omfattas den av GDPR. I vissa AI-system som används inom brottsbekämpning kan båda typerna av indata användas. Det är därför viktigt att avgöra om de tillämpliga informationsrättigheterna regleras av GDPR eller av LED. Situationen blir ännu mer komplex när privata företag deltar i utvecklingen av programvaran eller i behandlingen eller kontrollen av relevanta personuppgifter, se Erdogan (2023). Vid osäkerhet bör specialistkompetens inom juridik anlitas.

Register över behandlingsverksamhet

- Enligt artikel 24 i LED är den personuppgiftsansvarige skyldig att föra ett register över behandlingsverksamheten, vilket bland annat ska innehålla uppgifter om behandlingens ändamål, de kategorier av mottagare till vilka personuppgifterna har lämnats ut eller kommer att lämnas ut samt, i förekommande fall, användning av profilering. Registret ska på begäran göras tillgängligt för dataskyddsmyndigheterna (för möjligheter till samarbete med dataskyddsmyndigheter, se [Samarbete mellan sektorer](#)).

Konsekvensbedömning avdataskydd (DPIA)

I brottsbekämpningssammanhang kräver artikel 27 i LED att en konsekvensbedömning av dataskydd (Data Protection Impact Assessment, DPIA) genomförs när en typ av behandling som använder ny teknik sannolikt medför en hög risk för fysiska personers rättigheter och friheter. Skyldigheten att genomföra en DPIA varierar mellan medlemsstaterna beroende på hur direktivet har genomförts i nationell rätt. Arbetsgruppen för dataskydd rekommenderade i artikel 29 att nationella lagstiftare ska ålägga personuppgiftsansvariga att genomföra en DPIA i samband med automatiserade beslut. Precis som för konsekvensbedömningar av grundläggande rättigheter (FRIA) är brottsbekämpande myndigheter dock inte skyldiga att lämna in någon sammanfattning av DPIA:n till EU-databasen (artikel 49.4 c i AI-förordningen).



Steg 3

Metod för att bedöma den information som har erhållits

Steg 3 introducerar de olika typer av diskriminering som kan orsakas av AI-system, oavsett om de används för automatiserat beslutsfattande eller för att stödja beslutsfattare.

Introduktionen till centrala begrepp beskriver olika former av diskriminering och visar hur de kan uppstå i en AI-kontext. Nya begrepp för likabehandlingsorgan kan bland annat vara proxydiskriminering och variabler. Avsnittet introducerar också hur en misstanke om diskriminering eller ett *prima facie*-fall av diskriminering kan identifieras och fastställas.

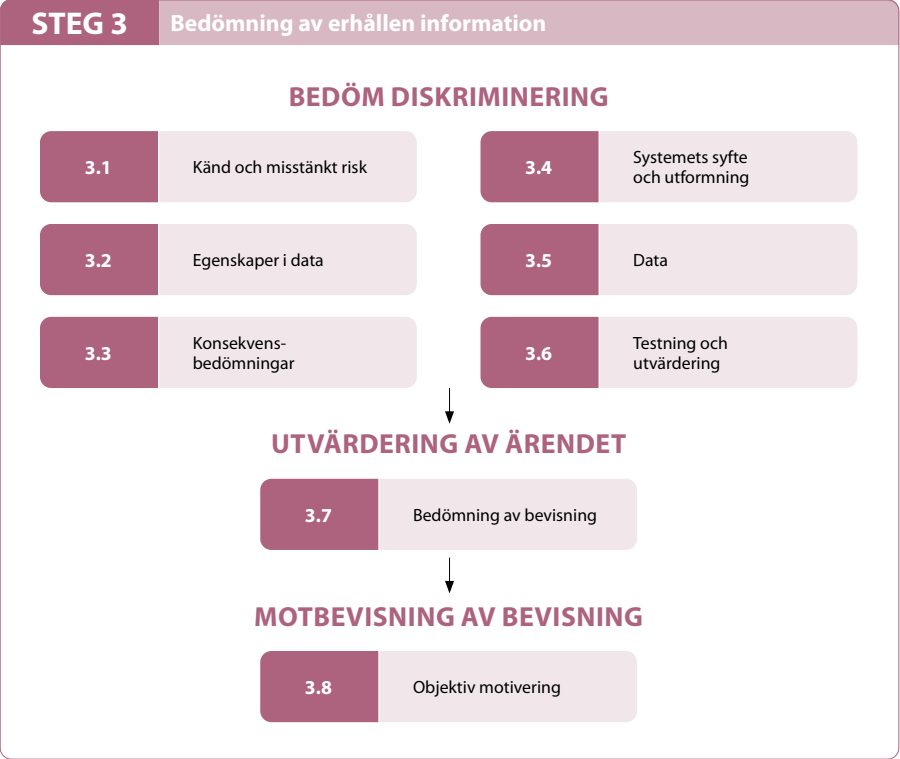
Den praktiska delen av steg 3 är utformad för att hjälpa handläggare att bedöma den information som samlats in enligt steg 2. Det exempel på dokumentationsblad som finns i bilaga K kan användas parallellt för att dokumentera informationen. I steg 3.2 till 3.6 får handläggaren en kort förklaring, en rad frågor att överväga, möjliga informationskällor som kan hjälpa till att besvara dessa frågor samt exempel på tecken på diskriminering eller diskriminerande praxis.

Steg 3.7 sammanför dessa slutsatser och ger vägledning om hur man bedömer om det finns en misstanke omdiskriminering. Avsnittet behandlar också, vilket är särskilt viktigt, hur man bör gå vidare när den tillgängliga bevisningen är otillräcklig eller när det råder bristande transparens.

Steg 3.8 ger vägledning om hur sannolikt det är att svaranden kan motbevisa eller rättfärdiga den påstådda diskrimineringen.

Det förutsätts inte att handläggare ska kunna besvara alla frågor eller ha tillgång till alla dokument eller all bevisning. Syftet är i stället att ge dem stöd för att göra en så välgrundad bedömning som möjligt utifrån den information som faktiskt finns tillgänglig och för att förstå dess betydelse.

Figur 4. Steg 3 – Bedömning av erhållen information



I steg 2 bör du ha påbörjat processen med att samla in relevant dokumentation och annan bevisning. I steg 3 ska denna bevisning användas för att bedöma om det går att fastställa en misstanke om diskriminering eller ett prima facie-fall av diskriminering, samt om det är sannolikt att svaranden kommer att kunna motbevisa en sådan misstanke eller lämna en godtagbar rättfärdigandegrund.

Gällande fler av delstegen i steg 3 kan teknisk expertis vara till hjälp. En sådan expert bör ha goda kunskaper inom statistik, datavetenskap och AI. Begrepp som är centrala för utvecklingen av algoritmiska system, såsom träningsdata, referensetiketter

(ground truth labels), utdata och indata, förklaras inte närmare i denna metodik, eftersom de förutsätts utgöra grundläggande kunskap inom statistik, datavetenskap och AI. När en teknisk expert anlitas bör det säkerställas att experten har förståelse för både datavetenskap och principerna om likabehandling.

I vissa fall kan MSA:er eller andra myndigheter i medlemsstaten ha den tekniska expertis som likabehandlingsorganen kan använda sig av (se [Samarbete mellan sektorer](#)).

Dessutom kan sektorsspecifik expertis ibland vara värdefull för att förstå mönster av historiska och befintliga ojämlikheter inom den sektor där systemet används. När sektorsspecifik expertis anlitas kan det vara lämpligt att välja experter med bakgrund inom samhällsvetenskap och empirisk forskning, eftersom en sådan bakgrund bidrar till att överbrygga klyftan mellan sektorspecifik kunskap och teknisk expertis. Om detta inte är möjligt kan EB:et behöva genomföra ytterligare forskning om historiska och andra former av ojämlikhet för att kunna bedöma om dessa kan ha påverkat systemet. Exempel kan vara välkänd överbevakning av vissa samhällsgrupper eller inkomstskillnader mellan kvinnor och män.

En rekommendation om att involvera experter anges, där det är relevant, under varje delsteg.

Observera: Som det framgår i denna metodik kan den nya artikel 77.1 b om krav på samarbete och ömsesidigt bistånd innebära att MSA eller en annan offentlig myndighet eller ett annat offentligt organ enligt AI-förordningen kan tillhandahålla den expertis som krävs.

Observera: Om det inte finns tillräcklig information tillgänglig behöver du överväga antingen:

- ▶ hur bristande transparens påverkar bevisbördan (se transparens under steg 3.7); eller
- ▶ att begära en utvärdering och testning av systemet av MSA (se steg 4.2.1)..

En rimlig kunskapsnivå om lagstiftning som berör likabehandling förutsätts för att använda denna metodik. De grundläggande begreppen återges dock i början av detta steg för att tydligare visa hur de är kopplade till algoritmiska system.

Introduktion till begrepp

Former av diskriminering

Direkt diskriminering

En person behandlas mindre förmånligt än hur någon annan i en jämförbar situation behandlas eller skulle ha behandlats, och skälet till detta är en särskild egenskap som personen har och som omfattas av en skyddad diskrimineringsgrund (och endast

i EU-rättsliga mål som rör ålder, funktionsnedsättning och verkliga och avgörande yrkeskrav kan en rättfärdigandegrund komma i fråga; i övriga fall finns ingen möjlighet att rättfärdiga direkt diskriminering).

Exempel på direkt algoritmisk diskriminering

I Finland nekades en man ett lån av ett finansinstitut som använde ett poängbaserat bedömningssystem baserat på uppgifter såsom bostadsort, kön, ålder och modersmål för att bedöma risken för att ett lån inte skulle återbetalas.

Det konstaterades att mannen skulle ha beviljats lånet om han hade varit kvinna eller haft svenska som modersmål. Dessutom gjordes ingen individuell bedömning av hans faktiska risk att inte kunna återbetala lånet; han nekades helt enkelt lånet på grundval av statistiska antaganden kopplade till hans kön och modersmål. Den finska diskriminerings- och jämställdhetsnämnden (National Non-Discrimination and Equality Tribunal) konstaterade år 2018, efter en begäran om utredning från Diskrimineringsombudsmannen, att detta utgjorde ett fall av direkt diskriminering.

Ytterligare information

- ▶ Bedömning av kreditvärdighet, myndighetsutövning, direkt flerfaldig diskriminering, kön, språk, ålder, bostadsort, ekonomiska skäl, föreläggande om vite, Finlands diskriminerings- och jämställdhetsnämnd (2018).

Indirekt diskriminering

En till synes neutral bestämmelse, ett kriterium eller ett förfaringsätt (provision, criterion or practice – PCP) som medför att personer som tillhör en skyddad grupp missgynnas särskilt i jämförelse med andra personer i en jämförbar situation, såvida inte:

- ▶ bestämmelsen, kriteriet eller förfaringsättet är objektivt motiverat av ett berättigat syfte; och
- ▶ medlen för att uppnå detta syfte är lämpliga och nödvändiga.

Det finns ännu inget bindande avgörande om huruvida vissa eller alla automatiserade beslutsprocesser utgör en bestämmelse, ett kriterium eller en praxis, men det råder en stark uppfattning bland vissa experter att så nästan säkert är fallet: "Enligt vår uppfattning utgör automatiserade beslutsprocesser som använder algoritmer nästan säkert bestämmelser, kriterier eller praxis i samband med en talan om indirekt diskriminering" (Allen och Masters 2019).

Exempel på indirekt algoritmisk diskriminering (Deliveroo-målet)

Deliveroos algoritmiska system för självbetjäningssbokning "self-service booking" av arbetspass befanns vara indirekt diskriminerande, särskilt på grund av religion eller annan trosuppfattning (deltagande i kollektiva aktioner såsom strejker) och funktionsnedsättning. Genom att basera tilldelningen av arbetspass på statistik

över deltagande och tillförlitlighet – till synes neutrala kriterier – missgynnade systemet i praktiken bud som av legitima skäl inte kunde arbeta, exempelvis på grund av strejk, sjukdom eller omsorgsansvar. Detta ledde till lägre poäng och därmed sämre möjligheter att få framtida arbetspass.

Förbiseendet av viktiga faktorer som präglade den statistik som användes – det vill säga systemets oförmåga att ta hänsyn till orsaker till frånvaro som hade samband med skyddade diskrimineringsgrunder – innebär att olika situationer behandlades på samma sätt. Detta ledde till en oproportionerligt negativ påverkan på skyddade grupper.

Ytterligare information

- *Filcams Cgil Bologna m.fl. mot Deliveroo Italia S.R.L.* (2020)
 - *Purificato* (2021)
-

Multipel diskriminering

Diskriminering kan förekomma på grund av mer än en orsak. Till exempel kan en person drabbas av diskriminering på grund av ras och på grund av kön eller funktionsnedsättning. Det finns flera olika termer som används för att beskriva diskriminering som förekommer på mer än en orsak. Några exempel på termer som används är "flera diskrimineringsgrunder", "kumulativ diskriminering", "sammansatt diskriminering", "kombinerad diskriminering" och "intersektionell diskriminering" (Fredman 2016). Dessa har alla subtilt olika betydelser. För denna metodik är det viktigt att inse att grunden för diskriminering kan vara mångfacetterad (Xenidis 2021).

Intersektionell diskriminering

"I grund och botten ... är intersektionalitet ett sätt att tänka kring identitet och dess förhållande till makt" (Crenshaw 2015).

Intersektionell diskriminering skiljer sig från andra former av flera diskrimineringsgrunder genom att den inte bara är en förstärkning av diskriminering på grund av enskilda orsaker, utan olika personliga egenskaper överlappar varandra och kan inte separeras. Sociala, ekonomiska, kulturella och institutionella strukturer låser sig samman och bildar ett komplext system som skapar (och vidmakthåller) ojämlikhet (Crenshaw, 1991). Till exempel kan sociala kategorier som ålder och kön överlappa varandra för att skapa fördel eller nackdel beroende på omständigheterna. Ojämlikhet kan förändras beroende på omständigheter och fakta och är ibland tids- och platsberoende. Till exempel kan en icke-migrantkvinna vara "mindre" jämlik än en man men mer jämlik än en migrantkvinna. Sociala ojämlikheter kan också forma maktrelationer. Dessa sammanlänkade strukturer av ojämlikhet återspeglas i den verkliga data som används för att utforma, bygga och testa algoritmsystem. Dessutom kan maskininlärningsmetoder som används för att bygga AI-system hitta interagerande mönster som inte är kända eller uppenbara från grundläggande dataanalyser. På så sätt kan algoritmsystem leda till intersektionell diskriminering (Xenidis, 2021). Även om intersektionell diskriminering avvisades av EU-domstolen i Parris-målet (C-443/17), anser vissa kommentatorer att detta behöver omprövas särskilt med

tanke på det ökande erkännandet av fenomenet i unionslagstiftning såsom Pay Transparency-direktivet 2023/970/EU och direktiv (EU) 2024/1385 om bekämpande av våld mot kvinnor och våld i hemmet (Howard 2024).

Proxy-diskriminering

Diskriminering som grundar sig på användning av egenskaper, så kallade proxies, som i sig själva inte tillhör någon av de uttryckligen förbjudna diskrimineringsgrunderna (se [skyddade egenskaper](#)), men som korrelerar med en skyddad grund, benämns proxy-diskriminering. Även om EU-lagstiftningen inte uttryckligen erkänner eller hänvisar till proxy-diskriminering är det ett teoretiskt begrepp som hjälper till att förstå hur algoritmisk diskriminering kan uppstå.

- ▶ Proxy-diskriminering kan utgöra direkt diskriminering om det finns en oskiljbar koppling mellan de valda – neutrala – egenskaperna och de uttryckligen skyddade (graviditet och kön till exempel: *Dekker v. Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus*, 1990; *Dita Danosa v. LKB Lizings SIA*, 2010; *Andersen v. Region Syddanmark*, 2010; *VL mot Szpital Kliniczny im. dra J. Babińskiego SPZOZ i Krakowie*, 2021).
- ▶ Alternativt kan proxy-diskriminering utgöra indirekt diskriminering om bestämmelsen, kriteriet eller förfarandet, inklusive proxies, påverkar personer med skyddade egenskaper på ett oproportionerligt sätt (för en förklaring av detta i praktiken, se *Jyske Finans v. Ligebehandlingsnævnet*, 2017 (C-668/15)).

Viktigt rättsfall: *Dekker v. Stichting Vormingscentrum* (1990)

En arbetsgivare handlade i direkt strid mot principen om likabehandling (rådets direktiv 76/207/EEG) om genomförande av likabehandling för män och kvinnor vad gäller tillgång till sysselsättning, yrkesutbildning och befordran samt arbetsvillkor när den vägrade att ingå ett anställningsavtal med en kvinnlig sökande som var lämplig för jobbet, där vägran grundade sig på de möjliga negativa konsekvenserna av att anställa en gravid kvinna eftersom det nationella sjukförsäkringssystemet likställde graviditet med sjukdom. Domstolen konstaterade att endast kvinnor kan vägras anställning på grund av graviditet och att detta utgjorde direkt diskriminering på grund av kön.

Proxyvariabler

I diskrimineringslagstiftning avser "proxies" eller "proxyvariabler" egenskaper eller kombinationer av sådana som i sig själva inte listas bland de uttryckligen förbjudna diskrimineringsgrunderna enligt det relevanta rättsliga ramverket, men som korrelerar med dessa grunder på ett sådant sätt att deras användning kan leda till förbjuden diskriminering. Ett exempel på detta är ett postnummer: ett postnummer är i sig inte diskriminerande; dock, om många migranter eller personer av ett visst etniskt ursprung bor i ett visst område, kan det fungera som en proxy för en förbjuden grund, nämligen etniskt ursprung.

Proxies kan identifieras genom överrepresentation: om personer med en specifik skyddad egenskap är överrepresenterade i en grupp är det troligt att indikatorer för denna grupp är proxies. Således kan, i exemplet ovan, eftersom personer av olika ursprung är överrepresenterade i vissa stadsdelar, ett postnummer erkännas som en proxy.

När en proxy används som indata till ett algoritmsystem kan detta indikera potentiell diskriminering. Observera att beroende på typen av proxy kan användningen av proxies indikera både direkt och indirekt diskriminering (Weerts m.fl. 2024). Vissa egenskaper är kända, genom forskning på området och relevant rättspraxis, att vara proxies för skyddade egenskaper. Proxies kan också identifieras genom att förstå och analysera det specifika sammanhang i vilket ett algoritmsystem tillämpas, med beaktande av befintliga historiska ojämlikheter och demografi.

För detaljer och exempel på olika typer av proxies, se [bilaga E](#).

I Chez-målet (C-83/14) infördes en kollektiv åtgärd (placering av elmätare i ett "romskt" område på en höjd mellan sex och sju meter i stället för mindre än två meter i andra områden) i ett område som huvudsakligen, men inte uteslutande, beboddes av personer av romskt ursprung. Om det visade sig att denna åtgärd hade införts och/eller upprätthållits av skäl som rör det etniska ursprung som är gemensamt för de flesta invånarna i det berörda området, skulle den anses utgöra direkt eller indirekt diskriminering, inklusive mot de invånare i det området som inte har romskt ursprung. Detta var en fråga för den nationella domstolen att avgöra genom att beakta alla relevanta omständigheter i målet och de regler om omvänd bevisbörda som föreskrivs i artikel 8.1 i direktiv 2000/43.

Algoritmexempel: kontroll av stipendier (college grant check)

Detta exempel bör hjälpa till att förklara hur proxies fungerar. I detta fall fungerade "bo nära föräldrarna" och "följa yrkesutbildning" som proxies för "icke-europeisk migrantbakgrund".

Den nederländska utbildningsmyndigheten (DUO) genomför kontroller för att säkerställa att utbetalda stipendier beviljas korrekt. Denna process är känd som "college grant check" (på nederländska förkortat till CUB). Misstanke uppstod att den nu avvecklade CUB-processen var diskriminerande mot studenter med migrantbakgrund. Denna misstanke grundade sig på överklagandeärenden som initierats av studenter som bedömts ha felaktigt beviljats stipendier, där advokater konstaterade att deras klienter kom från migrantbakgrunder. Utredningen visade att ett bedrägeriskbedömningssystem användes i denna process för att prioritera utredningar av (o)laglig användning av stipendierna.

Algorithm Audit genomförde en analys av algoritmsystemet, skillnader i behandling i de separata stegen i CUB-processen och undersökte om indata till systemet kunde

betraktas som proxies.⁴⁵ Systemet visade sig använda tre indata: studentens ålder, utbildningsnivå och avståndet mellan deras adress och deras föräldrar. Nationalitet, migrantbakgrund eller annan information om ras eller etniskt ursprung användes inte i bedrägeripoängssystemet.

Algorithm Audits viktigaste slutsatser var att:

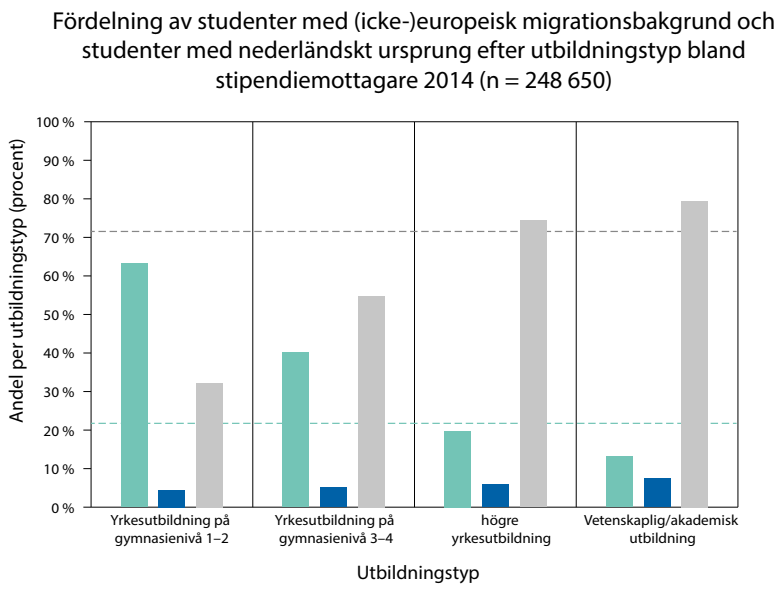
- ▶ CUB-processen som helhet ledde till ojämn behandling av studenter med icke-europeisk migrantbakgrund;
- ▶ riskprofilen tilldelade högre riskpoäng till studenter med migrantbakgrund; studenter med icke-europeisk migrantbakgrund var dubbelt så benägna att tilldelas en högriskkategori som studenter av nederländskt ursprung;
- ▶ det ytterligare manuella urvalet av studenter för kontroll förstärkte överrepresentationen snarare än minskade den. Studenter med icke-europeisk migrantbakgrund var 6,2 gånger mer benägna att väljas ut för en kontroll, som innebar ett inträngande hembesök, än studenter av nederländskt ursprung.

En nederländsk domstol kom fram till att denna praxis inte var objektivt motiverad och utgjorde indirekt diskriminering (Rechtbank Overijssel 2024).

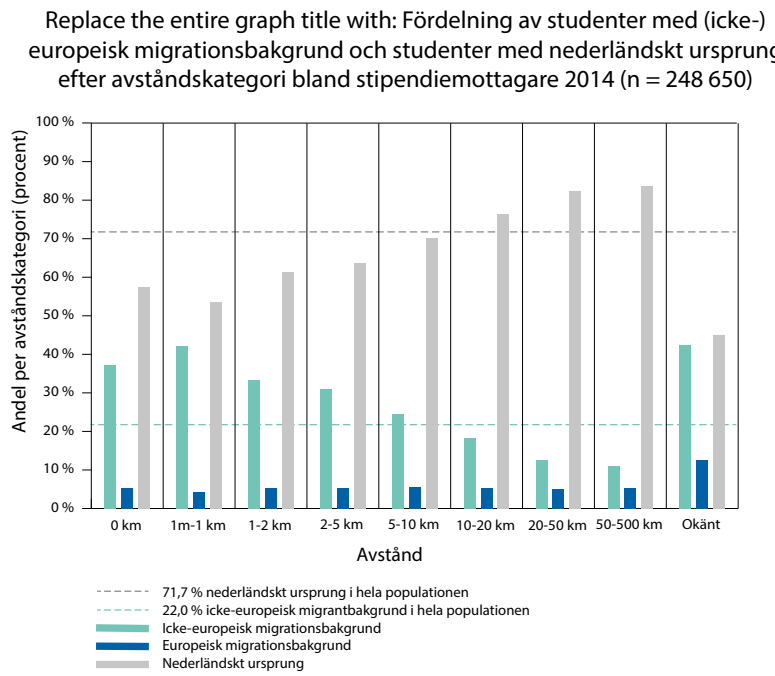
Detaljerade analyser visade att riskbedömningssystemet tilldelade högre riskpoäng till studenter på lägre utbildningsnivåer (yrkesutbildning) och studenter som bodde närmare sina föräldrar. Båda variablerna visade sig korrelera (starkt) med migrantbakgrund och är därför proxies för migrantbakgrund (se graferna nedan).

45. Algorithm Audit beviljades full tillgång till både algoritmen och DUO:s interna data. På grund av den allmänna dataskyddsförordningen (GDPR) samlar DUO inte in eller behandlar information om studenternas ursprung i syfte med CUB-processen. För att genomföra dessa analyser samarbetade DUO och Algorithm Audit med det nederländska nationella statistikmyndigheten (Netherlands Statistics eller CBS). CBS tillhandahöll aggregerad statistik om födelseland och ursprungsland baserat på populationer som definierats av Algorithm Audit för varje steg i processen (riskkategori, utvald för kontroll, kontrollresultat, överklagandeärende, överklaganderesultat) och indatavariabler (utbildning, ålder, avstånd till förälder/föräldrar).

Graf 1



Graf 2



Källa: Algorithm Audit

För en fullständig förklaring, se Algorithm Audit 2024a och 2024b.

Jämförelsepersoner

En jämförelseperson är en person i en liknande situation som klaganden men som inte delar den skyddade egenskap som antas vara skälet till diskrimineringen.

Observera: det är möjligt att två grupper av personer anses vara i en analog situation för syftet med ett klagomål men inte ett annat. Det är därför alltid viktigt att bedöma jämförelsepersonen mot bakgrund av syftet med den ifrågasatta åtgärden (Europarådet och Europeiska unionens byrå för grundläggande rättigheter 2018: 44–49).

Valet av rätt jämförelseperson är särskilt viktigt vid bedömningen av bevisning som rör algoritmiska system, särskilt när systemets syfte samt rättsmedlen för träning, testning och validering och den faktiska testningen och utvärderingen ska bedömas. Det kan särskilt vara nödvändigt att bedöma om uppgifter om skyddade egenskaper har använts för att träna det algoritmiska systemet och vilka val som har gjorts vid testningen av systemet. Detta behandlas mer ingående i avsnitten 3.3, 3.5 och 3.6.

Fastställande av en misstanke om diskriminering

Enligt antidiskrimineringsdirektiven ska det, när personer som anser sig ha blivit lidande av att principen om likabehandling inte har tillämpats på dem fastställa inför en domstol eller annan behörig myndighet fakta som ger anledning att misstänka att direkt eller indirekt diskriminering har förekommit, åligga svaranden att bevisa att det inte har skett någon överträdelse av principen om likabehandling.⁴⁶ Se även [steg 3.8](#).

Detta innebär att klaganden måste komma fram med tillräckliga bevis för att ge upphov till en misstanke om diskriminering.

Därefter övergår bevisbördan till svaranden. I praktiken har nationella domstolar ett visst utrymme för bedömning när bevisningen som presenteras av klaganden är tillräcklig för att överföra bevisbördan till svaranden.⁴⁷

Bevis kan inkludera statistiska uppgifter,⁴⁸ skriftliga dokument – om de har innehåll som rör diskrimineringen eller rör det specifika fallet för den sökande – samt rapporter och analyser från relevanta källor (se data som samlats in i steg 2), situationstestning (se under steg 3.6), sunt förnuft-bedömningar, vittnesmål och slutsatser dragna från indicier. EU-domstolen betonar vikten av att beakta "alla omständigheter kring den ifrågavarande praxisen" (C-83/14 CHEZ, punkterna 80–84). Följaktligen kan både statistiska och kvalitativa data användas för att fastställa en misstanke om diskriminering och det finns argument för att båda faktiskt bör användas.

Vad är "jämlikhetsdata"?

I metodikens sammanhang hänvisar vi till både kvantitativa (inklusive statistiska) och kvalitativa data som "jämlikhetsdata", med utgångspunkt från följande definition i

46. Direktiv 2000/78/EG, artikel 10, direktiv 2006/54/EG, artikel 19 och direktiv 2000/43/EG, artikel 8.

47. I praktiken skiljer sig detta ofta åt mellan direkt och indirekt diskriminering. För närmare information, se Farkas och O'Farrell (2015).

48. För mer information om statistik som bevismedel, se nedan under "[Metod för statistiska uppgifter](#)".

”Riktlinjer för förbättrad insamling och användning av jämlikhetsdata” (Europeiska kommissionen 2021) som:

all information som är användbar i syfte att beskriva och analysera tillståndet för jämlikhet. Informationen kan vara kvantitativ eller kvalitativ till sin natur. Den kan inkludera aggregerade data som återspeglar ojämlikheter eller deras orsaker eller konsekvenser i samhällen. Ibland kan data som samlas in främst av andra skäl än jämlikhetsrelaterade syften användas för att producera jämlikhetsdata. (Europeiska kommissionen 2021: 8)

Jämlikhetsdata används bland annat för ”att tillhandahålla tillförlitliga bevis i administrativa eller rättsliga fall om diskriminering genom data som pekar på direkt eller indirekt diskriminering; och för att stödja opinionsbildning och medvetandehöjande arbete på området jämlikhet och likabehandling” (Europeiska kommissionen 2021: 9).

Ytterligare läsning

- ▶ ”Handbok om att identifiera och använda jämlikhetsdata i rättsligt arbete” (Ilieva 2024);
- ▶ ”Handbok för strategiska rättsprocesser” (Equinet 2017);
- ▶ ”Handbok: artificiell intelligens, juridiskt beslutsfattande och grundläggande rättigheter” (JuLIA Project 2024).

Hur steg 3 används

Avsnitten nedan innehåller vägledande frågor som kan hjälpa likabehandlingsorganen EB:er att bedöma om det finns ett *prima facie*-fall av diskriminering. Syftet är att visa var i den information som samlats in under steg 2 den relevanta bevisningen kan finnas, vad man bör leta efter i bevisningen och hur den kan användas för att styrka diskriminering.

Varje avsnitt i denna del följer samma grundläggande struktur.

- ▶ en förklaring av vad EB:et ska bedöma;
- ▶ frågor att överväga vid bedömningen av informationen;
- ▶ möjliga informationskällor;
- ▶ indikatorer för diskriminering.

Avsnitt 3.7 och 3.8 innehåller därefter en ram för att bedöma om den bevisning som har samlats in är tillräcklig för att styrka ett *prima facie*-fall av diskriminering som involverar ett AI-system eller ett algoritmiskt system ([3.7 Bedömning av om bevisningen är tillräcklig](#)) samt för att bedöma sannolikheten för att svaranden kan motbevisa misstanken eller lämna en rättfärdigande grund för den påstådda diskrimineringen ([3.8 Kan svaranden visa att skillnaden i behandling inte är diskriminerande?](#)).

Detta hjälper till att avgöra om några åtgärder bör vidtas mot svaranden. Det kommer ofrånkomligen att finnas en viss överlappning mellan dessa moment och vid granskningen av bevisningen.

Det är ytterst osannolikt att ett EB kommer att kunna få tillgång till all information eller besvara samtliga frågor i detta steg. Frågorna är i stället avsedda att fungera som en vägledning om vilka typer av information som kan finnas tillgängliga och hur de kan bedömas, både var för sig och sammantaget.

Ett förslag till dokumentationsblad finns i [bilaga K](#). Det är avsett som ett hjälpmedel och kan anpassas efter det enskilda EB:ets behov.

Mer information

Några av begreppen i steg 3 kan vara nya eller obekanta. Nedanstående resurser förklarar dem mer ingående.

- ▶ Wachter, Mittelstadt och Russell (2020), *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*. *Computer Law & Security Review*, volym 41, juli 2021.
- ▶ Mittelstadt (2024), *How to Use the Artificial Intelligence Act to Investigate Bias and Discrimination: A Guide for Equality Bodies*, *Equinet*.

3.1. Vilka är de kända eller misstänkta riskerna för diskriminering?

Användningen av algoritmiska system i vissa sammanhang kan medföra en inneboende risk för diskriminerande metoder. Om algoritmisk diskriminering tidigare har konstaterats i ett liknande användningsfall inom samma sektor kan detta tyda på en bredare risk för diskriminering som även är relevant för det aktuella ärendet. Om det exempelvis har visats att ansiktsgenkänning som används inom polisverksamhet är diskriminerande, är det sannolikt att ansiktsgenkänning som används för andra ändamål också kan vara diskriminerande.

3.1.1. Omnämns den aktuella typen av algoritmiskt system eller det verksamhetsområde inom vilket det används uttryckligen i AI-förordningen på grund av dess diskrimineringsrisker?

AI-förordningen hänvisar både till "sektorer" och "användningsfall" inom dessa sektorer.

När algoritmiska system används inom sektorer där historiska ojämlikheter förekommer och/eller där diskriminering har visat sig vara vanligt förekommande, bör detta ge upphov till en ökad misstanke om diskriminering. Algoritmiska system bygger på data från verkligheten. Befintliga ojämlikheter och mänskliga fördomar återspeglas i dessa data och kan därför också återfinnas i – och i vissa fall förstärkas av – det algoritmiska systemet.

Exempel på bias i historiska data: Amazons AI-baserade rekryteringsverktyg

Amazon utvecklade ett AI-baserat rekryteringsverktyg som tränades på historiska meritförteckningar (CV:n) från en grupp av sökande som till övervägande del bestod av män. Som en följd av detta "lärde sig systemet självt att manliga kandidater var att föredra". Även om systemet inte uttryckligen använde kön som indata, förlitade det sig på proxyvariabler som korrelerade med att vara kvinna – till exempel ordet "kvinnors" (såsom i "ordförande för kvinnors schackklubb") eller examen från kvinnouniversitet – och nedgraderade systematiskt ansökningar från kvinnliga kandidater (Dastin 2018).

Många av de högriskanvändningsfall som omfattas av AI-förordningen har klassificerats som högrisk just därför att de medför en risk för diskriminering. I skälen 54 och 56–60 i AI-förordningen nämns uttryckligen risken för diskriminering, förstärkning av historiska ojämlikheter och särskilda sårbarheter. Även om ett system inte är tekniskt komplext och/eller inte utgör ett AI-system, innebär användningen av algoritmiska system (regelbaserade system) för dessa användningsfall en liknande risk för diskriminering. [Bilaga I](#) beskriver de högriskområden och användningsfall för AI som omfattas av förordningen.

Ett användningsfall som uttryckligen nämns i AI-förordningen motiverar en fördjupad granskning. Dessutom kan det för sådana fall finnas kompletterande jämlikhetsdata, både kvalitativa och kvantitativa, som ofta återfinns i rättspraxis, akademisk forskning, rapporter från civilsamhällesorganisationer eller myndigheter. Se även avsnitt 2.1.2, Befintlig rättspraxis, och Övriga informationskällor.

Observera: Om systemet är ett AI-system med hög risk och är registrerat i EU-databasen, inklusive i avsnitt B,⁴⁹ innebär detta per definition att det uttryckligen omfattas av AI-förordningen och därför potentiellt medför en hög risk för diskriminering.

Frågor att överväga

- ▶ Är det algoritmiska systemet en "typ" av teknik som är känd för att medföra en risk för diskriminering, såsom profilering (artikel 6.3 och skäl 53), prediktiv riskanalys (skäl 42) eller biometrisk teknik och system för känslighetskänning (bilaga III punkt 1 samt skälen 44 och 54)?
- ▶ Var det beslut eller den behandling som personen utsattes för kopplad till någon av följande sektorer?
 - utbildning och yrkesutbildning (bilaga III punkt 3, skäl 56);

49. Om ett system finns registrerat i avsnitt B i EU-databasen har leverantören genom en egen bedömning ansett att systemet används inom en högrisksektor, men att det inte medför en hög risk på grund av ett begränsat antal undantag (se AI-förordningen, artikel 6.3). Även om den dokumentation som krävs enligt AI-förordningen sannolikt inte kommer att vara tillgänglig, används systemet inom ett område som uttryckligen anges i AI-förordningen på grund av dess risker för diskriminering..

- anställning, personalledning och tillgång till egenföretagande (bilaga III punkt 4, skäl 57);
 - tillgång till och åtnjutande av offentliga tjänster och förmåner (social välfärd) (bilaga III punkt 5 a);
 - kreditbedömning (bilaga III punkt 5 b, skäl 58);
 - liv- och sjukförsäkring (bilaga III punkt 5 c, skäl 58);
 - tillgång till och behörighet för hälso- och sjukvårdstjänster (bilaga III punkt 5 a, skäl 58);
 - brottsbekämpning (bilaga III punkt 6, skäl 59);
 - migration, asyl och gränskontroll (bilaga III punkt 7, skäl 60);
 - rättskipning (bilaga III punkt 8 a, skäl 61);
 - administration av demokratiska processer (bilaga III punkt 8 b, skäl 61).
- Om ärendet rörde någon av ovanstående sektorer, omfattar ärendet då ett användningsfall som anges i den relevanta delen av bilaga III (inklusive tillhörande skäl)?

Potentiella informationskällor för att bedöma en allmänt fastställd risk för diskriminering

EU:s AI-förordning

- Artikel 6, bilaga III och skälen – dessa beskriver typer av AI-system med hög risk;
- Artikel 5 – denna beskriver förbjudna AI-metoder samt skälen 28–45.

Övriga källor

- Kommissionens riktlinjer om förbjudna metoder inom AI, enligt definitionen i AI-förordningen.
- Kommande: Kommissionens riktlinjer om klassificering av AI-system med hög risk och relaterade krav.⁵⁰

Indikatorer på en risk för diskriminering

Om någon av dessa indikatorer föreligger utgör detta i sig inte bevis för diskriminering. Indikatorerna kan dock tyda på att ett system medför en risk för diskriminering:

- om systemet är ett AI-system med hög risk i avsnitt A eller C;
- för system som inte är högriskklassificerade:
 - om systemet är avsett att användas inom en sektor där det finns en identifierad risk för diskriminering (bilaga III eller skälen); och/eller
 - om det är en typ av teknik där det finns en känd risk för diskriminering.

50. Vid tidpunkten för redigeringen av denna metodbeskrivning har kommissionen offentliggjort ett utkast till riktlinjer för samråd: [Kommissionen efterfrågar synpunkter på utkastet till riktlinjer för klassificering av högrisk-AI-system.](#)

Dokumentera om systemet är ett högrisk-system eller inte, eller om det är avsett att användas inom någon av de sektorer eller områden som i AI-förordningen har identifierats som områden som medför en risk för diskriminering.

Misstänkt felaktig klassificering enligt AI-förordningen

Om ett likabehandlingsorgan misstänker att ett system felaktigt har bedömts av leverantören som ett AI-system med utan hög risk eller att det felaktigt inte har identifierats som en förbjuden AI-metod (se Misstänkta förbjudna AI-system under [steg 1.4](#)), bör organet informera den relevanta MSA:en, oavsett om ett fall av diskriminering kan fastställas eller inte (se även [steg 4](#)).

3.1.2. Har en känd risk för diskriminering inom denna sektor eller teknik identifierats av nationella eller regionala myndigheter, nationella eller internationella civilsamhällesorganisationer eller människorättsorganisationer?

Rapporter från ombudsinstitutioner, civilsamhället och akademiska aktörer samt befintlig rättspraxis kan visa på ett mönster av ojämlika konsekvenser för en viss skyddad grund.

Använd den information som samlats in i steg 2 för att bedöma om detta system eller denna typ av system redan är känt eller misstänks leda till diskriminerande resultat. All relevant information kan även bidra till genomförandet av de efterföljande stegen.

I detta steg är sektorspecifik expertis värdefull.

Potentiella informationskällor för att bedöma en allmänt fastställd risk för diskriminering

- ▶ likabehandlingsorganens (EB:er) övervakning, forskning eller utredningar;
- ▶ rapporter från civilsamhällesorganisationer (CSO:er), nationella och internationella organ;
- ▶ befintlig rättspraxis;
- ▶ sökningar på internet;
- ▶ experter.

Indikatorer på diskriminering

- ▶ Diskriminering har identifierats eller misstänks av andra vid den specifika användningen av algoritmiska system av den berörda svaranden som det aktuella ärendet omfattar.
- ▶ Diskriminering har identifierats eller misstänks av andra vid mycket liknande användningar av algoritmiska system som det aktuella ärendet omfattar, vilket tyder på ett möjligt mönster av diskriminering.

Dokumentera det område där det finns en känd risk, tekniken samt all annan relevant information och bevisning. Exempel:

Sektor: brottsbekämpning

Teknik: ansiktsgenkänning, biometriska uppgifter

Informationskällor: forskning utförd av en icke-statlig organisation (NGO)/ akademiska aktörer,

AI-förordningen artikel 5 (förbjudna system), artikel 6 (högriskklassificering) och bilaga III.

Användningsområden som bör betraktas som högrisk eller förbjudna enligt AI-förordningen

Om det identifierade mönstret av diskriminerande resultat för ett specifikt algoritmiskt användningsområde är tydligt och konsekvent, men användningsområdet inte finns med i förteckningarna i bilaga III, bör likabehandlingsorganen (EB:er) överväga att rapportera användningsområdet till Europeiska kommissionen så att det kan läggas till i förteckningen enligt artikel 7 (se [steg 4.3.2](#)).

Om likabehandlingsorganet konstaterar att en AI-metod i grunden strider mot unionens värden om respekt för människans värdighet, frihet, jämlikhet, demokrati och rättsstatsprincipen samt de grundläggande rättigheter som fastställs i stadgan, men att metoden varken är förbjuden eller betraktas som högrisk enligt AI-förordningen (se [steg 4.3.2](#)).

3.2. Finns det några kända skyddade egenskaper eller proxyvariabler i indata eller datauppsättningar?

Indata för ett algoritmiskt system avgör utdata och därmed resultatet. Om en skyddad egenskap (se [skyddade egenskaper](#)) eller en proxyvariabel (se [Proxy-diskriminering](#)) används som indata i det algoritmiska systemet bör det därför finnas en misstanke om diskriminering.

I detta steg är sektorsspecifik expertis användbar.

Frågor att beakta

- ▶ Behandlas några personuppgifter som avslöjar skyddade egenskaper eller proxyvariabler för skyddade egenskaper?
- ▶ Hänvisas det till skyddade egenskaper eller kända proxyvariabler (se [bilaga E](#)) i någon av de relevanta dokumenten?
- ▶ Nämns skyddade egenskaper eller kända proxyvariabler direkt som indata till systemet (antingen vid träning, testning eller användning)? Detta kan exempelvis uttryckas som: "X, Y och Z används för att förutsäga/beräkna".

- ▶ Är någon av systemets indata, eller en kombination av flera indata, sannolikt en proxyvariabel för en skyddad egenskap baserat på sektorsspecifik kunskap? För att besvara denna fråga kan det vara nödvändigt att konsultera nationell statistik eller sektorsexperter med kunskap om hur ojämlikheter uppstår och hänger samman inom en viss sektor. Exempel:
 - i länder där det finns skatteavdrag för sjukvårdsrelaterade kostnader är personer med funktionsnedsättning sannolikt överrepresenterade bland dem som utnyttjar detta avdrag. Användningen av detta avdrag är därför sannolikt en proxyvariabel för funktionsnedsättning;
 - i många länder är personer med migrationsbakgrund underrepresenterade bland personer med universitetsutbildning. Utbildningsnivå är därför sannolikt en proxyvariabel för ras/etnisk bakgrund.
- ▶ Är någon av proxyvariablerna ouppslösligt kopplad till en skyddad egenskap? Se [bilaga E – Proxyvariabler](#).
- ▶ Kan misstanken om att en viss indata (eller ett urval av indata) fungerar som en proxyvariabel styrkas genom bevisning? Exempelvis: finns det statistik från en nationell statistikmyndighet eller akademisk forskning som visar att en grupp med en skyddad egenskap är över- eller underrepresenterad i en viss kategori av indata? Exempel:
 - om utbildningsnivå är en av de indata som misstänks vara en proxyvariabel för etnisk bakgrund, finns det då statistik som visar att personer med migrationsbakgrund är överrepresenterade bland personer med yrkesutbildning?

Observera: att samla in denna bevisning kan kräva expertmedverkan från personer med erfarenhet av dataanalys.

- ▶ Är systemet tränat, testat eller validerat med information som innehåller skyddade egenskaper? (Se [steg 3.5](#).)

Potentiella källor till information för att förstå indata till systemet

Klagande

- ▶ Förteckning över information som lämnats av klaganden till svaranden;
- ▶ förklaringar, motiveringar och underlag som lämnats till klaganden om behandlingen.

GDPR, artiklarna 13 och 14: information som lämnats till den registrerade

- ▶ information som lämnats till den registrerade (klaganden) vid den tidpunkt då de lämnade uppgifterna (artikel 13 i GDPR);
- ▶ information som lämnats av en personuppgiftsansvarig där personuppgifterna inte erhöles direkt från den registrerade (artikel 14 i GDPR);
- ▶ den logik som tillämpas i fall av automatiserat beslutsfattande (artiklarna 13.2 f och 14.2 g).

GDPR, svar på DSAR-begäran

- ▶ kategorierna av berörda personuppgifter (artikel 15.1 b): den personuppgiftsansvarige måste lämna en kopia av de personuppgifter som behandlas. Detta kan inkludera uppgifter som lämnats av tredje part (artikel 14 i GDPR);
- ▶ potentiellt: indatavariabler eller parametrar som använts för beslutsfattande och hur de uppkommit (till exempel genom användning av statistisk information) och deras respektive vikt på aggregerad nivå (yttrande från generaladvokat Pikamäe i SCHUFA Holding, 2023, punkt 58).

EU:s AI-förordning, artikel 86: begäran om förklaring

- ▶ indata och de huvudsakliga elementen i beslutet.

Offentligt tillgängligt

- ▶ EU-databas för högrisk-AI-system: avsnitt A högrisk-AI-system – en beskrivning av den information som används av systemet (data, indata) och dess funktionslogik;
- ▶ FAQ på plattformen;
- ▶ offentliga företagsmeddelanden.

Svaranden

Från leverantören:

EU:s AI-förordning, artikel 11 och bilaga IV: teknisk dokumentation

- ▶ bilaga IV 2 b, den allmänna logiken i AI-systemet och i algoritmerna, motiveringen och de antaganden som gjorts, inklusive ifråga om personer eller grupper av personer som systemet är avsett att användas för; de huvudsakliga klassificeringsvalen; vad systemet är utformat att optimera för, och relevansen av de olika parametrarna; beskrivningen av den förväntade outputen och outputkvaliteten hos systemet;
- ▶ bilaga IV 2 d, en allmän beskrivning av datamängderna, information om deras ursprung, omfattning och huvudsakliga egenskaper.

Från användaren:

EU:s AI-förordning, artikel 27: konsekvensbedömning av grundläggande rättigheter

- ▶ kategorierna av fysiska personer och grupper som sannolikt påverkas av algoritmsystemet (artikel 27.1 c i EU:s AI-förordning).

Övrigt

- ▶ avtalsklausuler;
- ▶ skrivbordsforskning, befintlig rättspraxis och statistik om över- och underrepresentation inom det relevanta området.

Indikatorer på diskriminering

- ▶ direkt användning av skyddade egenskaper i systemet;
- ▶ användning av proxies som är oskiljaktigt kopplade till en skyddad egenskap i systemet.
- ▶ användning av kända proxies i systemet;
- ▶ användning av skyddade egenskaper eller proxies för att träna, testa eller validera systemet.

Alla ovanstående kräver bevis för att dessa indata leder till ojämn behandling. Indikatorerna a och b indikerar direkt diskriminering. Indikatorerna c och d indikerar indirekt diskriminering.

- ▶ Dokumentera direkt användning av skyddade egenskaper. Ange var de används och på vilket sätt.
- ▶ Dokumentera all användning av skyddade egenskaper eller proxyvariabler som har, eller som bedöms kunna ha, en påverkan på systemets resultat. Undersök systemets tekniska utformning och hur de (misstänkta) proxyvariablerna påverkar systemets resultat. Det faktum att en proxyvariabel används som indata är i sig inte tillräckligt som bevis för diskriminering. Nästa steg är att utvärdera bevisningen för att fastställa om användningen av dessa indata leder till olika behandling.⁵¹
- ▶ För varje proxyvariabel ska det anges vilken bevisning eller vilka argument som stöder varför EB (likabehandlingsorganet) anser att den utgör en proxyvariabel.

3.3. Bedömning av risker för grundläggande rättigheter (riskbedömningar, HRIA, FRIA och DPIA)

Leverantörer och användare av AI-system med hög risk har olika och varierande skyldigheter att identifiera och dokumentera risker för diskriminering. För algoritmiska system som inte är högrisk kan AI-leverantörer och/eller användare ändå ha identifierat och dokumenterat diskrimineringsrisker som en del av sin efterlevnad av GDPR (särskilt i en konsekvensbedömning av dataskydd, DPIA) eller inom ramen för tillbörlig aktsamhet för företag i fråga om mänskliga rättigheter (Förenta nationerna 2011). Nya skyldigheter kan också införas genom Europarådets ramkonvention om AI.

Risker för diskriminering identifieras vanligtvis i en konsekvensbedömning av grundläggande rättigheter (FRIA), en konsekvensbedömning av dataskydd (DPIA) och/eller andra riskhanteringsprocesser. Om dessa processer genomförs korrekt bör de identifiera och dokumentera eventuella diskrimineringsrisker som ett AI-system eller

51. En indata till systemet som indikerar en skyddad grund eller en känd proxyvariabel, men som inte sammanfaller med en skyddad grund, är i sig inte nödvändigtvis tillräcklig för att styrka ett diskrimineringsanspråk. Se exempelvis STAR-fallet i Danmark. I det fallet använde prediktionsverktyget "härkomst" eller "ursprung" som en prediktiv variabel. Nämnden konstaterade att "härkomst", såsom begreppet användes i profilverktyget, "inte säger något om etniskt ursprung". Se [www.retsinformation.dk/eli/accn/W20220951825](https://retsinformation.dk/eli/accn/W20220951825). För ytterligare information, se Equinet 2022: 23–24.

behandling av uppgifter med hjälp av ett AI-system medför, samt vilka riskreducerande åtgärder som har vidtagits och vilken effekt dessa har haft.

I teknisk dokumentation och riskbedömningar kan begreppen bias eller rättvisa (fairness) användas i stället för ojämlikhet eller diskriminering.

I detta steg skulle både teknisk expertis och sektorspecifik expertis vara värdefull.

Konsekvensbedömning av grundläggande rättigheter: AI-förordningen

Kravet att genomföra en FRIA gäller endast vissa användare av högrisk-AI-system (se [steg 2.2](#) för mer detaljer).

FRIA:n (och andra bedömningar) bör innehålla en rad uppgifter som bör belysa potentiella luckor och kränkningar av rättigheter, inklusive likabehandling. Detta inkluderar:

- ▶ information om hur användaren kommer att använda högrisk-AI-systemet enligt leverantörens avsedda syfte;
- ▶ hur länge och hur ofta systemet kommer att användas;
- ▶ hur åtgärder för mänsklig tillsyn har genomförts;
- ▶ kategorierna av fysiska personer och grupper som sannolikt kommer att påverkas;
- ▶ de specifika riskerna för kränkningar av grundläggande rättigheter, inklusive likabehandling med beaktande av information som tillhandahållits av leverantören;
- ▶ arrangemangen för intern styrning och klagomålsmekanismer som användaren kommer att använda om risken materialiseras.

Om användaren borde ha genomfört en FRIA men inte har gjort det eller inte inkluderat relevant information, se [steg 4](#).

DPIA

En DPIA⁵² bör beakta alla risker för fysiska personers rättigheter och friheter till följd av behandling av personuppgifter, inte bara integritetsrisker. Detta inkluderar även diskrimineringsrisker (artikel 29 Data Protection Working Party, 2017a). EB:et kan behöva samarbeta med dataskyddsmyndigheterna.

Frågor att beakta

Information om de olika riskbedömningar som utförts av leverantören och användaren bör kors refereras med varandra och eventuella avvikelser noteras.

- ▶ Har leverantören eller användaren identifierat relevanta risker för grundläggande rättigheter som den andra parten inte har?
- ▶ Där en part har genomfört mer än en riskbedömning eller mer än en typ av riskbedömning – finns det konsekvens i fråga om identifierade risker? Saknas några risker som identifierats i en bedömning i en annan (om inte det finns

52. Europeiska dataskyddsstyrelsen höll ett offentligt samråd i maj 2026 om en DPIA-mall, tillgänglig här: www.edpb.europa.eu/our-work-tools/documents/public-consultations/2026/edpb-dpia-template_en.

en god förklaring; till exempel uppstår risken endast i frågan om ett barn och beaktas specifikt i barnriskbedömningen)?

För alla algoritmsystem, inklusive högrisk-AI:

- ▶ Har diskriminering (eller annan risk för grundläggande rättigheter) identifierats av leverantören eller användaren?
- ▶ Har leverantören infört ett riskhanteringssystem?
- ▶ Hur användes riskhanteringsförfaranden och konsekvensbedömningar för att fastställa bias eller ojämlikhet i data eller systemet? Hur definieras eller avgränsas bias?
- ▶ Finns det några risker som inte har hanterats? Kontrollera ifall möjligt risker som hänvisas till i den tekniska dokumentationen mot riskhanteringssystemet.

Observera: där diskriminering har identifierats och påstådda förvaltningsåtgärder har vidtagits kräver detta ytterligare utredning (se steg 3.6, [Testning och utvärdering](#)). Om systemet är högrisk och EB:et är ett artikel 77-organ kan de begära tillgång till dokumentationen för riskbedömningen enligt artikel 9.

I fall som rör ett högrisk-AI-system

- ▶ Är något riskhanteringsförfarande förenligt med kraven i artikel 9?

I fall som rör ett högrisk-AI-system, där en FRIA krävs från leverantören:

- ▶ Har FRIA:n genomförts korrekt? Den bör inte vara en ren avbockningsövning;
- ▶ Har leverantören involverat några relevanta intressenter vid genomförandet av riskbedömningen? (Skäl 96 ger möjlighet till detta);
- ▶ Borde intressenter ha konsulterats, till exempel personer som sannolikt påverkas av systemet? (se Bridges-målet nedan);
- ▶ Om en konsultationsprocess ägt rum, överväg att kontakta de konsulterade personerna men var medveten om att de kan vara bundna av sekretessplikt eller sekretessavtal;
- ▶ Kontrollera att alla risker för diskriminering som identifierats i den tekniska dokumentationen (artikel 11) finns hänvisade i riskhanteringssystemets dokumentation (artikel 9). Om dessa inte ingår kan en misstanke om diskriminering uppstå.

Potentiella källor till information om risker och riskbedömningar

Klagande

GDPR, artiklarna 13–15, Information som lämnas till den registrerade

- ▶ Förekomsten av automatiserat beslutsfattande och/eller profilering (artiklarna 13–15, 21 och 22 i GDPR).
- ▶ Den logik som tillämpas i det ADM som används (artiklarna 13.2 f, 14.2 g och 15.1 h i GDPR).

Offentligt tillgängligt

- EU-databas för högrisk-AI-system: avsnitt C användare: sammanfattning av DPIA och FRIA (om inte brottsbekämpning eller migrationskontroll).

Svaranden

Från leverantören:

EU:s AI-förordning, artikel 9, riskhanteringssystem

- Leverantörer är skyldiga att inrätta ett riskhanteringssystem för högrisk-AI-system. Det ska vara en kontinuerlig iterativ process och kräver:
 - riskidentifiering: identifiering och analys av kända och rimligen förutsebara risker för grundläggande rättigheter (artikel 9.2 a);
 - riskhantering: antagande av lämpliga och riktade riskhanteringsåtgärder (artikel 9.2 d), inklusive tillhandahållande av information och utbildning till användare för att eliminera eller minska risken (artikel 9.5 c och artikel 13);
 - testning: säkerställa att systemet har testats i syfte att identifiera de mest lämpliga och riktade riskhanteringsåtgärder (artikel 9.6);
 - särskilda grupper: vid genomförandet av systemet ska hänsyn tas till om systemet sannolikt kommer att ha en negativ inverkan på personer under 18 år och, i förekommande fall, andra sårbara grupper (artikel 9.9).

Riskhantering krävs dock endast beträffande de risker som "rimligen kan begränsas eller elimineras genom utveckling eller utformning av högrisk-AI-systemet, eller genom tillhandahållande av adekvat teknisk information".

EU:s AI-förordning, artikel 10, data och dataförvaltning (se steg [3.5 Tränings-, test- och valideringsdata](#))

EU:s AI-förordning, artikel 11 och bilaga IV, teknisk dokumentation

- Bilaga IV 2 g kräver att den tekniska dokumentationen innehåller information om eventuella potentiellt diskriminerande konsekvenser.
- Bilaga IV 3: detaljerad information om övervakning, funktion och kontroll av AI-systemet, i synnerhet gällande dess förmågor och begränsningar i prestanda, inklusive graden av noggrannhet för specifika personer eller grupper av personer som systemet är avsett att användas för och den övergripande förväntade noggrannhetsnivån i förhållande till sitt avsedda syfte; de förutsebara oavsiktliga utfallen och källor till risker för hälsa och säkerhet, grundläggande rättigheter och diskriminering med beaktande av det avsedda syftet med AI-systemet; de mänskliga tillsynsåtgärder som behövs i enlighet med artikel 14, inklusive de tekniska åtgärder som vidtagits för att underlätta driftsättarnas tolkning av AI-systemens utdata; specifikationer för indata, i förekommande fall.

Från användaren:

EU:s AI-förordning, artikel 13, transparens och tillhandahållande av information till användare

- ▶ Bruksanvisningar som tillhandahållits användare av leverantören (artikel 13.3):
 - kända eller förutsebara omständigheter för avsett bruk eller missbruk som kan innebära risk för grundläggande rättigheter;
 - när det är lämpligt, dess prestanda gällande specifika personer eller grupper av personer som systemet är avsett att användas för;
 - när det är lämpligt, specifikationer för indata, eller annan relevant information om de tränings-, validerings- och testdatamängder som använts, med beaktande av det avsedda syftet med högrisk-AI-systemet.

EU:s AI-förordning, artikel 26.9, dataskyddskonsekvensbedömning: utförd med hjälp av den information som tillhandahållits av leverantören/GDPR, artikel 35, dataskyddskonsekvensbedömning (DPIA)

EU:s AI-förordning, artikel 27.1 d, bedömning av konsekvenserna på grundläggande rättigheter

- ▶ De specifika riskerna för skada på grundläggande rättigheter, inklusive likabehandling med beaktande av information som tillhandahållits av leverantören.

Övrigt

- ▶ Bedömning som genomförts med hjälp av Europarådets risk- och konsekvensbedömning av AI-system ur perspektivet mänskliga rättigheter, demokrati och rättsstat (HUDERIA-metodiken);
 - ▶ Vissa organisationer genomför och publicerar frivilligt en HRIA eller FRIA, DPIA eller bedömning av barns rättigheter. Det kan finnas en "etikbedömning";
 - ▶ Eventuella andra (risk)bedömningar som delas av svaranden. Begäran om allmänna handlingar där detta är möjligt.
-

Viktigt rättsfall: *R (Bridges) v. Chief Constable of South Wales Police, 2020*

(Siffror inom hakparenteser anger paragrafnummer i domen.)

Bridges var ett mål som avgjordes av appellationsdomstolen i Storbritannien och som prövade om en offentlig myndighet hade brutit mot mänskliga rättigheter eller underlåtit att uppfylla sina skyldigheter att bedöma konsekvenser för lika-behandling enligt den inhemska Equality Act 2010 innan live-ansiktsgenkänning infördes. Den offentliga sektorns jämlikhetsskyldighet kräver att en offentlig myndighet ägnar "vederbörlig uppmärksamhet" åt behovet av att eliminera diskriminering, trakasserier, rpressalier, offerstatus och annat relevant beteende, främja jämlikhet i möjligheter och främja goda relationer mellan personer som delar en relevant egenskap och personer som inte gör det.

Domstolen fastställde att skyldigheten:

- ▶ måste uppfyllas innan och vid den tidpunkt då en policy (att använda ansiktsgenkänning) övervägs;
- ▶ måste utövas i sak, med noggrannhet och ett öppet sinne – det är inte en ren avbockningsövning;
- ▶ är icke-delegerbar, vilket kan skapa mycket verkliga svårigheter för offentliga användare som förlitar sig på tredje parts leverantörer, särskilt om otillräcklig information om ett systems prestanda tillhandahållits eller om myndigheten inte hade tillräcklig expertis eller erfarenhet för att förstå informationen eller systemet [199].

Dessutom, om material som är relevant för bedömningen inte är tillgängligt finns det en skyldighet att skaffa det – detta kan kräva ytterligare konsultation med relevanta grupper. Slutligen, förutsatt att det

finns rigorös beaktande av skyldigheten så att det finns en riktig uppskattning av det potentiella inflytandet av beslutet på jämlikhetsmålen och önskvärdheten av att främja dem är det upp till beslutsfattaren att avgöra hur mycket vikt som ska läggas vid faktorer som informerar beslutet.

Sammanfattningsvis fastställde domstolen att efterlevnad av skyldigheten "kräver att rimliga steg vidtas för att inhämta upplysningar om vad som ännu inte är känt för en offentlig myndighet om den potentiella konsekvensen av ett föreslaget beslut eller en policy på personer med de relevanta egenskaperna, i synnerhet i detta fall ras och kön" [181].

Att ha en människa i loopen ansågs i sig osannolikt kunna uppfylla den lagstadgade skyldigheten eftersom skyldigheten reglerar processen och inte substansen i ett beslut.

Domstolen konstaterade också att människor gör misstag [185].

Liknande iakttagelser kommer sannolikt att gälla för riskbedömningar enligt AI-förordningen.

Indikatorer på diskriminering

En risk för diskriminering som har identifierats men inte begränsats, övervakats eller motiverats bör betraktas som en stark indikator.

- ▶ Direkt hänvisning till diskriminering, bias eller orättvisa i FRIA.
 - ▶ Direkt hänvisning till diskriminering, bias eller orättvisa i DPIA.
 - ▶ I fall där bias nämns som en risk definieras inte betydelsen och typerna av bias eller definieras på ett olämpligt sätt i dokumentationen.
 - ▶ Inga eller otillräckliga riskhanteringsförfaranden på plats för identifierade risker för diskriminering, bias eller orättvisa.
 - ▶ Obehandlade risker för diskriminering syns i dokumentationen.
 - ▶ Otillräcklig testning.
 - ▶ Direkt hänvisning till en risk för brott mot grundläggande rättigheter i antingen FRIA eller DPIA:
 - där det överhuvudtaget finns en risk för brott mot grundläggande rättigheter bör detta flagga ett krav på ytterligare utredning för att säkerställa att systemet är icke-diskriminerande.
 - ▶ En ofullständig, ytlig eller felaktig FRIA eller DPIA motiverar ytterligare utredning.
 - ▶ Underlåtenhet att konsultera relevanta intressenter för att identifiera risker där de tydligt borde ha inkluderats (till exempel ett system som är utformat för att fatta beslut om personer med funktionsnedsättningar och som inte konsulterade några personer med funktionsnedsättningar eller representativa organisationer innan driftsättning kan ha underlåtit att genomföra FRIA med den omsorg som krävs).
- ▶ Anteckna eventuella risker för diskriminering som identifierats av leverantören eller användaren och eventuella åtgärder som identifierats för att begränsa den risken.
 - ▶ Beakta och anteckna om det finns några tydliga och uppenbara brister i riskbedömningen, DPIA eller FRIA.

3.4. Vad är algoritmsystemet avsett att göra och hur fungerar det?

För att förstå om algoritmisk diskriminering har förekommit bör EB förstå systemet självt, åtminstone i allmänna termer. Att förstå systemets syfte och utformning och hur det fungerar hjälper till att utvärdera följande.

- ▶ Vad är syftet med systemet?
 - Är syftet relaterat till ett område där det har förekommit historisk diskriminering beträffande den/de skyddade egenskapen eller egenskaperna?
 - Är de data som används lämpliga för syftet (se [steg 3.5](#))?
- ▶ Hur bidrar indata till utdata?

- Påverkar en indata som identifierats i steg 3.2 utdatan och leder till en potentiell skillnad i behandling?
- ▶ Huruvida algoritmsystemet påverkar antingen hela beslutsprocessen eller den slutliga konsekvensen för den drabbade gruppen:
 - Vilken storlek har systemets roll? Till exempel, är det den enda beslutsfattaren eller bistår det beslutsfattare?
 - Är algoritmsystemet hela systemet eller bara en del av ett mycket större system som kan involvera andra algoritmsystem?

I informationen om avsett bruk och utformning bör du generellt leta efter om designvalen verkar lämpliga för det avsedda syftet och den faktiska användningen. I detta steg är både teknisk expertis och sektorexpertis användbar.

Den information som är tillgänglig för EB:erna bör i slutändan möjliggöra för EB:et att förstå vad AI-systemet gör, hur det gör det genom meningsfull information om den logik som tillämpas, betydelsen av behandlingen och de planerade konsekvenserna för den enskilde samt varför olika val i fråga om utformning och data gjordes. När algoritmsystemet bedöms vara kvalificerat ADM eller ett högrisk-AI-system som används för beslut med stor påverkan är tillgång till denna information obligatorisk. Om dokumentationen dock inte möjliggör denna förståelse kan det finnas frågor som rör brist på transparens (se nedan, Transparens, under steg 3.7).

Utformningen av systemet kan också vara särskilt viktig senare när man tittar på sätt på vilka svaranden kan försöka bemöta en presumtion om diskriminering eller motivera fattade beslut.

Frågor att beakta

Syfte och användning

- ▶ Hur är algoritmsystemet avsett att användas?
 - Vad är det avsedda syftet?
 - Hur används det i praktiken?
 - I fall där systemet tillhandahålls av en leverantör: används det av användaren enligt leverantörens avsedda syfte?
- ▶ För vem är systemet utformat?
 - Vilka designval (se rutan nedan) gjordes beträffande de personer eller grupper av personer som systemet är avsett att användas för?
- ▶ Finns det några antaganden om personer eller grupper av personer som systemet är avsett att användas för? Till exempel att detta system förväntas användas endast av vuxna, eller att detta system antas användas inom en specifik sektor eller för en specifik population. Är dessa antaganden rationella och motiverade?
- ▶ Hur påverkar systemet behandlingen av personer?
- ▶ I fall där det finns en mänsklig beslutsfattare som är ansvarig för att fatta beslut som påverkar behandlingen av personer (human-in-the-loop):
 - Finns det några riktlinjer eller protokoll som utfärdats till beslutsfattare för att hjälpa dem att förstå hur man använder AI-systemet eller tolkar utdatan?

- Har den relevanta personen den nödvändiga kompetensen, utbildningen och befogenheten att åsidosätta ett beslut som fattats av ett AI-system?
- Har den relevanta personen utbildats för att förstå och tolka systemets utdata och känna igen var bias kan uppstå och ingripande kan krävas?

Observera: enbart förekomsten av en human-in-the-loop innebär inte att systemet inte påverkar behandlingen av personer, eftersom human-in-the-loop i större eller mindre utsträckning påverkas av systemets utdata.

Hur systemet fungerar

- ▶ Vad är systemet utformat att optimera för?
 - Var det utformat för att optimera för något (lite) annat än sitt avsedda användningsfall? Till exempel, var det optimerat för att förutsäga fel, men används nu för att förutsäga bedrägerier?
- ▶ Använder systemet gruppegenskaper för att producera utdata om en individ (profilering)?
- ▶ Hur hanterar systemet avvikelser?
 - Hur hanterar systemet avvikelser från förväntat beteende (fel eller ovanlig aktivitet till exempel)?
 - Behandlas alla avvikelser på samma sätt, oavsett den underliggande orsaken?
- ▶ Var systemet utformat för rättvisa och likabehandling?
 - Var rättvisa eller likabehandling en del av utformningen?
 - Är den kontextspecifika uppfattningen av rättvisa uttryckligen definierad?
 - Vilka steg vidtogs för att utforma för rättvisa och likabehandling?

Indatas påverkan

- ▶ Hur gjordes valet av någon (förtränad) modell eller algoritm?
 - Är detta förfarande dokumenterat och tillgängligt för användare av systemet?
- ▶ Vilka kriterier påverkar systemets utdata?
 - Finns relevanta skyddade egenskaper och/eller proxyvariabler i de data som behandlas eller som ligger till grund för ett ADM eller profilering (såsom indata för att skapa en profil, information om profilen)?
 - Finns det mätetal tillgängliga om vikten/påverkan av specifika indata på utdatan?
- ▶ Vilka kriterier ges störst vikt i systemet eller beslutet?
 - Kan svaranden argumentera för att skyddade egenskaper och/eller proxies har en försumbar påverkan på utdatan?
- ▶ Vad är den kumulativa konsekvensen av flera proxyvariabler tillsammans, såsom faktorer som postnummer + utbildning?
- ▶ Är de data som används eller ges störst vikt godtyckliga eller irrelevanta för det syfte de används för?

Potentiella källor till information för att förstå systemets syfte, utformning och hur det fungerar

Klagande

GDPR, artiklarna 13–15, information som lämnas till den registrerade

- ▶ Förekomsten av automatiserat beslutsfattande och/eller profilering (artiklarna 13–15, 21 och 22 i GDPR).
- ▶ Den logik som tillämpas i det ADM som används (artiklarna 13.2 f, 14.2 g och 15.1 h i GDPR).

GDPR, artikel 15.3: DSAR

- ▶ Omfattningen, arten och detaljnivån på den information som lämnas i svar på en DSAR-begäran kan variera. I fall av kvalificerat ADM krävs följande, men för andra algoritmsystem kan (delar av) denna information också inkluderas:
 - indatavariabler eller parametrar som används för beslutsfattande och hur de uppkommit (såsom genom användning av statistisk information) och deras respektive vikt (yttrande från generaladvokat Pikamäe i SCHUFA Holding, 2023, punkt 58);
 - effekten av en indatavariabel på poängen; förklaring till varför den registrerade tilldelades en viss poäng; förteckning över möjliga profilkategorier och kontrafaktiska scenarier;
 - den underliggande beräkningsprincipen: detaljerade förklaringar av metoden för att beräkna poängvärdet; tilldelning av den registrerade till ett utvärderingsresultat, förteckning över profilkategorier (Barros Vale och Zanfir-Fortuna 2022);
 - förekomsten av automatiserat beslutsfattande och/eller profilering och meningsfull information om den logik som tillämpas samt betydelsen och de planerade konsekvenserna av sådan behandling för den registrerade (artikel 15.1 h).

EU:s AI-förordning, artikel 86, AI-förordningen: rätt till förklaring

- ▶ AI-systemets roll i beslutsprocessen. Till exempel: "AI-systemet [namn/beteckning på AI-systemet] användes för [beskrivning av den specifika funktionen hos AI-systemet, såsom att analysera data, utvärdera kriterier, utföra röstanalys för igenkänning av känslor m.m.]. AI-systemet bidrog till beslutsprocessen genom [att analysera data och tillhandahålla en bedömning av de relevanta parametrarna, till exempel]".
- ▶ Behandlingslogiken: "Systemet analyserade indata med hjälp av [beskrivning av algoritmer eller modeller, såsom beslutsträd] för att utföra [beskrivning av målet, såsom riskbedömning]"
- ▶ Utdata: "Baserat på denna analys genererade AI-systemet följande utdata, som införlivades i det slutliga beslutet: [beskrivning av utdatan, en poäng, en klassificering, till exempel]".
- ▶ Konsekvensen av det fattade beslutet på den drabbade personen: "Beslutet har följande potentiella konsekvens för dig: [konsekvens gällande kreditvärdighet,

tillgång till vissa tjänster, ekonomisk konsekvens m.m.]. Dessa konsekvenser är relevanta eftersom de kan påverka din förmåga att [förklara konsekvenserna, såsom deltagande i program eller mottagande av erbjudanden om finansiering]”.

- Viktning av kriterier: ”De viktigaste påverkande kriterierna viktades enligt följande: [kort förklaring av varje kriteriums betydelse]”.

Svaranden

Från leverantören

EU:s AI-förordning, avsnitt A 5 och 6 i bilaga VIII om högrisk-AI-system

- Beskrivning av det avsedda syftet med AI-systemet, grundläggande och kortfattad beskrivning av den information som används av systemet (data, indata) och dess funktionslogik.
- För brottsbekämpning/migrationskontroll kan ytterligare information i databasen vara tillgänglig för den relevanta MSA.

EU:s AI-förordning, artikel 11 och bilaga IV, teknisk dokumentation

- 1 a Det avsedda syftet.
- 1 d Beskrivningen av alla former i vilka AI-systemet släpps ut på marknaden eller tas i bruk (såsom mjukvarupaket inbäddade i hårdvara, nedladdningar eller API:er).
- 1 h En grundläggande beskrivning av det användargränssnitt som tillhandahålls användaren.
- 2 b Designspecifikationer och allmän logik. Se ovan, 3.2: Finns det några kända skyddade egenskaper eller proxies i indata eller datamängder?

Från användaren

Artikel 13, Transparens och tillhandahållande av information till användare

- Leverantören måste tillhandahålla (artikel 13.2) bruksanvisningar till användaren;
- Artikel 13.3 b IV, när det är lämpligt, de tekniska förmågorna och egenskaperna hos högrisk-AI-systemet för att tillhandahålla information som är relevant för att förklara dess utdata, och 3 b v, dess prestanda beträffande specifika personer eller grupper av personer som systemet är avsett att användas för.;
- Artikel 13.3 c, de ändringar av högrisk-AI-systemet och dess prestanda som har förutbestämts av leverantören vid tidpunkten för den initiala bedömningen av överensstämmelse, om sådana finns.

Övrigt

- Dokumentstudie
- Avtalsklausuler
- FAQ på plattformen
- Offentliga företagsmeddelanden

Indikatorer på diskriminering

Starka indikatorer

Dessa indikatorer tillsammans kan stödja ett *prima facie*-fall.

- ▶ Misstänkta kriterier (som identifierats i [steg 3.2](#)) har en inte obetydlig påverkan på systemets utdata:
 - när en skyddad egenskap eller en proxy som är oskiljaktigt kopplad till en skyddad egenskap har en inte obetydlig påverkan på systemets utdata är detta en tydlig indikation på direkt diskriminering;
 - när en (misstänkt) proxy har en inte obetydlig påverkan på systemets utdata är detta en indikation på indirekt diskriminering.
- ▶ Systemets utdata (åtminstone delvis) avgör behandlingen av individer.

Andra indikatorer

- ▶ Avsett bruk, antaganden och/eller designval misstänks återspegla eller förstärka historiska ojämlikheter, vilket leder till ojämna behandling;
- ▶ designval gynnar en grupp människor framför en annan;
- ▶ systemet optimerar för historiskt gynnade grupper;
- ▶ systemet används inte för det avsedda syftet det utvecklades/tillhandahölls för.

- ▶ Anteckna systemets syfte och vad det optimerar för.
- ▶ Anteckna proxies och skyddade egenskaper som påverkar systemets utdata och, när sådana finns tillgängliga, mätetal om deras vikt/påverkan på utdatan.

3.5. Tränings-, test- och valideringsdata

Data används för att utveckla och testa algoritmsystem. Även regelbaserade algoritmsystem som utformas manuellt valideras med hjälp av data. Data är alltid en ofullständig representation av verkligheten. Bias och befintliga ojämlikheter återspeglas i datan. Bedömning av bias och ojämlikheter i datan kan användas för att argumentera för att det resulterande systemet leder till diskriminerande utfall. Observera att förståelse för bias i datan i sig inte visar ojämlik behandling. Tillsammans med andra kompletterande analyskällor bidrar det dock till ett *prima facie*-fall och kan användas för att motivera förfrågningar om ytterligare testning.

När statistiska jämlighetsdata finns tillgängliga är detta avsnitt dessutom relevant för att beakta validiteten och tillförlitligheten hos de data som används för att beräkna dessa statistikuppgifter. I så fall bör detta avsnitt läsas i samband med [avsnitt 3.6](#).

I datan bör du generellt leta efter vilken information svaranden har lämnat om hur de har tränat och testat sitt system. Du kan möjligen identifiera eventuella viktiga frågor kring diskriminering utifrån detta; till exempel att inga kvinnor inkluderades i datamängden, att träningsdatan kom från en stad men systemet används på landsbygden, osv.

I andra fall kan du behöva använda teknisk och sektorexpertis för att identifiera om systemet är tränat och testat på representativa, relevanta, korrekt och lämpligt mätta samt generaliserbara datamängder.

Några viktiga frågor att beakta inkluderar en rad olika bias. Dessa bias är inte oberoende av varandra: en bias kommer att påverka en annan.

- ▶ Huruvida det är troligt att det finns historisk bias i datan. I alla områden där det finns en historisk eller aktuell ojämlikhet återspeglas denna ojämlikhet i datan, om den inte har korrigerats mycket noggrant. Exempel är tidigare överdriven polisbevakning av svarta samhällen och perioder av arbetslöshet på grund av barnomsorg eller annat omsorgsansvar som påverkar jämställdheten mellan könen.
- ▶ Huruvida det finns representativ bias i datan. Till exempel, speglar målpopulationen användarpopulationen? Till exempel kan data som är representativa för Lissabon inte vara representativa för Porto. Data som är representativa för Helsingfors för 30 år sedan behöver inte spegla dagens befolkning. Det kan finnas ytterligare problem som skapas av vilka som är representerade i datan och i vilken utsträckning. Till exempel kan det i hälsodata finnas en underrepresentation av kvinnor och/eller en överrepresentation av gravida kvinnor. Något av detta kan snedvrider datan och påverka utfallen. Var dessutom uppmärksam på hur minoriteter är representerade i datan. Till exempel, när vissa accenter inte har inkluderats i ett röstigenkännings-system kommer detta system att fungera sämre för dessa accenter. Om personer som bär hijab inte har inkluderats i datan för ansiktsgenkänning finns det möjlighet att personer med hijab inte känns igen väl.
- ▶ Huruvida det finns ett urvals-bias i datan. Detta är fallet när det bara finns data för en specifik undergrupp av datan. Till exempel, i lånebeviljande finns det bara information om vilka som misslyckats med att betala tillbaka sina lån för den grupp som beviljades lån tidigare. Polisens patrulldata inkluderar endast information om de områden där polisen faktiskt patrullerade. Bedrägeridata kan endast inkludera den grupp som utreddes för bedrägeri tidigare. Rekryteringsdata innehåller mindre information om sökande vars CV aldrig valdes ut för vidare steg. Detta kan snedvrider datan, särskilt om personer med en specifik skyddad egenskap historiskt sett var över- eller underrepresenterade i detta sammanhang.
- ▶ Huruvida det finns ett mät-bias i datan. Detta är fallet när etiketterna (mätningen) inte återspeglar det begrepp de är avsedda att mäta. Till exempel, om du vill mäta bedrägeri i välfärdsförmåner men beslutar att alla felaktiga utbetalningar av förmåner är relevanta att beakta, kommer resultatet att bli att du inkluderar oavsiktliga misstag såväl som bedrägeri. Detta kan i sin tur ha en oproportionerlig konsekvens för vissa grupper, till exempel de med låg läskunnighet, låg utbildningsnivå och de som inte har ett nationellt språk som modersmål. En annan form av mät-bias är när mänskliga fördomar (eller mänsklig diskriminering) påverkar de registrerade etiketterna som nu betraktas som grundläggande sanning. I alla situationer där etiketterna är resultatet av en mänsklig utvärdering med viss individuell diskretion finns det en risk för att mänskliga fördomar resulterar i ett mät-bias. Exempel är lärarutvärderingar, rekryteringsdata eller bedrägeriutredningar.

Frågor att beakta

Beakta nedanstående frågor med bias och ojämlikhet i åtanke. Alla frågor nedan är subjektiva, nyanserade och kontextspecifika. Att svara på dessa kräver en god förståelse för sektorn och dess befintliga ojämlikheter samt för systemen och den teknik som används. Resultatet av denna bedömning kommer mestadels att ge riktning för ytterligare bedömning eller kan användas för att bygga ett *prima facie*-fall.

Observera att det i fallet med manuellt utformade algoritmsystem inte kommer att finnas någon träningsdata. I vissa AI-system innehåller inte träningsdatan etiketter. I så fall finns det fortfarande en förväntan på testdata, som innehåller någon form av etiketter för att testa om modellen ger korrekta eller acceptabla utdata.

Tränings-, test- och valideringsdata

- ▶ Hur valdes tränings-, test- och valideringsdatamängderna ut?
- ▶ Vilka tränings-/test-/valideringsdata samlades in?
- ▶ Vilka är källorna till dessa data?
- ▶ För vilket syfte samlades datan ursprungligen in? Uppstår några potentiella problem av att använda dem i ett annat sammanhang? Till exempel, om data samlades in i syfte att bedöma hälsa, väcker återanvändning av dem för försäkringssyften några potentiella mät-bias-problem?
- ▶ Vilka metoder användes för att samla in datan?
 - Verkar dessa metoder rimliga, i den meningen att de rimligen resulterar i tillförlitliga, relevanta data för algoritmsystemets avsedda uppgift?
 - Säkerställde de att datan var representativ och lämpligt mätt? Kan det rimligen förväntas att datan generaliserar till det område och det sammanhang där algoritmsystemen nu tillämpas?
 - Samlades ytterligare data in (eller etiketterades) för att användas som tränings-/test-/valideringsdata? Användes ett slumpmässigt urval för att samla in data?
- ▶ Där datamängden kommer från en tredje part, hur anskaffades den? Granskades den tredje partens data?
- ▶ Finns det historiska ojämlikheter som sannolikt återspeglas i datan (se avsnittet om bias ovan)? Beakta även flera/intersektionell diskriminering här.
- ▶ Är minoriteter tillräckligt representerade i datan? Detta är särskilt relevant för algoritmsystem som tillämpas i sektorer där kultur, bakgrund och fysiskt utseende kan påverka utdata.

Etiketter

- ▶ Hur tillförlitliga är etiketterna?
- ▶ Stämmer etiketterna med systemets syfte? Med andra ord, mäter etiketterna det som systemet är avsett att förutsäga?
- ▶ Kan korrekthet avgöras objektivt? Har detta avgjorts?
- ▶ I fall där mänskligt omdöme är involverat i etiketteringen, kan mänskliga fördomar ha snedvridit etiketterna? Finns det några skäl att misstänka diskriminering?

Befintliga begränsningsåtgärder

- ▶ Hur bedömdes datakvaliteten?
 - Vilka steg vidtogs för att åtgärda eventuella kvalitetsproblem som upptäcktes, såsom att komplettera eller ta bort data?
- ▶ Vilka åtgärder vidtogs för att säkerställa att datan var representativ, relevant, korrekt mätt och generaliserbar?
- ▶ Hur begränsades eventuell bias och ojämlikhet i datamängden?
 - Implementerades några bias-begränsande åtgärder⁵³
 - För vilka bias implementerades dessa?
 - Finns det några erkända biaser för vilka inga begränsande åtgärder implementerades?
 - Vilka tekniker användes?
 - Finns information tillgänglig om effektiviteten hos dessa bias-begränsande åtgärder?

Potentiella källor till information för att bedöma det data som används

Klagande

EU:s AI-förordning, artikel 86, rätt till förklaring

GDPR, artiklarna 13–15, information som lämnas till den registrerade

Offentligt tillgängligt

EU-databas för högrisk-AI-system

- ▶ Avsnitt A, högrisk-AI-system – information som används av systemet, inklusive data och indata.
- ▶ Avsnitt C, användare – sammanfattning av DPIA och FRIA (om inte använt för brottsbekämpning eller migrationskontroll).

Svaranden

Från leverantören:

EU:s AI-förordning, bilaga IV 2 d, teknisk dokumentation

- ▶ "Där det är relevant, datakraven i form av datablad som beskriver träningsmetoderna och teknikerna och de träningsdatamängder som använts, inklusive en allmän beskrivning av dessa datamängder, information om deras ursprung, omfattning och huvudsakliga egenskaper; hur datan erhöles och valdes ut; etiketteringsförfaranden (såsom för övervakad inlärning) och metoder för datarensning (identifiering av avvikande värden till exempel)."

EU:s AI-förordning, artikel 10, datastyrning och datahanteringsrutiner

53. Bias-begränsande åtgärder kan vara alla åtgärder som vidtagits för att förhindra bias i data eller i ett system. Dessa kan inkludera slumpmässigt urval, ytterligare datainsamling för att öka datans representativitet, strikta protokoll för utvärdering av fall, fyra-ögon-granskning vid utvärdering av fall, väga om data och många andra åtgärder.

- ▶ Leverantörer är skyldiga att ha infört lämpliga rutiner i den utsträckning att datamängderna är "relevanta, tillräckligt representativa och i största möjliga utsträckning fria från fel och fullständiga med beaktande av det avsedda syftet. De ska ha lämpliga statistiska egenskaper, inklusive, där det är tillämpligt, beträffande de personer eller grupper av personer som högrisk-AI-systemet är avsett att användas för. Dessa egenskaper hos datamängderna kan uppfyllas på nivån för enskilda datamängder eller på nivån för en kombination av dem".

Detta omfattar åtminstone:

- ▶ relevanta designval;
- ▶ datainsamling och dataanvändning (artikel 10.2 b, e och g);
- ▶ processer för databeredning och databehandling, inklusive annotering och etikettering (artikel 10.2 c);
- ▶ formulering av antaganden särskilt gällande mätning och representation (artikel 10.2 d);
- ▶ uttryckligen med avseende på bias, är de skyldiga att undersöka och "vidta lämpliga åtgärder för att upptäcka, förhindra och begränsa eventuella bias" (artikel 10.2 g) i datamängderna som sannolikt "kommer att ha en negativ inverkan på grundläggande rättigheter eller leda till diskriminering som är förbjuden enligt unionslagstiftning" (artikel 10.2 f).

EU:s AI-förordning, artikel 17.1 f, kvalitetsledning

- ▶ Skriftliga policyer, förfaranden och instruktioner ska inkludera: "system och förfaranden för datahantering, inklusive dataanskaffning, datainsamling, dataanalys, dataetikettering, datalagring, datafiltrering, dataminning, dataaggregering, datalagring och alla andra åtgärder gällande data som utförs före och i syfte med utsläppande på marknaden eller ibruktagande av högrisk-AI-system".

Från användaren:

EU:s AI-förordning, artikel 13 b VI, bruksanvisningar

- ▶ Leverantören bör tillhandahålla "när det är lämpligt, specifikationer för indata, eller annan relevant information om de tränings-, validerings- och testdatamängder som använts, med beaktande av det avsedda syftet med högrisk-AI-systemet" (i den information som tillhandahålls användaren av leverantören).

EU:s AI-förordning, artikel 26.4

- ▶ I den utsträckning användaren utövar kontroll över indata ska användaren säkerställa att indata är relevanta och tillräckligt representativa med beaktande av det avsedda syftet med högrisk-AI-systemet.

Övrigt

- ▶ Dokumentationsinitiativ
 - datablad för datamängder;
 - modellkort för modellrapportering;

- tillämpade datastyrnings- och kvalitetsstandarder – när svaranden nämnt användning av standarder kan dessa standarder användas för att förstå vad svaranden åtagit sig.
 - För brottsbekämpning/migrationskontroll kan ytterligare information i databasen vara tillgänglig för den relevanta MSA:n.
-

Indikatorer på diskriminering

Indikatorerna nedan visar problem med datan, som kan bidra till ett prima facie-fall genom att generellt indikera att opålitliga eller bias-drabbade data används och kan användas för att motivera förfrågningar om ytterligare testning. Detta är troligen en ofullständig förteckning. Det kan finnas andra saker som behöver tas i beaktande som inte listas här.

- Allmänna datafrågor:
 - inte representativ (geografiskt område, sektor);
 - felaktigt anskaffad (återanvänd);
 - inte aktuell;
 - det finns urvals-bias;
 - det finns mät-bias såtillvida att etiketten återspeglar ett annat begrepp än det som förutsägs eller att etiketterna generellt är opålitliga.
- Indikatorer på datakvalitetsproblem relaterade till diskriminering:
 - data är inte representativa för alla som påverkas;
 - hög sannolikhet för att återspegla historiska eller bestående ojämlikheter;
 - urvals-bias sannolikt leder till över- eller underrepresentation av grupper med skyddade egenskaper;
 - mät-bias sannolikt påverkar grupper med skyddade egenskaper;
 - det finns en indikation på diskriminering i anskaffning, skapande eller etikettering av datan.
- Beslut beträffande data är otillräckligt dokumenterade.
- Inga begränsningsåtgärder finns eller begränsningsåtgärderna kan inte visas vara effektiva.

Indikatorer på bristande efterlevnad av AI-förordningens krav för högrisk-system

Även om bristande efterlevnad av AI-förordningen inte är fokus för denna bedömningsmetodik kan varje indikation på att kraven i AI-förordningen inte uppfylls informera om korrigerande åtgärder (se [steg 4](#)).

- Leverantören kan inte visa att de har infört rutiner som säkerställer att datamängderna är "relevant, tillräckligt representativa och i största möjliga utsträckning fria från fel och fullständiga med beaktande av det avsedda syftet. De ska ha lämpliga statistiska egenskaper, inklusive, där det är tillämpligt, beträffande de personer eller grupper av personer som högrisk-AI-systemet är avsett att användas för. Dessa egenskaper hos datamängderna kan uppfyllas på nivån

för enskilda datamängder eller på nivån för en kombination av dem”. Eller så kan de inte tillhandahålla dokumentation som bevis för detta.

- ▶ Leverantören har inte utfört bias-testning, eller har inte testat alla relevanta bias.
- ▶ Leverantören har inte infört effektiva korrigeringsåtgärder för att förhindra och begränsa bias som de upptäckt eller misstänker.

- ▶ Dokumentera alla allmänna problem med data, eventuella brister i datakvaliteten, brister i dokumentationen samt eventuella farhågor beträffande riskbegränsande åtgärder
- ▶ Dokumentera alla datarelaterade problem angående ojämlikhet och diskriminering, eventuella farhågor beträffande riskbegränsande åtgärder samt – där sådan information finns – en motivering till hur dessa kan leda till ojämlik behandling i det aktuella fallet.
- ▶ Dokumentera alla indikatorer på bristande efterlevnad av AI-förordningens bestämmelser som är relevanta för data.

3.6. Testning och utvärdering

Resultaten av testning och utvärdering kan ge värdefull information för att bedöma om diskriminering har förekommit. Den centrala frågan här är om algoritmsystemet ger ojämna resultat för olika grupper av individer. Ett sätt att visa detta är genom statistik på systemets utdata.

När statistiska jämlikhetsdata finns tillgängliga bör dessa data utvärderas för validitet och tillförlitlighet (se nedan). Det är därför relevant att beakta validiteten och tillförlitligheten hos de data som används för att beräkna dessa statistikuppgifter. Därför bör detta avsnitt läsas i samband med [avsnitt 3.5](#).

I detta steg kan teknisk och sektorexpertis vara till hjälp.

För att testa utfall för olika grupper med olika skyddade egenskaper behöver det finnas data tillgängliga om dessa skyddade egenskaper. Till exempel, för att testa om ett system producerar olika utdata för män och kvinnor behöver man veta för varje testperson vem som är man och vem som är kvinna. Dessa data kanske inte alltid är tillgängliga. I så fall bör testningen fokusera på egenskaper som är tillgängliga, inklusive sådana som är eller kan vara relaterade till skyddade egenskaper. Till exempel kanske data om etnisk bakgrund inte är tillgängliga, men data om nationalitet är det, eller för ansiktsbilder kanske det inte finns data tillgängliga personer som utsätts för rasdiskriminering men färgvärdet på ansiktsbilder kan användas.

Inhämtande av (ytterligare) statistiska data

När statistiska data inte är tillgängliga eller är otillräckliga för att bedöma algoritmisk diskriminering finns det några alternativ som kan hjälpa till att inhämta (ytterligare) statistiska data.

Utvärdering och testning av MSA

Där ett EB inte kan fastställa om en överträdelse av rättigheter enligt unionslagstiftning har skett möjliggör artikel 77.3 i AI-förordningen för EB:et att lämna en "motiverad begäran till MSA att organisera testning av systemet genom tekniska medel". EB:et bör vara nära delaktigt i testningen.

Skäl 157 klargör att förfarandet enligt artikel 77 ska tillämpas på högrisk-AI-system som utgör en risk för grundläggande rättigheter, förbjudna system som släppts ut på marknaden (artikel 6), tagits i bruk eller använts i strid med förbuden i förordningen (artikel 5) och AI-system som antingen är högrisk eller inte högrisk och har gjorts tillgängliga i strid med transparenskraven i förordningen och utgör en risk (artiklarna 6.4 och 50).

Testning av EB:et eller i samarbete med tredje part

Likabehandlingsorganen kan skapa egna jämlikhetsdata när de utövar sina övervaknings- och andra mandat enligt antidiskriminerings- och/eller EB:ens direktiv och förordningar eller på annat sätt. Dock kan det krävas försiktighet vid användning av dessa data eftersom det finns en risk att de inte skulle betraktas som oberoende vid utövandet av en processföringsfunktion. Det är för närvarande

oklart hur sekretessskyldigheterna i artikel 78 i AI-förordningen skulle påverka användningen av information som erhållits för ett syfte i civilprocess (jfr brottsliga och administrativa förfaranden som uttryckligen nämns).

Beakta om den bevisning som du söker stödja dig på erhölls:

1. till följd av att likabehandlingsorganet utövar sin oberoende och opartiska beslutsfunktion eller sin övervaknings- och databehandlingsfunktion; eller
2. i samband med rättsliga förfaranden.

Om datan skapades för 1., kan oberoendet hos dessa data ifrågasättas. Som Equinet noterat:

Där ett likabehandlingsorgan processar baserat på egna data kan domstolen i teorin ha svårt att se sådana data som externa och oberoende. Ännu mer där ett likabehandlingsorgan prövar fall och förlitar sig (delvis) på egna jämlikhetsdata kan parter i ett mål och andra observatörer ifrågasätta sådan praxis på åtminstone två grunder: är organet en opartisk domare om det proaktivt identifierar och använder bevis som gynnar den påstått utsatta personen[?]; och är sådan bevisning trovärdig om den har tagits fram av organet självt, som per definition verkar för likabehandling och till stöd för personer som anser sig ha utsatts för diskriminering[?]. Sådana potentiella (upplevda) problem i fråga om rätten till en rättvis rättegång och principen om parternas likställdhet (equality of arms) kan påverka ett likebehandlingsorgans övergripande trovärdighet i samhället, även när det gäller dess rapporteringsfunktion. (Ilieva 2024)

Statistiska analyser av det algoritmiska systemet

Om det algoritmiska systemet och relevanta data är tillgängliga för EB:et kan organet genomföra en egen statistisk utvärdering. Både rättvisemått (*fairness metrics*) och metoder för kontroll av efterlevnad (*enforcement methods*) har implementerats i ett antal programverktyg med öppen källkod för analys av rättvisa. Dessa verktyg kan användas för att mäta skillnader i prestanda mellan olika grupper i algoritmiska system eller i de datamängder som görs tillgängliga för EB:et. Vanligt använda exempel är IBM AI Fairness 360 och Microsoft Fairlearn, men det finns många andra tillgängliga verktyg. Det finns inte något enskilt verktyg som generellt kan rekommenderas, eftersom de alla har sina respektive fördelar och begränsningar (Mittelstadt 2024).

Situationstestning

”Situationstestning” kan också användas för att ta fram bevis på diskriminering. Situationstestning är en experimentell metod där man bildar par av testpersoner som endast skiljer sig åt i fråga om ett skyddat kännetecken, i syfte att undersöka skillnader i behandling.

I vissa sammanhang, till exempel vid algoritmisk rekrytering, är ett genomförbart alternativ att samla in statistiska data om ett specifikt algoritmiskt system genom (diskriminerings-) situationstestning. I flera länder, däribland Storbritannien, Nederländerna, Frankrike, Danmark, Finland och Sverige, har situationstestning godkänts som en giltig form av bevisning (Europeiska kommissionen 2019). Situationstestning i stor skala kan också ge statistiska data och har tillämpats i många olika sammanhang, såsom tillgång till sysselsättning, bostäder och andra

typer av varor och tjänster. Testresultaten betraktas som godtagbara bevis i flera EU-länder. För att säkerställa trovärdiga resultat kan likabehandlingsorganet behöva samarbeta med experter för att genomföra testningen (Makkonen 2016).

Samarbete med statistikmyndigheter/användning av allmän statistik

Allmän statistik kan användas för att visa att användningen av proxyvariabler leder till skillnader i behandling. Ett exempel är de nederländska kontrollerna av studiebidrag:

- ▶ algoritmen gjorde åtskillnad utifrån utbildningsnivå;
- ▶ enligt allmän statistik från den nederländska statistikmyndigheten (Statistics Netherlands) är studerande med utomeuropeisk migrationsbakgrund överrepresenterade på de lägsta utbildningsnivåerna. Antalet studerande med utomeuropeisk migrationsbakgrund är ungefär dubbelt så stort som vad som skulle kunna förväntas utifrån studentpopulationen som helhet ([StatLine – Mbo; studenten, niveau, leerweg, migratie 2005/06–2021/22](#));
- ▶ därav följer att användningen av utbildningsbakgrund i det algoritmiska systemet leder till ojämlik behandling, utan att det krävs tillgång till individuella uppgifter om studenternas migrationsbakgrund.

Domstolarnas syn på statistiska data

Om nödvändigheten av statistiska data

Statistik i sig avslöjar inte nödvändigtvis diskriminering eller en diskriminerande praxis (*Hugh Jordan v. the United Kingdom*, Application No. 24746/94, European Court of Human Rights).⁵⁴ De är bland de relevanta bevisformer som kan användas för att fastställa eller bemöta ett prima facie-fall av direkt eller indirekt diskriminering. Domstolar kan generellt vara mer benägna att godta statistisk direkt diskriminering än att förlita sig på sådan bevisning från svarande för att bemöta det, vilket återspeglar den informationsobalans som finns mellan parterna och de erkända svårigheter klagande möter när de ska underbygga diskrimineringsklagomål (Makkonen 2007). Det kommer att vara upp till den nationella domstolen att bedöma om den tillgängliga statistiken är "giltig" och kan beaktas (se Enderby, mål C-127/92, punkt 16).

Statistiska data och direkt diskriminering

I ett fall av direkt diskriminering kan statistik avslöja förekomsten av ett diskriminerande mönster och därmed utgöra bevis för direkt diskriminering i förhållande till ett enskilt klagomål. Användningen av statistisk bevisning är dock vanligtvis inte nödvändig för syftet att skapa ett prima facie-fall av direkt diskriminering.

54. Statistik visade en skillnad mellan antalet personer som dödades och tillhörde olika religioner, men inte huruvida dödandet var olagligt eller skedde genom säkerhetsstyrkornas överdrivna våldsanvändning (punkt 154).

Statistiska data och indirekt diskriminering

Statistiska data är särskilt användbara i fall av indirekt diskriminering, där en klagande kan använda statistik för att visa att tillämpningen av en till synes neutral praxis eller regel har en oproportionerlig konsekvens eller inverkan på en specifik grupp av personer i jämförelse med andra i en liknande situation (FRA 2018: 242).

I synnerhet "där en sökande kan visa, på grundval av oomtvistade officiella statistikuppgifter, förekomsten av en prima facie-indikation på att en specifik regel – även om den är formulerad på ett neutralt sätt – i själva verket påverkar en klart högre procentandel kvinnor än män, åligger det den svarande regeringen att visa att detta är resultatet av objektiva faktorer som inte har samband med någon diskriminering på grund av kön" (*Hoogendijk mot Nederländerna*, 2005).

Ingen strikt tröskel har fastställts men en väsentlig siffra måste vanligtvis uppnås för att visa differentierad inverkan, det vill säga "mycket större antal" (Ingrid Rinner-Kühn, 1989) eller "avsevärt lägre procentandel" (Helga Nimz, 1991, och Maria Kowalska, 1990) eller "mycket fler" (M. A. De Weerd, 1994), även om inlämnande av statistiska data för att bevisa indirekt diskriminering inte är ett formellt krav (se *D.H. och andra mot Tjeckien*, 2007, § 188).

Exakta statistiska data krävs dock inte, särskilt om den part som påstår sig vara diskriminerad inte har tillgång till dessa statistikuppgifter eller har svårt att få tillgång till dem (Minoo Schuch-Ghannadan, 2019).

Användning av andra data och bevis

Bevis om andras praxis eller uppfattningar gentemot dem som tillhör samma skyddade kategori kan vara tillräckligt. Till exempel tillämpades en praxis att placera barn i separata klasser baserat på deras otillräckliga kunskaper i kroatiska endast på romska elever, vilket gav upphov till en presumtion om differentierad behandling trots att statistiken inte fastställde diskriminering. Europeiska domstolen för de mänskliga rättigheterna har klargjort att "indirekt diskriminering kan bevisas utan statistisk bevisning" (*Oršuš och andra mot Kroatien*, 2010).

I ärenden som rör algoritmiska system kan även andra sakuppgifter och bevis vara tillräckliga för att fastställa att det föreligger en prima facie-diskriminering. I Deliveroomålet krävde domstolen till exempel inte omfattande statistiska tester eller komplex dataanalys. Istället räckte själva systemets utformning – såsom den rekonstruerats på en grundläggande nivå utifrån tillgänglig information såsom avtalsklausuler, plattformens vanliga frågor, svarandens externa och interna kommunikation samt vittnesmål från berörda personer – i kombination med en bedömning av dess förutsebara konsekvenser för att påvisa en oproportionerlig nackdel och fastställa en *prima facie*-diskriminering.

Bedömning av jämlikhetsdata som lagts fram av svaranden

För att uppfylla kraven enligt diskrimineringslagstiftning måste statistiska uppgifter vara både valida och tillförlitliga.

Validitets- och tillförlitlighetstestningen bör tillämpas genom hela bedömningen. Följande trestegsprocess kan vara till hjälp:

- ▶ Har den metod som leverantören eller användaren har tillämpat vid testning och utvärdering grundats på, eller resulterat i, valida och tillförlitliga statistiska uppgifter?
- ▶ Är den statistiska bevisning som åberopas för att ifrågasätta svarandens ståndpunkt valid och tillförlitlig?
- ▶ Är den statistiska motbevisning som åberopas valid och tillförlitlig?

För att statistiska uppgifter ska anses vara giltiga bör följande kriterier vara uppfyllda:

- ▶ de bör i allmänhet uppvisa statistisk signifikans;
- ▶ de bör omfatta ett tillräckligt stort antal individer;
- ▶ de bör fånga företeelser som inte är tillfälliga eller kortvariga;
- ▶ de bör sammantaget vara relevanta och tillräckliga.

Tillförlitlighet hos statistiska uppgifter avser stabiliteten eller konsekvensen hos den mätning eller det test som har tillämpats. Ett mått på diskriminering är exempelvis tillförlitligt i den utsträckning som mätmetoden ger samma resultat vid upprepade mätningar. Om man exempelvis mäter antalet anmälda likabehandlingsärenden per år, men antalet behöriga EB förändras över tid, påverkar detta mätningens tillförlitlighet negativt och försämrar även möjligheten att jämföra resultaten över tid. Inget mått är fullständigt tillförlitligt; tillförlitlighet är därför alltid en fråga om omfattning.⁵⁵

Val av mätetal

Inom datavetenskap har flera mätetal identifierats för att mäta olika uppfattningar om vad som är rättvist. Dessa mätetal har inte utformats utifrån diskrimineringsslagstiftning och överlappningen mellan mätetal på rättvisa inom datavetenskap och AI, och föreställningar om jämlikhet, är inte alltid tydliga (Weerts m.fl. 2023). Tyvärr är det matematiskt omöjligt att skapa ett system som är rättvist enligt alla tekniska mått.

Därför är det avgörande att fastställa vad en skillnad i behandling och en särskild nackdel (som en integrerad del av diskriminering) betyder i det aktuella fallets sammanhang och att välja de mätetal som mest nära överlappar med detta. Observera att det inte räcker med att bara presentera ett rättvisemätetal som visar att systemet är "rättvist" enligt detta mätetal för att det ska utgöra ett system som resulterar i likabehandling. Där det finns orättvisa kommer det att vara viktigt att beakta varför beslutet att använda just den uppsättningen jämlikhetsmätetal fattades och vilka åtgärder som vidtagits för att motverka den uppkomna ojämlikheten.

Demografisk likabehandling mäter om sannolikheten för en viss förutsägelse är densamma för två grupper. Detta mått är mest relevant när ett visst resultat är

55. Se Europeiska kommissionen 2021: 20; Enderby, 1993, punkt 17; Seymour-Smith, 1999, punkt 62. Om de allmänna kraven på statistik i diskrimineringsmål; se *Enderby v. Frenchay Health Authority and Secretary of State for Health*, punkt 17.

mindre fördelaktigt än andra resultat för de berörda personerna. Detta återspeglar begreppet ojämн behandling bäst i exempel som bedrägeriförutsägelse (när en grupp misstänks för bedrägeri oftare än den andra gruppen) eller rekrytering (när en grupp väljs ut för intervju oftare än den andra). I många fall är demografisk likabehandling det relevanta mätetal, om det inte finns övertygande skäl att förlita sig på ett annat mätetal. Observera att demografisk likabehandling t ibland även kallas oberoende, gruppriktvisa, statistisk likabehandling, indirekt diskriminering, lika urvalsfrekvens eller lika antagningsfrekvens.

Andra relevanta mätetal bygger på fel i systemet och testar om systemen resulterar i fler fel för en grupp än för en annan. Denna typ av mätetal är mest relevant när skada härrör övervägande från fel i systemen, inte från utdatan i sig. Exempel där felbaserade mätetal är mest relevanta är ansiktsgenkänning och diagnostisk testning inom medicin. Se tabell 6 nedan för en översikt över potentiellt relevanta mätetal.

Det är osannolikt att ett enhetligt tillvägagångssätt för testning kommer att vara tillräckligt. I stället kommer det att vara relevant att beakta både de statistiska modeller som används och deras begränsningar samt de omgivande omständigheterna såsom beslutsfattande organisationen och de processer de tillämpar (Barocas et al. 2023). Som sådant kommer "siffror" utan sammanhang att ha otillräcklig mening för att vara giltiga eller tillförlitliga.

Även när demografisk likabehandling väljs som primärt mätetal rekommenderas det att dessutom beakta ett mätetal som tar hänsyn till fel. Det finns ett krav på leverantörer att fastställa en "förväntad noggrannhetsnivå" för specifika grupper av personer (AI-förordningen, bilaga IV 3).

Jämlikhetsstatistik kan presenteras som ett förhållande eller som separata mätetal mellan grupper. Några exempel följer.

- ▶ Demografisk likabehandlingsstatistik i ett rekryteringssammanhang kan uttryckas som: 25 % av manliga sökande bjuds in till intervju och 15 % av kvinnliga sökande bjuds in efter automatiserad CV-urval. Detta kan också uttryckas som ett förhållande: $0,15/0,25 = 0,6$.
- ▶ Lika möjligheter för cancerscreening kan uttryckas som: 75 % av män med cancer upptäcks av screeningverktöget och 65 % av kvinnor med cancer upptäcks av screeningverktöget. Detta kan också uttryckas som ett förhållande: $0,65/0,75 = 0,87$.

Tabell 6. Beskrivning av rättvisemätetal

(anpassad med tillstånd från Mittelstadt 2023)

Metrisk	Beskrivning	När tillämpbar	Exempel
Demografisk likabehandling	Olika grupper bör få del av de positiva resultaten i samma utsträckning. Demografisk jämlikhet kan även benämnas oberoende, grupprättvisa, statistisk likabehandling, indirekt diskriminering, lika urvalsfrekvens eller lika antagningsfrekvens.	▶ Situationer där ett visst resultat i sig (och inte fel) är ofördelaktigt. ▶ Situationer där historiska data förväntas vara skadliga. ▶ Situationer där det saknas tillförlitliga etiketter (ingen gemensamt överenskommen referenssanning).	Rekrytering, godkännande av lån, upptäckt av bedrägerier, tillgång till utbildning.
Lika möjligheter	Olika grupper bör uppvisa samma andel falska negativa resultat. Detta kan även kallas balans i andelen falska negativa resultat.	▶ Situationer där den övervägande skadan beror på falska negativa resultat, och ▶ Det finns tillförlitliga klassificeringar (en överenskommen referenssanning).	Sjukdoms-screening.
Förväntad likabehandling	Andelen falska positiva resultat, räknat av alla positiva resultat, bör vara densamma för olika grupper. Detta kan även kallas utfallstest eller jämlikhet i positivt prediktivt värde.	▶ Situationer där den övervägande skadan beror på falska positiva resultat, och ▶ Det finns tillförlitliga etiketter (en överenskommen referenssanning).	Ansiktsgen-känning som polisen använder för att identifiera en person av intresse.
Balans mellan andelen falska positiva resultat	Olika grupper bör få falska positiva resultat i samma utsträckning. Detta kan även kallas prediktiv jämlikhet.	▶ Situationer där den övervägande skadan beror på falska positiva resultat, och ▶ Det finns tillförlitliga etiketter (en överenskommen referenssanning).	Ansiktsgen-känning som polisen använder för att identifiera en person av intresse.

Metrisk	Beskrivning	När tillämpbar	Exempel
Ujämnade odds	Andelen sanna positiva och falska positiva bör vara densamma för olika grupper. Detta kan även benämnas "likhet i villkorad procedurnoggrannhet" och "olika behandling".	► Situationer där man bör ta hänsyn till både skadorna av falska positiva resultat och falska negativa resultat, och ► Det finns tillförlitliga etiketter (en överenskommen referenssanning).	Behandling av sjukdom genom att utföra en riskfylld operation.

Utvärdering av statistik som tillhandahållits av svaranden

För att utvärdera validiteten hos statistiska jämlikhetsmätetal bör EB:er åtminstone beakta tre frågor.

1. Generellt, kvaliteten och bias i den data som används för att beräkna mätetalet (se [steg 3.5](#)).
2. Utformningen av mätetalet och valet av jämförelsegrupper.

Vid beaktande av hur en leverantör av ett system har identifierat och mätt eventuella prestandaskillnader bör EB:er ägna särskild uppmärksamhet åt hur beslutet att mäta har fattats och vilka grupper av personer som jämförs. Detta är ett viktigt beslut i diskrimineringslagstiftningens sammanhang där definiering av lämpliga "missgynnade" och "jämförelsegrupper" ofta är viktigt. (Mittelstadt 2024). Valet av grupper kan (medvetet) dölja verkliga ojämna utfall.

EB:er rekommenderas att:

- beakta vilka lämpliga "missgynnade" och "jämförelsegrupper" som bör beaktas, innan man tittar på svarandens statistik;
- beakta om svaranden har valt och definierat samma grupper och, om inte, beakta vilken information som läggs till eller förloras i svarandens val;
- beakta intersektionella grupper, eftersom ett fokus enbart på grupper med en enskild egenskap döljer intersektionella ojämlikheter;
- för att förstå de jämförelsegrupper som valts av svaranden, titta på förklaringar som tillhandahållits av svaranden, i synnerhet all information om i vilken utsträckning en variation i de personuppgifter som beaktats skulle ha lett till ett annat resultat;
- i fall där svaranden inte har rapporterat om relevanta rättvisemätetal eller har definierat påverkade grupper snävt, överväga att inhämta ytterligare statistiska data (se rutan om Inhämtande av (ytterligare) statistiska data ovan) (Mittelstadt 2024).

Statistiska jämlikhetsmätetal är ofta en del av bredare prestandarapporter som inkluderar en rad olika testresultat som mäter olika aspekter av ett systems prestanda, såsom noggrannhet (hur ofta systemet gör rätt förutsägelse) och robusthet

(motståndskraft mot fel, oantagna beteenden och förändringar i verkligheten (drift)). Det är värdefullt för EB:er att förstå prestandan i allmänhet, eftersom systemets prestanda är mycket relevant i bedömningen av algoritmisk diskriminering, i synnerhet när man beaktar en potentiell objektiv motivering (se [steg 3.8](#)) där centrala frågor kommer att involvera att fastställa om systemet är lämpligt eller ändamålsenligt och nödvändigt.

3. Relevans hos mätetalen för det aktuella systemet som tillämpas i praktiken.

Mätetalen bör återspegla det aktuella algoritmsystemet. De tester som tillämpats kan ha blivit föråldrade eller processen kring systemet kan ha förändrats (till exempel tillämpar systemet en statlig policy och den policyn har ändrats).

Observera att de rapporterade mätetalen bör återspegla systemet efter eventuella begränsande åtgärder för att "eliminera partiskhet" eller förhindra diskriminering har tillämpats. Om dessa åtgärder är manuella bör EB:er säkerställa att statistiken återspeglar utfallet efter denna manuella intervention.

Där ett högrisk-AI-system har erhållit ett överensstämmelsecertifikat efter en bedömning kommer det att anses att det uppfyller datakraven i artikel 10.4 i AI-förordningen, nämligen "att de i den utsträckning som krävs av det avsedda syftet beaktar de egenskaper eller element som är särskilda för den specifika geografiska, kontextuella, beteendemässiga eller funktionella miljö inom vilken högrisk-AI-systemet är avsett att användas".

Det är troligt att dessa överväganden kommer att vägleda tillämpningen av "validitetstestet" (se mer om statistiska data under steg 3.6), men det är viktigt att notera att detta antagande kan bemötas. Likabehandlingsorgan kan önska att beakta, i synnerhet, om de data som använts i själva verket uppfyller dessa krav för den givna användningen av AI-systemet. Om systemet används på annat sätt än för sitt avsedda syfte kommer detta antagande inte att gälla.

Frågor att beakta

Om allmän prestandatestning

- ▶ Vilka aspekter av prestanda har testats?
- ▶ Hur och var har systemets prestanda över individer och grupper av personer registrerats?
 - Kommer dessa siffror att uppdateras under systemets livscykel? I så fall, hur?
- ▶ Vilka potentiella konsekvenser för individer och grupper av personer har beaktats vid genomförandet av prestandatestning?
- ▶ Valdes någon form av skada som fokuspunkt för testning (såsom falska positiva eller falska negativa)?
- ▶ Vilka mätetal användes för prestandatestning?
- ▶ Har konfusionsmatriser producerats genom prestandatestning?
- ▶ Hur validerades testresultaten?

- ▶ Är de data som används för att genomföra prestandatestning tillräckligt representativa, relevanta och korrekt och lämpligt mätta (se [steg 3.5](#))?
- ▶ Uppfyller svarandens prestandatestning validitets- och tillförlitlighetstesten?

Jämlikhetsdata

- ▶ Vilken (typ av) jämlikhetsdata, rättvisemätetal eller bias-tester användes?
 - Återspeglar de mätetal som valts av svaranden likabehandling för det specifika sammanhanget? (Vanligtvis demografisk likabehandling om det inte finns tydliga motiveringar för andra mätetal, se tabell 6 ovan).
 - I fall där det finns felbaserade mätetal, beaktades samma former av skada som i vanlig prestandatestning? (Till exempel, när vanlig prestandatestning använde ett mätetal som fokuserade på falska positiva, gjorde jämlikhetstestningen detsamma?) När man testar för jämlikhet baserat på prestanda bör det mest relevanta prestandamätetale logiskt sett också användas för detta jämlikhetstest. Ett omotiverat val av ett annat mätetal kan peka på en brist i jämlikhetstestningen.
 - Vad är motiveringen eller grunden på vilken dessa valdes?
 - Övervägdes andra och/eller avvisades? I så fall, varför/inte?
 - Var det sortiment av tester som användes tillräckligt brett för att fånga olika typer av potentiell diskriminering eller konsekvenser som påverkar olika (intersektionella) grupper?
 - Har relevanta "jämförelsegrupper" valts för varje test? Inkluderas dessa överväganden i systemets tekniska dokumentation?
- ▶ Har någon form av metod för att säkerställa rättvisa eller för att minska snedvridning (eliminering av partiskhet) använts?
 - Vad är motiveringen eller grunden på vilken dessa valdes?
 - Övervägdes andra och/eller avvisades? I så fall, varför/inte?
 - Inkluderas dessa överväganden i systemets tekniska dokumentation?
- ▶ Vilka fördelar och skador resulterade i säkerställande av rättvisa eller snedvridning i?
 - Vilka grupper av personer påverkades/ kommer detta att påverka?
 - Undersöktes de sekundära konsekvenserna av likabehandling för andra grupper?
- ▶ Hur validerades testresultaten?
- ▶ Är de data som används för att beräkna statistiken tillräckligt representativa och relevanta, och korrekt och lämpligt mätta (se [steg 3.5](#))?
- ▶ Uppfyller de statistiska data som tillhandahållits av svaranden validitets- och tillförlitlighetstesten?

Potentiella källor till information för att bedöma statistiska jämlighetsdata

Klagande

GDPR, artikel 15.1 h, meningsfull information om den logik som tillämpas

Svaranden

Från leverantören

EU:s AI-förordning, artikel 10, data och datastyrning

EU:s AI-förordning, artikel 11, teknisk dokumentation (bilaga IV 2 g, 3):

- ▶ "validerings- och testförfarandena som använts, inklusive information om de validerings- och testdata som använts och deras huvudsakliga egenskaper; mätetal som används för att mäta noggrannhet, robusthet och efterlevnad av andra relevanta krav som anges i kapitel III, avsnitt 2, samt eventuella potentiellt diskriminerande konsekvenser; testloggar och alla testrapporter daterade och undertecknade av ansvariga personer, inklusive beträffande förutbestämda ändringar som avses under punkt f"

EU:s AI-förordning, artikel 11, loggar

Annan information som kan begäras från svaranden:

- ▶ tillgång till systemet;
- ▶ exempel gällande(syntetiska) datamängder;
- ▶ eventuella rapporter om systemets prestanda eller tillförlitlighet, inklusive felfrekvensmatriser;
- ▶ information om typen av testning och testmetodiken, inklusive eventuella (harmoniserade) standarder som tillämpats;
- ▶ eventuella rapporter om bias- och rättvisetestning som genomförts;
- ▶ information om typen av testning och testmetodiken, inklusive eventuella (harmoniserade) standarder som tillämpats;
- ▶ frivillig HRIA eller FRIA, DPIA, etik- eller barnkonsekvensbedömning;
- ▶ eventuella andra (risk)bedömningar som delas av svaranden;
- ▶ begäran om allmänna handlingar.

Indikatorer på diskriminering

- ▶ Det finns skillnader i utfall eller felnivåer mellan grupper av personer med skyddade egenskaper, baserat på de mått som är mest relevanta för att bedöma likabehandling i det aktuella fallet.
- ▶ De mått som är mest relevanta för att bedöma likabehandling i det aktuella fallet finns inte tillgängliga, eller också används mått som saknar relevans.
- ▶ Det finns skillnader i utfall eller felnivåer mellan grupper av personer med skyddade egenskaper, baserat på andra mått än de som bäst speglar likabehandling i det aktuella fallet.

- ▶ Samtliga berörda grupper har inte inkluderats, eller också har de berörda grupperna definierats alltför snävt, utan hänsyn till intersektionell diskriminering.
- ▶ Rätt grupper har identifierats och relevanta mått har använts, men slutsatserna håller inte vid en närmare granskning, exempelvis på grund av brister i datamaterialet.
- ▶ Endast ett rättvisemått (fairness metric) redovisas, i stället för resultaten från flera rättvisemått (t.ex. precision, återhämtning och felnivåer⁵⁶). En utvärdering med flera rättvisemått kan behövas för att identifiera olika typer av skillnader i systemets prestanda.

Kumulativa indikatorer på diskriminering

Samtliga indikatorer som beskrivits ovan kan verka kumulativt. Nedan följer några exempel.

- ▶ Risk för diskriminering identifierad av leverantör/användare (se [steg 3.3](#)) men testningen täckte inte hela uppsättningen av individer och/eller grupper som löper risk för skada – beakta i synnerhet proxies, multipel och intersektionell diskriminering. Till exempel testades könsdiskriminering men inte diskriminering baserat på rastillhörighet eller etniskt ursprung trots att risk för rasmässig diskriminering identifierades.
- ▶ (Risk för) diskriminering identifierad av tredje part (se [steg 3.1](#)) men testningen täckte inte hela uppsättningen av individer och/eller grupper som löper risk för skada – beakta i synnerhet proxies, multipel och intersektionell diskriminering. Till exempel testades för könsdiskriminering men inte diskriminering baserat på ras eller etniskt ursprung trots att personer i befintliga liknande fall har missgynnats baserat på ras och graviditet.
- ▶ Indatakriterier är starkt relaterade till skyddade egenskaper eller proxies används i algoritmer (se [steg 3.3](#)) och ingen testning genomförs för att förstå om kraven på likabehandling uppfylls för den skyddade gruppen.

- ▶ Dokumentera eventuella utdata- eller behandlingsluckor som framkommer i statistiken.
- ▶ Dokumentera eventuella problem med den rapporterade statistiken (såsom fel mätetal eller datafrågor).
- ▶ Dokumentera eventuella luckor i testningen.
- ▶ Kombinera resultaten från detta steg med resultaten från föregående steg för att förstå om det tillsammans kan argumenteras för ett fall av *prima facie*-diskriminering och/eller om svaren framgångsrikt har bemött misstanken eller motiverat diskrimineringen (om man tittar på motivering, se även 3.4 för att beakta systemets syfte).
- ▶ Dokumentera slutsatsen.

56. I dataanalyser kan olika mätetal användas för att återspegla olika aspekter av ett systems prestanda. Till exempel fokuserar precision på noggrannheten i positiva förutsägelser (undvikande av falska alarm), medan återkallningsgraden mäter förmågan att upptäcka alla faktiska positiva instanser (undvikande av förbisedda händelser).

Exempel: Smart Check Amsterdam

Denna modell utformades för att granska ansökningar om välfärdsstöd och identifiera de som mest sannolikt involverade stora fel, flagga dem och skicka ärendena till en utredningsavdelning för ytterligare granskning. Syftet var att minska antalet som flaggades för utredning.

Femton egenskaper beaktades, inklusive en rad faktorer såsom om sökanden tidigare ansökt om eller erhållit förmåner, antalet adresser de haft och summan av deras tillgångar för att fastställa en riskpoäng. Den undvek medvetet att använda skyddade grunder och försökte undvika "proxies".

Enligt författarna "tränades den på en datamängd som omfattade 3 400 tidigare utredningar. Detta ledde till en risk för att inbädda historisk bias, t.ex. om mänskliga handläggare historiskt sett hade gjort högre felprocent med en specifik etnisk grupp kunde modellen felaktigt lära sig att förutsäga att denna grupp begår bedrägeri i högre utsträckning".

Tjänstepersoner i Amsterdam antog en definition av rättvisa baserad på att fördela bördan av felaktiga utredningar lika över olika demografiska grupper. De genomförde också omfattande konsultation. Alla intressentgrupper från Amsterdams dataskyddsmyndighet till externa konsultfirmor godkände projektet, förutom den grupp som representerade dem som direkt påverkades (förmånstagare). Ändå sattes systemet i testning. Den initiala modellen var dubbelt så benägen att felaktigt flagga icke-nederländska sökande och nästan dubbelt så benägen att flagga en sökande med icke-västerländsk nationalitet som en med västerländsk nationalitet. Den var också 14 % mer benägen att felaktigt flagga män för utredning.

Amsterdam stad testade också de data som samlats in från mänskliga handläggare. I den processen upptäckte de att människor var mer benägna att felaktigt flagga nederländska medborgare och kvinnor.

På grund av förekomsten av bias försökte teamet använda omviktning av träningsdata för att åtgärda problemet. Detta innebar att sökande med icke-västerländsk nationalitet som bedömdes ha gjort betydande fel i sina ansökningar gavs mindre vikt i datan och de med västerländsk nationalitet gavs mer. När modellen hade omviktats var nederländska och icke-nederländska medborgare lika benägna att felaktigt flaggas.

Problemet med omviktningen av datan var att även om det kunde minska orättvisa mot personer med migrantbakgrund garanterade det inte rättvisa över intersektionella identiteter.

Trots detta togs systemet i bruk i begränsad omfattning för att användas på verkliga sökande. Systemets prestanda skulle bedömas utifrån två huvudsakliga kriterier: i) kunde det bedöma sökande utan snedvridning? ii) kunde det upptäcka bidragsfusk bättre och mer rättvist än mänskliga handläggare?

I stället för att flagga färre sökande för utredning flaggade det fler. Systemet var inte bättre än en människa på att identifiera de som krävde extra granskning. Dessutom flaggade systemet trots omviktningen nu felaktigt sökande med

nederländsk nationalitet och kvinnor. Dessutom var välfärdssökande med barn också mer benägna att flaggas för utredning.

Projektet gick slutligen inte vidare.

Ovanstående är en sammanfattning av "Inside Amsterdam's high-stakes experiment to create fair welfare AI", 11 juni 2025, MIT Technology Review, Lighthouse Reports and Trouw E. G., Geiger G. and Braun J-C. För mer detaljerad information, se Lighthouse Reports 2025.

3.7. Bedömning av om bevisningen är tillräcklig för att stödja ett *prima facie*-fall av diskriminering som involverar ett AI- eller algoritmsystem

- ▶ Granska all information och svaren på frågorna i steg 3 för att beakta om ett *prima facie*-fall av diskriminering har fastställts. Använd indikatorerna på diskriminering som vägledning.
- ▶ Vid detta tillfälle, förutsatt att det inte finns någon adekvat förklaring till de fakta som visar ett *prima facie*-fall (det kan finnas; dock skulle detta vara relevant för bedömningen av svarandens försvar (se [3.8 nedan](#))).
- ▶ På samma sätt, där det finns brist på bevis eller brist på transparens om ett system, bör detta inte hindra att ett fall väcks eller fortsätter vid detta tillfälle. Kom ihåg att det i steg 1 identifierades att tidig avvisning av klagomål där det finns en informationsobalans sannolikt skulle vara olämplig: se [1.4 Manjang v. Uber Eats UK Ltd & Others \(2021\)](#).

Du kan behöva beakta alla eller några av följande faktorer.

- ▶ Vem är offret och vilka skyddade grunder är engagerade?
- ▶ Vem är den lämpliga jämförelsepersonen? Varför har denna jämförelseperson valts?
- ▶ Har direkt eller indirekt diskriminering förekommit?
- ▶ Finns det flera eller intersektionella former av diskriminering?
- ▶ Har du fastställt någon proxy-diskriminering?
- ▶ Varför har diskrimineringen förekommit (kausalitet)? Beakta till exempel om diskrimineringen har förekommit på grund av systemets syfte, utveckling av systemet, användningen av data inom systemet eller för att systemet inte används för sitt avsedda syfte eller används felaktigt av användaren.

För indirekt diskriminering beakta även följande.

- ▶ Vilken/Vilket är den neutrala bestämmelsen, kriteriet eller praxisen?
- ▶ Hurdana är konsekvenserna av bestämmelsen, kriteriet eller praxisen betydligt mer negativa för en skyddad grupp?
- ▶ Har klaganden visat på fakta som det kan dras slutsatser från att personen har behandlats mindre gynnsamt på grund av skyddade grunder?

- ▶ Om skillnaden i behandling inte är "direkt" relaterad till en skyddad egenskap, är den tillräckligt kopplad? Detta kommer att vara särskilt viktigt när man beaktar proxy-diskriminering.
- ▶ Vad var sammanhanget för diskrimineringen? Kan svaranden framgångsrikt argumentera för att beslutet fattades av ett annat skäl än förbjuden diskriminering?

Med beaktande av alla omständigheter, har ett prima facie-fall av diskriminering fastställts?

Om ja, fortsatt med avsnitt [3.8](#).

Om nej, beakta följande.

- ▶ Har det förekommit några underlåtelser att uppfylla dokumentations-, testnings-, övervaknings- eller rapporteringsskyldigheter som kan förklara diskrimineringen eller resultera i luckor i bevisningen?
- ▶ Finns det några förberedande steg som kan vidtas för att fylla eventuella luckor i bevisningen?
 - Överväg att göra ansökningar om utlämnande av information baserat på överväganden kring transparens- och rättvisa enligt artikel 13 i Europeiska konventionen om de mänskliga rättigheterna (rätt till ett effektivt rättsmedel inför en nationell myndighet) och artikel 47 i Europeiska unionens stadga om de grundläggande rättigheterna (rätt till ett effektivt rättsmedel och en rättvis rättegång).
 - Observera att det i vissa omständigheter där svaranden söker argumentera kommersiell sekretess eller tredje parts rättigheter kan konfidentiella förfaranden vara tillgängliga via MSA, DPA eller domstolar. Aktuell rättspraxis tyder på en växande benägenhet att kräva att denna information görs tillgänglig; se [Dun & Bradstreet](#) (i GDPR-sammanhang):

Där den personuppgiftsansvarige anser att denna information inte kan tillhandahållas utan att avslöja tredje parts uppgifter eller företagshemligheter är den personuppgiftsansvarige skyldig att tillhandahålla den påstådda skyddade informationen till den behöriga tillsynsmyndigheten eller domstolen, som måste väga de rättigheter och intressen som står på spel för att avgöra omfattningen av den registrerades rätt till tillgång enligt artikel 15 i den förordningen (se även [Konfidentiell behandling](#)).

För relevansen av att tillhandahålla statistisk information för att fastställa ett prima facie-fall av diskriminering, se [steg 3.6 om statistiska data](#).

Transparens

En domstol kan dra slutsatser från en underlåtenhet att tillhandahålla relevant statistisk eller annan data. Slutsatser om diskriminering kan också dras från omgivande fakta och omständigheter såsom svarandens svar på frågor eller en underlåtenhet att följa en relevant praxiskod. Genom att inte tillhandahålla information kan svaranden i praktiken underlåta att uppfylla bevisbördan, vilket leder till att domstolen utgår från att diskriminering förekommit (se [3.8 för de omständigheter där en rättslig bevisbörda att tillhandahålla information uppstår](#)).

- ▶ Slutsatser kan ofta dras från ett tvetydigt, otillräckligt eller undvikande svar på en enkät eller redovisning under förhandsprocessen eller rättegången.
- ▶ Slutsatser kan också dras från att vägra bekräfta eller förneka förekomsten av information.

Frågor att beakta

- ▶ Har det förekommit några relevanta underlåtelse att lämna ut information från svaranden som lett till brist på transparens? Till exempel, fanns det ett adekvat svar på eventuella begäran enligt GDPR, eventuella begäran om allmänna handlingar, korrespondens före rättsprocessen, begäran enligt EU:s AI-förordning eller annat?
- ▶ Kan några slutsatser dras från den tillgängliga bevisningen? Till exempel statistik, omgivande fakta, känd diskriminering?
- ▶ Kan några slutsatser dras från brist på bevis? Till exempel en underlåtenhet att svara på frågor, underlåtenhet att följa riktlinjer eller praxiskoder?
- ▶ I alla omständigheter, ger omgivande fakta och omständigheter, inklusive eventuell underlåtenhet att svara på frågor eller begäran om information såsom en DSAR-begäran eller att tillhandahålla en förklaring enligt artikel 86 i EU:s AI-förordning, anledning att dra en slutsats om diskriminering?

Exempel från EU-domstolen

- ▶ C-109/88, *Handels- og Kontorfunktionærernes Forbund I Danmark v. Danfoss*: facket väckte talan å kvinnliga arbetstagares vägnar i ett företag eftersom de i genomsnitt tjänade 7 % mindre än sina manliga kollegor i samma eller liknande befattning. I detta fall beslutade EU-domstolen att eftersom lönesystemet som användes helt saknade transparens måste man utgå från att diskriminering har förekommit, och bevisbördan måste övergå till svaranden/arbetsgivaren för att visa att det finns objektiva och icke-diskriminerande skäl för skillnaden i lön.
 - ▶ C-415/10, *Galina Meister v. Speech Design Carrier Systems GmbH*: den sökande hävdade att ett förekommit diskriminering på grund av kön, ålder och etniskt ursprung i rekryteringsprocessen och begärde att domstolen skulle ålägga arbetsgivaren att offentliggöra information om huruvida en annan kandidat anställdes i slutet av processen. EU-domstolen ansåg att även om likabehandlingsdirektiven inte föreskriver någon rätt till tillgång till denna typ av information kan svarandens vägran att ge någon tillgång till information för den sökande vara en av de faktorer som kan beaktas i samband med fastställandet av de fakta som det kan antas att direkt eller indirekt diskriminering har förekommit.
-

Exempel på konsekvenserna av brist på transparens från nationell rättspraxis

► **Nederländskt mål NL24.38560 (2025) (interimistiskt domslut)**

Domstolen beslutade på eget initiativ att utreda förekomsten av en visumalgoritm och ålade den relevanta ministern att:

- tillhandahålla information som är relevant för beslutsfattande;
- genomföra en HRIA;

på grund av transparens i beslutsfattande och rätt till effektivt rättsligt skydd enligt både artikel 13 i Europeiska konventionen om de mänskliga rättigheterna och artikel 47 i Europeiska unionens stadga om de grundläggande rättigheterna.

► **UK Tax Tribunal case (First tier): Elsbury v. The Information Commissioner (2025) UKFTT 915 (GRC) (2 augusti 2025)**

”HMRC:s underlåtenhet att antingen bekräfta eller förneka att de innehar den begärda informationen förstärker tron – baserat på indikatorer i HMRC:s korrespondens om R&D-anspråk – att AI används av HMRC-tjänstemän – kanske på ett obehörigt sätt – vilket underminerar skattebetalarnas förtroende och tillit till HMRC:s behandling av anspråk, vilket i sin tur avskräcker legitima sökande från att lämna in anspråk och därmed hindrar själva policymålen för R&D-skatteavdragsordningen.”

► **Filcams Cgil Bologna and Others v. Deliveroo Italia S.R. (2020):**

Domstolen i Bologna fann att företaget inte hade uppfyllt sin bevisbörda, eftersom det varken visade eller förklarade hur algoritmen faktiskt fungerade och inte heller lade fram någon bevisning om dess faktiska funktion. Företaget klargjorde inte vilka specifika kriterier som användes för att fastställa budens statistik, och dessa kriterier offentliggjordes inte heller på plattformen, som endast hänvisade till ”tillförlitlighet” och ”deltagande” som indikatorer för fördelning av framtida arbetsmöjligheter. För ytterligare detaljer om detta mål, se [Purificato \(2021\)](#).

3.8. Kan svaranden visa att skillnaden i behandling inte är diskriminerande?

I praktiken kommer du troligen att behöva beakta samma bevis och ställa samma frågor som anges ovan (3.1–3.8). Det är dock möjligt att det vid detta tillfälle nu finns ytterligare bevis att granska eller att svarandens förklaring av bevisningen eller systemet är mer detaljerad. Detta kommer att kräva en granskning av det ursprungliga materialet och den tidigare genomförda analysen för att säkerställa att den förblir giltig i ljuset av den nya informationen.

När den person som påstår sig ha blivit utsatt för diskriminering har visat att det föreligger ett prima facie-fall av diskriminering, övergår bevisbördan därefter till

svaranden, som måste visa att skillnaden i behandling inte är diskriminerande. Detta kan göras på följande sätt:

- ▶ Genom att visa att det i själva verket inte föreligger någon skillnad i behandling, exempelvis eftersom jämförelsepersonen har valts felaktigt, eftersom beslutsunderlaget inte grundades på den skyddade grunden, eller eftersom omständigheterna och sakförhållandena i det aktuella fallet skiljer sig från jämförelsefallet.
- ▶ Genom att visa att det inte fanns något orsakssamband mellan den förbjudna diskrimineringsgrunden och den särbehandling som skett.
- ▶ Genom att visa att även om särbehandlingen har samband med den förbjudna diskrimineringsgrunden kan den motiveras enligt något av de undantag som anges i EU-rätten (såsom genuina och avgörande yrkeskrav).
- ▶ Genom att visa att den åtgärd, lagstiftning eller praxis som har antagits är lämplig och nödvändig för att uppnå det relevanta legitima målet och att eventuella nackdelar som uppstår inte står i oproportionerlig proportion till de eftersträlvade målen (rimlig och objektiv motivering), i enlighet med Europakonventionen vid direkt och indirekt diskriminering samt med EU-rätten i fall av indirekt diskriminering.

Om svaranden inte kan visa något av detta hålls de ansvariga för diskriminering.

I samband med algoritmiska system eller AI-system innebär detta att svaranden, när ett prima facie-fall har visats, måste lämna ut relevant information om systemet. Denna information kan syfta till att visa att de val som gjorts har varit förenliga med objektiva och icke-diskriminerande kriterier. För betydelsen av att tillhandahålla statistiska uppgifter vid bemötande av bevisbördan, se avsnitt [3.6 tillvägagångssätt för statistiska uppgifter](#).

Tabell 7. Direktiv om likabehandling och förskjutningen av bevisbördan, med exempel på EU-domstolens tillvägagångssätt för att dra slutsatser av underlåtenhet att lämna ut bevisning

Direktiv	Skäl/ artikel	Skyddade områden	Tillämpnings- område	Ett exempelfall som illustrerar EU-domstolens tillvägagångssätt när det gäller att dra slutsatser av att en svarande underlåtit att lämna ut relevant eller nödvändig information och/eller andra krav på öppenhet
2000/43/EG (direktivet om lika- behandling oavsett ras)	Skäl 15 och 21 Artikel 8	Ras och etniskt ursprung – inte medborgarskap och utan att det påverkar den behandling som följer av den rättsliga ställningen för tredjelands- medborgare och statslösa personer	Sysselsättning, yrkesutbildning, socialt skydd, utbildning, varor och tjänster, inklusive bostäder	Mål C-415/10: vägran att lämna ut uppgifter (till exempel om rekryteringskriterier) kan ge upphov till en slutsats om diskriminering
Direktiv 2000/78/EC	Skäl 31 Artikel 10	Religion eller övertygelse, funktions- nedsättning, ålder, sexuell läggning	Anställning och yrke	Kelly C-104/10: ingen rätt till uppgifter, men underlåtenhet att lämna ut uppgifter kan göra det möjligt för domstolarna att dra slutsatser
Direktiv 2006/54/EG (Omarbet- ning av jämfäll- hets- direktivet)	Skäl 30 Artikel 19	Kön (inklusive graviditet och moderskap)	Sysselsättning och yrke (inklusive socialförsäkrings- system)	Danfoss C-109/88: när lönesystemet är otydligt läggs bördan på arbetsgivaren Enderby C-127/92: motiveringar till löneskillnader kan krävas att uppfylla test av objektiv motivering Kenny C-427/11: uppgifter om jämförelsegruppens löner kan krävas för att motbevisa den preliminära bevisningen

Direktiv	Skäl/ artikel	Skyddade områden	Tillämpnings- område	Ett exempelfall som illustrerar EU-domstolens tillvägagångssätt när det gäller att dra slutsatser av att en svarande underlåtit att lämna ut relevant eller nödvändig information och/eller andra krav på öppenhet
Direktiv 2004/113/EG (direktivet om jäm- ställdhet, varor och tjänster)	Skäl 22 Artikel 9	Kön	Varor och tjänster	Test-Achats C-236/09 Se även det särskilda förbudet mot att använda kön som en faktor vid beräkningen av premier för försäkringar och därmed sammanhängande finansiella tjänster, samt beaktandet av försäkringstekniska faktorer i artikel 5
2010/41/EU (direktivet om egen- företagande ur ett jäm- ställdhets- perspektiv)		Kön	Egenföretagare, makar eller sambor till egenföretagare	(Pågående mål) Loredas C-531/23: rörande lagligheten av undantag från skyldigheten att föra arbetstidsregister
2023/970/EU Direktivet om lika lön (införlivande senast i juni 2026)	Artiklar 5, 6 och 8–9		Sysselsättning	Ger en direkt rätt till öppenhet. Kapitel II föreskriver en rätt till löneöppenhet (artikel 5), lönesättning och karriärutvecklingspolicy (artikel 6), en rätt till information som måste vara tillgänglig (artiklarna 8 och 9) samt rapportering av löneskillnader (artiklarna 9 och 10)

Finns det objektiv motivering?

För att visa att en skillnad i behandling inte är diskriminerande kan svaranden i ett fall av indirekt diskriminering⁵⁷ argumentera för att det finns en objektiv motivering för praxisen. Detta avsnitt belyser punkter att utvärdera en objektiv motivering som framförs av en svarande.

För att motivera differentierad behandling måste det visas att:

57. Enligt EU-rätt är motivering normalt inte tillgänglig i fall av direkt diskriminering, förutom i fall som rör ålder/funktionsnedsättning och verkliga yrkeskrav

- ▶ algoritmsystemet har ett legitimt syfte;
- ▶ algoritmsystemet är lämpligt eller ändamålsenligt för att uppnå det syftet;
- ▶ algoritmsystemet är nödvändigt för att uppnå det syftet;
- ▶ algoritmsystemet är proportionerligt, det vill säga att algoritmsystemet är effektivt för att uppnå syftet, det inte finns andra sätt att uppnå det legitima syftet som skulle leda till mindre skillnad i behandling, och det legitima syftet är tillräckligt viktigt för att motivera skillnaden i behandling.

Exempel från EU-domstolen: C-167/97 Seymour-Smith

I fall där talan om osaklig uppsägning endast kunde väckas efter två års anställning och där statistik visade att detta innebar indirekt diskriminering av kvinnor, var den medlemsstat som ansvarade för den indirekt diskriminerande bestämmelsen, kriteriet eller praxisen skyldig att visa att regeln återspeglade ett legitimt mål inom dess socialpolitik, att målet inte hade något samband med könsdiskriminering och att staten rimligen kunde anse att de valda medlen var lämpliga för att uppnå detta mål. Domstolen fann i det målet att "enbart allmänna antaganden om en viss åtgärds förmåga att främja rekrytering är inte tillräckliga för att visa att syftet med den omtvistade regeln saknar samband med könsdiskriminering eller för att tillhandahålla underlag på grundval av vilket det rimligen kan anses att de valda medlen var lämpliga för att uppnå detta syfte" (punkterna 76–77). Domstolen noterade även att statistik som tagits fram efter det att åtgärden antogs kunde beaktas för att ge en indikation på åtgärdens påverkan på kvinnor och män (punkt 49).

Systemets syfte

Bedömningen av en rimlig och objektiv motivering börjar med att identifiera ett legitimt mål. Det kan här vara värdefullt att notera om svaranden har angett olika syften i olika dokument; se avsnitt 3.4 *Vad är det algoritmiska systemet avsett att göra?*

Beakta olika typer av mål:

- ▶ Mål på lägre nivå (t.ex. att förutsäga bedrägerier, fastställa om en student befinner sig framför sin dator, förutsäga sannolikheten för att en enskild person inte kan återbetala ett lån).
- ▶ Mål på högre nivå (t.ex. att minimera statliga utgifter för bidrag till personer som inte är berättigade till dessa, övervakning vid digitala examinationer, att minimera den övergripande andelen fallissemang).

Det kan även vara nödvändigt att beakta den nationella rättsliga ram inom vilken det algoritmiska systemet infördes och användes. Se *Glukhin v. Ryssia*, ansökan nr 11519/20 (4 juli 2023) (nedan).⁵⁸

58. Se även *Beghal v. United Kingdom* för ytterligare diskussion om de krav som ställs på nationella lagar för att en ifrågasatt åtgärd ska betraktas som "i enlighet med lagen".

Lämplighet eller ändamålsenlighet

Svaranden bör visa att det algoritmiska systemet kan uppnå målet. Detta innebär att svaranden, när målet är att förbättra effektiviteten i en beslutsprocess, bör visa att algoritmen har förkortat handläggningstiden, minskat behovet av resurser eller på annat sätt ökat effektiviteten. När det gäller ett prediktivt system kan visad lämplighet innebära att systemets prediktiva träffsäkerhet kan påvisas. Observera att påståenden om lämplighet måste överensstämma med systemets syfte. Både bredare mål och mål på högre nivå bör beaktas. Ett system kan vara olämpligt om det inte uppfyller något av målen:

- ▶ mål på lägre nivå, exempelvis om systemet har utformats för att optimera något annat än det användningsområde som det faktiskt används för. Exempelvis om systemet har optimerats för att förutsäga fel men därefter används för att förutsäga bedrägerier;
- ▶ mål på högre nivå, exempelvis om målet är att minska utgifterna för felaktigt utbetalda bidrag, men de kontroller som stöds av algoritmen är dyrare än de besparingar som uppnås genom minskade utbetalningar. I så fall är det algoritmiska systemet inte lämpligt för just detta mål.

Nödvändighet

- ▶ Målet kan inte uppnås (lika effektivt) utan algoritmen:
 - jämfört med situationen innan algoritmen infördes;
 - jämfört med icke-algoritmiska metoder.
- ▶ Målet kan inte uppnås (lika effektivt) med ett annat algoritmiskt system:
 - jämfört med algoritmer som använder andra indata (exempelvis genom att utesluta de starkaste proxyvariablerna);
 - jämfört med andra utvecklingsdatamängder (med mindre inslag av historiska ojämlikheter eller mänskliga snedvridningar)

Exempel från Europeiska domstolen för de mänskliga rättigheterna: *Glukhin v. Ryssia*, ansökan nr 11519/20 (4 juli 2023)

Europeiska domstolen för de mänskliga rättigheterna beaktade användningen av live-ansiktsgenkänningsteknik (LFR) för att hitta och lokalisera en demonstrant i Rysslands tunnelbana under artikel 8 (privatliv) och artikel 10 (yttrandefrihet). Medan domstolen erkände vikten av moderna vetenskapliga utrednings- och identifieringstekniker (punkt 85) och identifierat brist på rättslig grund för de vidtagna åtgärderna på grund av brist på begränsning i lag av användningen av LFR, bedömde domstolen också "nödvändigheten i ett demokratiskt samhälle" av de vidtagna åtgärderna, i synnerhet med beaktande av deras potentiella "kylande effekt" på rätten till yttrandefrihet och mötesfrihet (artikel 10). Domstolen kom fram till:

90. Med beaktande av alla ovanstående överväganden drar domstolen slutsatsen att användningen av mycket inträngande ansiktsgenkänningsteknik i sammanhanget av den sökandes utövande av sin konventionsrätt till yttrandefrihet är oförenlig med de ideal och värden som präglar ett demokratiskt samhälle styrt av rättsstaten, vilka konventionen

var utformad för att upprätthålla och främja. Behandlingen av den sökandes personuppgifter med användning av ansiktsigenkänningsteknik inom ramen för förfaranden för administrativa brott – först för att identifiera honom utifrån fotografierna och videon som publicerats på Telegram och, för det andra, för att lokalisera och gripa honom medan han reste med Moskvas tunnelbana – kan inte betraktas som "nödvändig i ett demokratiskt samhälle".

Proportionalitet

I alla fall måste den diskriminering som orsakats av användningen av systemet som svaranden söker motivera vara proportionerlig i förhållande till det syfte som eftersträvas. Sammantaget kräver detta att en domstol väger skadans allvar mot de andra faktorer som beaktas i detta avsnitt.

Frågor att beakta

- ▶ Kan svaranden visa att någon diskriminering inte har skett?
- ▶ Kan svaranden visa att de inte ska anses ha blivit diskriminerade?
 - Löser förekomsten av en människa i beslutsprocessen (human in the loop) i tillräcklig grad eventuell kvarstående diskriminering?
 - Observera att diskrimineringen kan bero på mänskliga fel snarare än på det algoritmiska systemet.
- ▶ Kan svaranden hävda att fel jämförelseperson har valts?
- ▶ Kan svaranden visa att den klagande personen inte befinner sig i en liknande eller jämförbar situation som jämförelsepersonen?
- ▶ Finns det en hållbar alternativ förklaring till skillnaden i behandling som inte beror på diskriminering?
- ▶ Vilka åtgärder har vidtagits för att undvika diskriminering (skäliga anpassningar)?
- ▶ Kan svaranden, när det är relevant, visa att skillnaden i behandling inte grundas på den skyddade grunden utan på andra objektiva skillnader?
- ▶ Omfattas dessa objektiva skillnader uttryckligen av direktiven om likabehandling?
- ▶ Eftersträvar den aktuella bestämmelsen eller praxisen ett legitimt mål?
- ▶ Se avsnitt 3.4 om vad AI-systemet är avsett att göra och hur det gör detta – var de valda medlen lämpliga?
- ▶ Är de medel som används för att uppnå målet lämpliga, nödvändiga och proportionerliga i förhållande till det mål som eftersträvas?



Steg 4

Åtgärder

Steg 4 beskriver de olika åtgärder och handlingsalternativ som står till buds för EB:er. Befogenheter och mandat kommer sannolikt att skilja sig avsevärt mellan medlemsstaterna, och detta steg har därför utformats särskilt för att undersöka samspelet mellan standarddirektiven, som ska vara införlivade senast den 19 juni 2026, AI-förordningen, GDPR och annan relevant unionsrätt.

Det kommer vanligtvis att vara nödvändigt att överväga mer än en åtgärd, beroende på frågan eller ärendet.

Figur 5. Steg 4 – Åtgärder

STEG 4 Åtgärder		
4.1	Åtgärder för alla algoritmiska system	MÖJLIGA ÅTGÄRDER • Åtgärder enligt standarddirektiven • Inledande bistånd • Klagomål till MSA • GDPR-åtgärder • Alternativ tvistlösning • Rättegång • Insamling och rapportering av jämställdhetsdata
4.2	Ytterligare åtgärder för AI-system med hög risk	MÖJLIGA ÅTGÄRDER • MSA:s svar på allvarlig incident • MSA-utredning av produkt med allvarlig risk för grundläggande rättigheterna • Felaktiga AI-klassificeringar • Konforma AI-system som utgör risk grundläggande rättigheterna
4.3	Allvarlig risk för de grundläggande rättigheterna, men AI utgör inte en hög risk och är inte ett förbjudet system	MÖJLIGA ÅTGÄRDER • Begär ändring av AI-förordningen • Tillägg till bilaga III i AI-förordningen

Avsnitt 4.1 behandlar åtgärder för alla algoritmiska system.

Avsnitt 4.2 innehåller ytterligare åtgärder för högriskklassificerade system.

Avsnitt 4.3 beskriver vad ett EB bör göra i ett fall där ett system eller användningsområde innebär en risk för grundläggande rättigheter men inte klassificeras som högrisk enligt AI-förordningen, inklusive åtgärder som en nationell tillsynsmyndighet (MSA) kan vidta och som skulle kunna kräva eller innebära medverkan av ett EB.

Enligt antidiskrimineringsdirektiven och standarddirektiven ges EB:er breda befogenheter att genomföra aktiviteter för att förhindra diskriminering och främja likabehandling. Dessa befogenheter inkluderar medvetandehöjande arbete (artikel 5), tillhandahållande av stöd till offer (artikel 6), alternativ tvistlösning (artikel 7), genomförande av utredningar om brott mot principen om likabehandling (artikel 8), utfärdande av yttranden och beslut, inklusive rätt att fastställa fakta (artikel 9) samt agerande i domstolsförfaranden, inklusive rätt att lämna in synpunkter (artikel 10).

När diskriminering har identifierats i ett algoritmiskt system bör det övervägas att utöva dessa befogenheter för att säkerställa att:

- ▶ svaranden upphör med att använda AI-systemet på ett diskriminerande sätt, eller, om detta inte är möjligt, upphör med att använda AI-systemet helt och hållet;
- ▶ kunskap om den algoritmiska diskrimineringen kommuniceras på ett lämpligt sätt till berörda grupper och organ;
- ▶ råd och stöd, inklusive när det gäller alternativ tvistlösning (*alternative dispute resolution, ADR*) och rättsliga processer, tillhandahålls enskilda personer som utsatts för diskriminering,

- ▶ avhjälpande åtgärder vidtas, bland annat genom tillhandahållande av utbildning, rådgivning och stöd;
- ▶ en offentlig debatt om frågan om diskriminering i algoritmiska system främjas och att positiva åtgärder samt integrering av ett likabehandlingsperspektiv (*equality mainstreaming*) främjas bland offentliga och privata aktörer;
- ▶ kommunikation med relevanta intressenter, inklusive arbetsmarknadens parter, upprätthålls och utbyte av god praxis främjas.

Vid genomförandet av sådana aktiviteter kan EB beakta särskilda situationer av missgynnande som följer av intersektionell diskriminering, vilket förstås som diskriminering som grundas på en kombination av diskrimineringsgrunder som skyddas enligt direktiven om likabehandling.

EB:er måste beakta lämpliga kommunikationsverktyg och kommunikationsformat för varje målgrupp. De ska särskilt fokusera på grupper vars tillgång till information kan hindras, exempelvis på grund av osäker ekonomisk situation, ålder, funktionsnedsättning, läs- och skrivkunnighet, nationalitet eller uppehållsstatus, eller på grund av bristande tillgång till digitala verktyg.

Viktiga resurser

- ▶ *Strategic litigation handbook* (Equinet 2017)
- ▶ *How to use the Artificial Intelligence Act to investigate AI bias and discrimination: a guide for equality bodies* (Mittelstadt 2024)

4.1. Åtgärder för alla algoritmiska system

Genom att följa processen i steg 2 och 3 bör man ha kunnat fastställa om diskriminering förekommer och vad orsaken till diskrimineringen kan ha varit. EB:er bör använda den information som samlats in under dessa faser för att föreslå åtgärder och anta en plan för att avskaffa eller ändra ifrågasatta rutiner, algoritmer eller system. EB kan till exempel ha konstaterat att valet av rättvisemått var olämpligt för användningsfallet eller att uppgifterna inte var tillräckligt representativa.

Om EB:et beslutar att risken för diskriminering eller den identifierade diskrimineringen är olaglig men kan åtgärdas, bör man överväga åtgärder för att uppnå:

- ▶ förbud mot användning av systemet tills diskrimineringen har åtgärdats;
- ▶ eventuella korrigerande åtgärder som bör vidtas.

Om likabehandlingsorganet beslutar att risken för diskriminering är olaglig och inte kan åtgärdas, bör rättsliga åtgärder övervägas för att uppnå ett av följande resultat:

- ▶ förbud mot användning av det algoritmiska systemet helt och hållet;
- ▶ förbud mot användning av det algoritmiska systemet inom en viss sektor;
- ▶ förbud mot användning av det algoritmiska systemet för ett visst ändamål.

Sådana rättsliga åtgärder I skulle kunna genomföras genom att likabehandlingsorganen utövar sina egna befogenheter när det gäller att avge yttranden, lämna

rekommendationer och genomföra samråd, i samverkan med MSA, antingen enligt GPSR eller AI-förordningen eller genom domstolarna.

4.1.1. EB:er befogenheter att vidta rättsliga åtgärder

EB:er bör först överväga sina befogenheter att vidta rättsliga åtgärder enligt nationell rätt, enligt ovan. Dessa kan omfatta några av eller samtliga rättsliga åtgärder som ska formaliseras genom standarddirektiven, såsom att lämna ett yttrande eller en rekommendation eller att kräva samråd eller genomföra testning.

Om standarddirektiven införlivas i nationell rätt bör EB:et tilldelas befogenheter att meddela rättelseåtgärder. Dessa kommer att inkludera artikel 9 (utfärdande av ett bindande eller icke-bindande yttrande, beroende på nationell kontext) och artikel 15, konsultations- och rekommendationsbefogenheter.

EB:er bör också ha befogenhet att tillhandahålla och dokumentera sin bedömning av fallet, inklusive fastställande av fakta och en motiverad slutsats om förekomsten av diskriminering. I fallet med algoritmisk diskriminering kan detta inkludera samarbete med MSA:et.

Där det är lämpligt ska både icke-bindande yttranden och bindande beslut inkludera specifika åtgärder för att avhjälpa varje överträdelse av principen om likabehandling som fastställts och för att förhindra framtida förekomster. EB:er kan behöva experthjälp för att tillhandahålla specifika åtgärder och om de inte har expertisen kan de vilja rådfråga MSA. Lämpliga mekanismer för uppföljning av icke-bindande yttranden bör inrättas, såsom skyldigheter att ge återkoppling, och för verkställighet av bindande beslut.

Konsultationer kan användas för att identifiera förekomsten av problem, för att testa de sociala och diskriminerande konsekvenserna av system på drabbade eller potentiellt drabbade samhällen och för att samla in bevis.

Mer generellt kan rekommendationer och samråd inriktas på att säkerställa att problemen inte uppstår i framtiden, exempelvis genom rekommendationer om:

- ▶ konsekvensbedömningar av dataskydd och likabehandling;
- ▶ samråd med berörda grupper och en samhällsdriven bedömning av konsekvenser för grundläggande rättigheter (*FRIA – Fundamental Rights Impact Assessment*);
- ▶ användning av sandlådemiljöer (*sandboxes*) och tester i verkliga miljöer.

Ytterligare läsning

- ▶ Crowley (2023), "Informing the policy agenda: equality bodies making recommendations."

4.1.2. Ge inledande stöd till en person som utsatts för diskriminering

Artikel 6 i standarddirektiven anger att EB ska ge inledande stöd genom att informera personer som utsatts för diskriminering om följande:

- ▶ den rättsliga ramen för diskriminering och algoritmiska system, inklusive rådgivning som är anpassad till deras specifika situation;
- ▶ de tjänster som erbjuds av EB:et och relaterade processuella aspekter;
- ▶ tillgängliga rättsmedel, inklusive möjligheten att driva ärendet i domstol, bland annat:
 - enligt diskrimineringslagstiftningen;
 - enligt AI-förordningen – se exempelvis rätten att lämna in ett klagomål till den nationella tillsynsmyndigheten (MSA) (artikel 85 i AI-förordningen) och rätten till en förklaring (artikel 86 i AI-förordningen);
 - enligt dataskyddsrätten – exempelvis rätten att lämna in ett klagomål till dataskyddsmyndigheten (DPA) (artikel 77 i GDPR) och rätten att göra en begäran om tillgång till personuppgifter (artikel 15 i GDPR);⁵⁹
- ▶ tillämpliga regler om konfidentialitet och skydd av personuppgifter (exempelvis EB:ens dataskyddspolicyer och, när så är relevant, eventuella konflikter mellan att EB:et inhämtar bevisning inom ramen för utövandet av en del av sitt mandat och att kunna tillhandahålla denna bevisning för att en enskild person ska kunna använda den i domstol).
- ▶ möjligheten att få psykologiskt stöd eller andra relevanta former av stöd från andra organ eller organisationer.

EB:er ska inom rimlig tid informera den klagande om huruvida klagomålet kommer att avslutas eller om det finns skäl att gå vidare med ärendet.

Efter att inledande stöd har lämnats kan de rättsmedel som står till buds för en enskild klagande omfatta:

- ▶ att lämna in ett klagomål till MSA (artikel 85 i AI-förordningen);
- ▶ att överväga alternativ tvistlösning;
- ▶ att väcka talan om diskriminering vid nationella domstolar.

Alla beslut om vilken handlingsväg som är bäst bör fattas tillsammans med den berörda personen eller gruppen. Om EB:et saknar kapacitet, kompetens (exempelvis i fråga om dataskyddsanspråk) eller resurser för att driva den åtgärd som den berörda personen efterfrågar, bör ärendet hänvisas till en annan behörig juridisk tjänsteleverantör.

4.1.3. Bistå en juridisk eller fysisk person med att lämna in ett klagomål till MSA

Om utredningen enligt steg 3 har identifierat brister i efterlevnaden av AI-förordningen har en juridisk eller fysisk person rätt att lämna in ett klagomål till den nationella tillsynsmyndigheten (artikel 85 i AI-förordningen). Detta innebär att rätten även

59. Rätten till en förklaring föreskrivs både i GDPR (artiklarna 15 och 22) och i AI-förordningen (artikel 86). I [steg 2](#) och bilagorna [A](#) och [B](#) finns ytterligare förklaringar och mallar för hur man lämnar in en begäran.

omfattar ett organ med juridisk personlighet, såsom ett företag eller en organisation. EB:et kan bistå den klagande personen med den bevisning som identifierats i steg 3.

Denna rätt påverkar inte befintliga rättigheter och rättsmedel enligt unionsrätten och nationell rätt (skäl 170). Den är inte begränsad till AI-system med hög risk.

Ett sådant klagomål ska "beaktas" vid genomförandet av marknadskontrollrättsmedel enligt förordning (EU) 2019/1020 om marknadskontroll och överensstämmelse för produkter. Tidsfrister kommer sannolikt att fastställas i nationell lagstiftning (se artikel 11 i förordning (EU) 2019/1020).

AI-system som inte är högriskklassificerade omfattas av den allmänna produktsäkerhetsförordningen (EU) 2023/988, förordning (EU) 2023/988 och marknadskontrollförordningen (EU) 2019/1020.

Det är viktigt att AI-system som är kopplade till produkter och som enligt denna förordning inte klassificeras som högrisk och därför inte behöver uppfylla de krav som gäller för högrisk-AI-system ändå är säkra när de släpps ut på marknaden eller tas i bruk. För att bidra till detta mål ska Europaparlamentets och rådets förordning (EU) 2023/988 (53) fungera som ett skyddsnät (skäl 166 i AI-förordningen).

Rätten att lämna in ett klagomål påverkar inte andra administrativa eller rättsliga åtgärder och bör därför användas tillsammans med andra tillämpliga rättsmedel. Enbart denna åtgärd är sannolikt inte den bästalösningen för en enskild person, inte minst eftersom det inte finns någon bestämmelse om skadestånd. Den är dock en viktig kompletterande åtgärd och bör övervägas i lämpliga fall, eftersom den nationella tillsynsmyndigheten har mycket omfattande befogenheter och åtgärdsalternativ enligt förordning (EU) 2019/1020. Ett klagomål till den nationella tillsynsmyndigheten bör leda till en mer strukturell utredning av myndigheten och till bredare åtgärder som påverkar en större grupp berörda personer än ett enskilt diskrimineringsärende. Den nationella tillsynsmyndigheten har omfattande befogenheter enligt både AI-förordningen och förordning (EU) 2019/1020 och kan genomföra dokumentkontroller, fysiska kontroller och laboratoriekontroller (se avsnitt 4.2).

Om klagomålet bifalles kan den nationella tillsynsmyndigheten kräva att en ekonomisk aktör vidtar korrigerande åtgärder, under förutsättning att åtgärden är lämplig och proportionerlig. Om de korrigerande åtgärderna inte vidtas eller inte får avsedd effekt kan myndigheten förbjuda eller begränsa att en produkt tillhandahålls på marknaden eller besluta att produkten ska dras tillbaka eller återkallas.

För produkter som innebär en risk och vars konsekvenser sträcker sig utanför utanför den egna medlemsstaten ska MSA skyldig att omedelbart underrätta Europeiska kommissionen och de övriga EU-medlemsstaterna om de åtgärder som vidtagits. Förordning (EU) 2019/1020 innehåller omfattande bestämmelser om gränsöverskridande samarbete mellan medlemsstaternas myndigheter.

4.1.4. Åtgärder för registrerade enligt GDPR eller nationella dataskyddsregler

Enligt GDPR har en enskild registrerad person (en fysisk person) följande rättigheter när det gäller personuppgifter, utöver rätten till en förklaring:

- ▶ rätten att invända mot automatiserad behandling av personuppgifter (artikel 21);
- ▶ rätten till rättelse av felaktiga personuppgifter som rör personen och att få ofullständiga personuppgifter kompletterade genom att lämna en kompletterande uppgift (artikel 16);
- ▶ rätten till radering av personuppgifter utan onödigt dröjsmål när något av följande gäller:
 - samtycket har återkallats och det saknas annan rättslig grund för behandlingen;
 - personen har utövat sin rätt att invända enligt artikel 21.1;
 - personuppgifterna har behandlats på ett olagligt sätt.

Dessa rättigheter kan vara relevanta när ett algoritmiskt system har lett till diskriminering. Eventuella åtgärder kan kräva medverkan av en dataskyddsexpert eller dataskyddsmyndigheten.

Om medlemsstater har antagit nationella dataskyddsregler utöver GDPR, eller om det gäller stater som inte omfattas av GDPR, är det viktigt att beakta relevant nationell dataskyddslagstiftning.

4.1.5. Erbjuda alternativ tvistlösning i enlighet med lag

Enligt standarddirektiven ska EB:er enligt artikel 7 kunna erbjuda parterna möjlighet att söka en alternativ lösning på sin tvist. Processen kan ledas av EB:et självt eller av ett annat behörigt organ i enlighet med nationell rätt och praxis. I det senare fallet kan EB:et lämna synpunkter till detta organ. Sådana alternativa tvistlösningar kan ta olika former, exempelvis medling eller förlikning, i enlighet med nationell rätt och praxis. Att en lösning inte nås ska inte hindra parterna från att utöva sin rätt att väcka talan i domstol. Medlemsstaterna ska säkerställa att det finns en tillräckligt lång preskriptionstid för att garantera att parterna i en tvist har tillgång till domstol, exempelvis genom att preskriptionstiden avbryts eller skjuts upp medan parterna deltar i ett alternativt tvistlösningsförfarande.

4.1.6. Diskriminerings- och människorättsprocesser

EB:er bör ha rätt att agera i domstolsförfaranden i civilrättsliga och förvaltningsrättsliga frågor, inklusive rätt att lämna synpunkter som rör direktiven om likabehandling i enlighet med nationell rätt och praxis. Detta kommer att variera mellan de enskilda medlemsstaterna.

Enligt artikel 10 i standarddirektiven ska EB:ets rätt att agera i domstolsförfaranden även omfatta åtminstone en av följande möjligheter:

- ▶ rätten att inleda domstolsförfaranden på en eller flera utsatta personers vägnar;
- ▶ rätten att delta i domstolsförfaranden till stöd för en eller flera utsatta personer;
- ▶ rätten att inleda domstolsförfaranden i eget namn för att tillvarata allmänintresset.

Detta kommer även att omfatta verkställighetsbeslut avseende bindande utlåtanden som utfärdats av EB:et enligt artikel 9 (se nedan). Vid övervägande av om strategiska rättsprocesser bör inledas hänvisas EB:er till de utförliga överväganden som anges i Equinets "Strategic litigation handbook" (Equinet 2017:17).

Det finns sannolikt betydande möjligheter att bedriva strategiska rättsprocesser när det gäller AI-system, eftersom både tekniken och rättsområdet är relativt nya. Strategiska rättsprocesser är dock inte nödvändigtvis den bästa metoden i varje enskilt fall. Det kan exempelvis ta för lång tid att nå ett resultat för en berörd person, den information som krävs för att väcka en framgångsrik talan kan saknas, eller det kan vara mer effektivt att inleda testnings- och förhandlingsprocesser. Det finns även andra faktorer som behöver beaktas från fall till fall.

4.1.7. Datainsamling, tillgång till jämlikhetsdata och rapportering

Tillgång till jämlikhetsdata är centralt för att fastställa diskriminering genom AI- och ADM-system. EB:er kan utnyttja de nya befogenheterna för att samla in data om förekomsten av algoritmisk diskriminering på en bredare basis. Rekommendationer om de typer av information som ska registreras har gjorts genomgående i denna metodik, vilket kan bistå EB:er i att sammanställa jämlikhetsdata som rör algoritm-system. Insamlad statistik kan vara till hjälp när man fastställer ett fall av prima facie-diskriminering i domstol, till exempel. Om EB:et noterar att det finns frekventa förfrågningar om stöd i förhållande till en viss sektor eller ett användningsfall av ett algoritmsystem kan det kräva ytterligare utredning eller prioritering.

Insamling av jämlikhetsdata är otillräcklig i många länder, inte minst eftersom den ofta är lagligt begränsad eftersom EB:er ofta inte kan behandla känsliga kategorier av personuppgifter för bias-mätning och begränsningsändamål enligt artikel 9.2 g i GDPR och artikel 10.5 i AI-förordningen. EB:er kan rekommendera reformering av datainsamlingsregler och praxis för att förbättra den befintliga kunskapen om algoritmisk diskriminering (se Europarådet 2025b).

Att identifiera strukturella problem och rapportera om dessa på årsbasis

Datainsamling bör möjliggöra för EB:et att identifiera och rapportera om strukturell ojämlikhet i algoritmsystem. Då EB:et använder en databas för att registrera detaljer om ärenden och rådgivning bör en ny ärendekategori läggas till för att möjliggöra för EB:et att spåra den faktiska eller misstänkta diskriminerande användningen av AI-system. På längre sikt kan denna metodik användas för att utforma specifika fält i databasen för att möjliggöra insamling och analys av data om dessa ärenden.

Som inspiration identifieras genomgående i steg 3 av metodiken indikatorer på den typ av information som kan registreras i en databas i slutet av varje avsnitt och ett förslag på arbetsblad tillhandahålls i [bilaga K](#).

Strukturella problem kan inkludera:

- ▶ bevis för diskriminering som uppstår i en viss typ av algoritmsystem eller teknik;
- ▶ bevis för diskriminering som uppstår när algoritmsystem används i särskilda sektorer;
- ▶ bevis för misstanke om att diskriminering uppstår i en sektor eller en typ av algoritmsystem;

- ▶ bevis för att grupper missgynnas oproportionerligt genom införandet av algoritmsystem över flera sektorer;
- ▶ AI-system eller användningsfall som borde, men för närvarande inte är:
 - förbjudna;
 - klassificerade som högrisk;
 - modeller för allmänna ändamål med systemrisk;
 - listade under någon av dessa kategorier men likväl utgör en oacceptabel risk för diskriminering;
- ▶ systematiska brister i efterlevnaden av AI-förordningen som är relevanta för diskriminering, såsom otillräcklig bias- eller rättvisetestning och otillräckliga datainsamlingsrutiner.

4.2. Ytterligare åtgärder för AI-system med hög risk

4.2.1. Åtgärder som vidtas av MSA i samarbete med EB:et

Utöver de befogenheter som ges till EB:er enligt artikel 77 i AI-förordningen i sig självt har MSA skyldigheter att samarbeta med ett EB som ett artikel 77-organ när man hanterar en "allvarlig incident" (artikel 73 i AI-förordningen), när AI-system utgör risker (artikel 79 i AI-förordningen) och när man utreder ett system som uppfyller kraven men utgör en risk för grundläggande rättigheter (artikel 82 i AI-förordningen). En MSA kan också behöva konsultera ett EB när man beaktar ett AI-system som felaktigt klassificerats av en leverantör som icke-högrisk (artikel 80 i AI-förordningen). Dessa processer diskuteras mer i detalj nedan.

Ett EB kan vilja hänvisa ett system till MSA med bevis som erhållits enligt denna metodik för att rapportera en allvarlig incident, för att genomföra testning, för att utföra en utvärdering eller för att utöva befogenheter enligt antingen AI-förordningen eller MSA-förordningen för att kräva korrigerande åtgärder avstängning eller återkallelse av produkten från marknaden eller andra åtgärder.

MSA:s svar på en anmälan om en allvarlig incident (artikel 73 i AI-förordningen)

MSA har betydande befogenheter och ansvar enligt förordning (EU) 2019/1020 när man hanterar en allvarlig incident.

MSA måste informera artikel 77-organet när det underrättas av en leverantör om en allvarlig incident (artikel 73.3 i AI-förordningen). En allvarlig incident definieras som en incident eller fel i ett AI-system som till exempel direkt eller indirekt leder till överträdelse av skyldigheter enligt unionslagstiftning som är avsedd att skydda grundläggande rättigheter, inklusive rätten till likabehandling (artikel 3.49 c i AI-förordningen). Vägledning kommer att utfärdas av kommissionen i god tid.

I sådana fall, beakta följande:

- ▶ Inkluderade det allvarliga incidenter en risk för diskriminering?
 - Om ja, bör de drabbade personerna ha rätt till en korrigerande åtgärd?

- Om ja, överväg den mest effektiva korrigerande åtgärden (beakta till exempel ADR, rättsprocess eller en utredning).
- Om de drabbade personerna har rätt till monetär ersättning, säkerställ att den valda åtgärden inte utesluter detta. Säkerställ att alla relevanta tidsfrister följs.
- ▶ Är det troligt att detta har ytterligare och pågående konsekvenser för andra individer? Till exempel, används systemet av banker för att avgöra kreditvärdighet?
- ▶ Vidta omedelbara åtgärder för att höja medvetenheten om frågan (artikel 5 i standarddirektiven eller via MSA med stöd av produktsäkerhetsförordningar).
- ▶ Överväg om AI-systemet interagerar med andra enheter/mjukvara och höj medvetenheten om eventuella liknande påverkade system.

Dokumentera typen av teknik och den sektor där frågan uppstod för användning i EB:ens rapportering och analys av jämlikhetsdata.

Den korrigerande åtgärd som krävs kommer att bero på de identifierade problemen. Dock bör den information, de steg och tillvägagångssätt som beskrivs i steg 3 bistå EB:et i att identifiera potentiella korrigerande åtgärder.

MSA:s utredning enligt artikel 79.3 i AI-förordningen: produkt som utgör en allvarlig risk för grundläggande rättigheter, inklusive diskriminering

Ifall MSA:t (ensam eller i samarbete med ett artikel 77-organ) finner att AI-systemet inte uppfyller de krav som fastställts i förordningen ska den kräva, utan onödigt dröjsmål:

- ▶ lämpliga korrigerande åtgärder för att få AI-systemet att uppfylla kraven;
- ▶ att informera kommissionen och andra medlemsstater om resultatet av utvärderingen och de åtgärder som leverantören ska vidta;
- ▶ att återkalla AI-systemet från marknaden (artikel 79.5 i AI-förordningen);
- ▶ en återkallelse inom en period som MSA kan föreskriva, inom 15 arbetsdagar eller som föreskrivs i unionslagstiftning (detta kommer att kräva en kontroll av relevant lagstiftning vid detta tillfälle);
- ▶ att MSA informerar det relevanta anmälda organet.

MSA:s förfarande för att hantera AI-system som klassificerats av leverantören som icke-högrisk enligt tillämpningen av bilaga III (artiklarna 6.4 och 80 i AI-förordningen)

Om man beaktar om ett AI-system har felaktigt klassificerats enligt undantagsförfarandet från självbedömning (artikel 6.4 och förfarandet i artikel 80) måste MSA genomföra en utvärdering för att avgöra om artikel 6.3 tillämpas.

I detta fall är det möjligt att det inte finns någon relevant bedömning av grundläggande rättigheter och inget beaktande av diskriminering. De frågor som beaktas är om AI-systemet:

- ▶ utför en snäv proceduruppgift;
- ▶ förbättrar resultatet av tidigare genomförd mänsklig aktivitet;

- ▶ upptäcker beslutsfattandemönster eller avvikelser från tidigare beslutsfattandemönster;
- ▶ utför en förberedande uppgift för en bedömning.

Det är dock möjligt att även dessa system kan utgöra risker för grundläggande rättigheter och likabehandling. EB:er kan hänvisa ett system till MSA för att beakta om AI-systemet har felaktigt klassificerats och tillhandahålla de bevis som de baserar hänvisningen på.

AI-system som uppfyller kraven och som utgör en risk för personers hälsa eller säkerhet, grundläggande rättigheter eller för andra aspekter av skyddet av allmänintresset (artikel 82 i AI-förordningen)

MSA har befogenheter beträffande ett system som uppfyller kraven men utgör en risk enligt artikel 82 i AI-förordningen. Innan dessa steg vidtas måste MSA ha både:

- ▶ genomfört en utvärdering enligt artikel 79 i AI-förordningen; och
- ▶ konsulterat det relevanta artikel 77-organet.

Den aktör som är ansvarig kommer att krävas att vidta "alla lämpliga åtgärder för att säkerställa att det berörda AI-systemet, när det släpps ut på marknaden eller tas i bruk, inte längre utgör den risken".

Även om det inte uttryckligen anges i AI-förordningen bör EB:et konsulteras åtminstone när MSA söker avgöra om riskerna för grundläggande rättigheter och/eller jämlikhet framgångsrikt har avlägsnats.

4.3. Vad bör man göra när ett system eller användningsfall utgör en risk för grundläggande rättigheter men inte klassificeras som "högrisk" eller "förbjudet"

Denna situation kan uppstå när:

- ▶ en leverantör har självbedömt sitt algoritmsystem som inte utgör något AI-system enligt definitionen i artikel 3 och det därför ligger utanför förordningens tillämpningsområde helt och hållet;
- ▶ ett algoritmsystem som borde förbjudas på grund av den risk det utgör inte ingår i förteckningen över förbjudna system i artikel 5 i AI-förordningen;
- ▶ ett algoritmsystem som borde klassificeras som högrisk inte ingår i bilaga III i AI-förordningen eftersom antingen sektorn eller användningsfallet inte ingår;
- ▶ en leverantör har tillämpat undantaget i artikel 6.3 på ett AI-system som annars skulle klassificeras som högrisk.

Observera: användare har en skyldighet att övervaka sina högrisk-system och underätta MSA där de har skäl att anse att användningen av AI-systemet i enlighet med instruktionerna kan leda till att systemet utgör en allvarlig risk för grundläggande rättigheter (artikel 26 i förening med artikel 79 i AI-förordningen).

Observera: om ett AI-system är en "produkt som utgör en risk" som definieras i artikel 3 punkt 19 i förordning (EU) 2019/1020 är MSA skyldig att utvärdera systemet och, där risker för grundläggande rättigheter identifieras, att samarbeta med de relevanta artikel 77-myndigheterna. Korrigerande åtgärder för en sådan produkt enligt AI-förordningen är begränsade till kompatibla högrisk-system. Alternativa korrigerande åtgärder enligt MSA-förordningen och den allmänna produktsäkerhetslagstiftningen, avtalsrätt eller skadeståndsrätt, diskrimineringslagstiftning eller människorättslagstiftning kan vara tillgängliga.

För enskilda personer eller berörda grupper innebär det faktum att ett system inte klassificeras som högrisk eller förbjudet inte att det inte kan leda till diskriminering eller resultera i, eller användas på ett sätt som kränker de grundläggande rättigheterna. De berörda grupperna har tillgång till alla de vanliga rättsmedlen mot diskriminering som föreskrivs i direktiven för likabehandling och i nationell lagstiftning.

4.3.1. Föreslå ändringar av AI-förordningen

Om EB:et konstaterar att en AI-praxis står i grundläggande strid med unionens värden om respekt för människans värdighet, frihet, jämlikhet, demokrati, rättsstatsprincipen och de grundläggande rättigheterna enligt stadgan, bör framställningar göras till MSA och Europeiska kommissionen med begäran om att AI-förordningen ändras. AI-förordningen innehåller ett förfarande genom vilket ändringar av förordningen kan göras under olika omständigheter.

Enligt artikel 112 ska kommissionen varje år bedöma behovet av att ändra förteckningen i bilaga III (AI-system med hög risk) och artikel 5 (förbjudna AI-system) samt lägga fram resultaten av denna bedömning för Europaparlamentet. Det är därför viktigt att, i enlighet med de förfaranden som anges i AI-förordningen, uppmärksamma kommissionen på AI-system som bör förbjudas eller AI-system med hög risk inom områden eller användningsfall som inte omfattas av bilaga III.

Under de första fem åren av AI-förordningens tillämpning (till och med den 1 augusti 2029, med automatisk förlängning därefter om ingen invändning görs) har kommissionen befogenhet att anta delegerade akter under följande omständigheter:⁶⁰

Artikel 6.6	Att lägga till, ändra eller ta bort villkor för egenbedömning av att inte utgöra hög risk
Artikel 7	Att ändra bilaga III genom att lägga till eller ändra användningsfall (men inte områden) samt ta bort system från förteckningen
Artikel 11.3	Att ändra den tekniska dokumentation som krävs enligt bilaga IV (särskilt mot bakgrund av den tekniska utvecklingen)

60. Kommissionen har delegerade befogenheter när det gäller GPAI, men detta faller utanför tillämpningsområdet.

Artikel 47.5	Att uppdatera innehållet i EU-försäkran om överensstämmelse i bilaga V
--------------	--

Även om kommissionen inte har befogenhet att anta en delegerad akt för att ändra områdesrubriker i bilaga III är den skyldig att bedöma behovet av eventuella sådana ändringar senast den 2 augusti 2028 av befintliga områdesrubriker eller tillägg av nya områdesrubriker i bilaga III, ändringar av listor över system som kräver ytterligare transparens och ändringar för att förbättra effektiviteten hos övervaknings-/styrningssystemet. Kommissionen är sedan skyldig att lämna in sin bedömning till Europaparlamentet och rådet.

AI-byrån kommer att utarbeta en metod för bedömningen av dessa kriterier.

4.3.2. AI-system som används inom ett högriskområde för ett användningsfall som inte anges i bilaga III till AI-förordningen

Om ett visst "användningsfall" inte finns upptaget i bilaga III till AI-förordningen har kommissionen enligt artikel 7 befogenhet att ändra bilagan genom att lägga till eller ändra användningsfall när:

- ▶ systemet är avsett att användas inom ett område som omfattas av bilaga III; och
- ▶ den negativa påverkan på grundläggande rättigheter (inklusive diskriminering) är likvärdig med eller större än risken för skada eller annan negativ påverkan från de system med hög risk som redan omfattas av bilaga III.

Om ett användningsfall som uppfyller dessa kriterier av EB:et bedöms ha en negativ inverkan på de grundläggande rättigheterna men inte finns upptaget i förteckningarna i bilaga III, bör EB:et överväga att rapportera användningsområdet till kommissionen så att det kan läggas till i förteckningen enligt artikel 7 i AI-förordningen.

De faktorer som kommissionen ska beakta anges i artikel 7.2 i AI-förordningen och bör behandlas i en sådan rapport:

- ▶ AI-systemets avsedda ändamål;
- ▶ i vilken utsträckning AI-systemet har använts eller sannolikt kommer att användas;
- ▶ arten och omfattningen av de uppgifter som behandlas och används av AI-systemet, särskilt om särskilda kategorier av personuppgifter behandlas;
- ▶ i vilken utsträckning AI-systemet agerar autonomt samt möjligheten för en människa att åsidosätta ett beslut eller en rekommendation som kan leda till potentiell skada;
- ▶ i vilken utsträckning användningen av AI-systemet redan har orsakat skada på hälsa eller säkerhet, haft en negativ inverkan på grundläggande rättigheter eller gett upphov till betydande farhågor om sannolikheten för sådan skada eller negativ påverkan, vilket exempelvis framgår av rapporter eller dokumenterade anmälningar som lämnats till behöriga nationella myndigheter eller, när så är lämpligt, av andra rapporter;

- ▶ den potentiella omfattningen av sådan skada eller negativ påverkan, särskilt när det gäller dess allvarlighetsgrad och dess förmåga att påverka ett stort antal personer eller att i oproportionerligt hög grad påverka en viss grupp personer;
- ▶ i vilken utsträckning de personer som riskerar att skadas eller påverkas negativt är beroende av det resultat som AI-systemet genererar, särskilt när det av praktiska eller rättsliga skäl inte rimligen är möjligt att avstå från detta resultat;
- ▶ i vilken utsträckning det råder en maktobalans, eller de personer som riskerar att skadas eller påverkas negativt befinner sig i en utsatt ställning i förhållande till användaren, särskilt på grund av status, auktoritet, kunskap, ekonomiska eller sociala förhållanden eller ålder;
- ▶ i vilken utsträckning det resultat som AI-systemet genererar enkelt kan rättas eller återställas, med beaktande av de tekniska lösningar som finns tillgängliga för att rätta eller återställa resultatet. Resultat som har en negativ inverkan på hälsa, säkerhet eller grundläggande rättigheter ska dock inte anses vara enkla att rätta eller återställa;
- ▶ storleken på och sannolikheten för de fördelar som användningen av AI-systemet medför för enskilda personer, grupper eller samhället i stort, inklusive möjliga förbättringar av produktsäkerheten;
- ▶ i vilken utsträckning unionsrätten redan innehåller:
 - effektiva rättsmedel mot de risker ett AI-system medför, med undantag för skadeståndsanspråk;
 - effektiva rättsmedel för att förebygga eller väsentligt minska dessa risker.

Referenser

Litteratur

Adams-Prassl J., Binns R. och Kelly-Lyth A. (2022), "Directly discriminatory algorithms", *The Modern Law Review*, 86(1), s. 144-175. Tillgänglig på: <https://doi.org/10.1111/1468-2230.12759> (hämtad 17 oktober 25).

Ahmed M. (2020), "UK passport photo checker shows bias against dark-skinned women", *BBC News*. Tillgänglig på: www.bbc.co.uk/news/technology-54349538.amp (hämtad 13 oktober 2025).

Allen R. och Masters D. (2019), "Artificial intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making", *Journal of the Academy of European Law*. DOI: 10.1007/s12027-019-00582-w (hämtad 19 juni 2025).

Allen R. och Masters D. (2020), "Regulating for an equal AI: a new role for equality bodies" *Equinet*. Tillgänglig på: https://equineteurope.org/publications/regulating-for-an-equal-ai-a-new-role-for-equality-bodies/?__cf_chl_f_tk=3KwuXAuyY1rQSy6-KKjf6WQaiw8Fzq0tVZpxUk2K39Qs-1782916549-1.0.1.1-nTAmM4fI0DMeHv8UOjuY4hye5_yL66T80oAESRgXorA (hämtad 20 oktober 2025).

Algorithm Audit (2024a), "Preventing prejudice". Tillgänglig på: https://algorithmaudit.eu/algoprudence/cases/aa202401_preventing-prejudice/ (hämtad 11 oktober 2025).

Algorithm Audit (2024b), "Addendum to preventing prejudice". Tillgänglig på: https://algorithmaudit.eu/algoprudence/cases/aa202402_preventing-prejudice_addendum/ (hämtad 1 oktober 2025)

AlgorithmWatch (2022), "Automated decision-making systems and discrimination: A guidebook for anti-discrimination counselling". Tillgänglig på: <https://algorithmwatch.org/en/autocheck/> (hämtad 29 september 2025).

Amnesty International (2021), "Xenophobic machines: discrimination through unregulated use of algorithms in the Dutch childcare benefits scandal". Tillgänglig på: <https://www.amnesty.org/en/documents/eur35/4686/2021/en/> (hämtad 13 oktober 2025).

Article 29 Data Protection Working Party (2017a), "Guidelines on data protection impact assessment (DPIA) and determining whether processing is 'likely to result in a high risk for the purposes of Regulation 2016/679'". Tillgänglig på: <https://ec.europa.eu/newsroom/article29/items/611236/en> (hämtad 8 november 2025).

Article 29 Data Protection Working Party (2017b), "Opinion on some key issues of the law enforcement directive (EU 2016/680) – wp258", European Commission. Tillgänglig på: <https://ec.europa.eu/newsroom/article29/items/610178/en> (hämtad 17 oktober 2025).

Article 29 Data Protection Working Party (2018a), "Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679

(WP251rev.01, 17/EN)”, European Commission. Tillgänglig på: <https://ec.europa.eu/newsroom/article29/items/612053/en> (hämtad 10 november 2025).

Article 29 Data Protection Working Party (2018b), “Guidelines on transparency under Regulation 2016/679 (WP260, 17/EN)”, European Commission. Tillgänglig på: www.edpb.europa.eu/our-work-tools/our-documents/article-29-working-party-guidelines-transparency-under-regulation_en (hämtad 10 november 2025).

Autoriteit Persoonsgegevens (2024), Final recommendation on supervision of AI: sector and centrally coordinated. Tillgänglig på: www.autoriteitpersoonsgegevens.nl/en/current/final-recommendation-on-supervision-of-ai-sector-and-centrally-coordinated (hämtad 1 oktober 2025).

Barocas S., Hardt M. och Narayanan A. (2023), Fairness and machine learning: limitations and opportunities, MIT Press, Cambridge, MA.

Barros Vale S. och Zanfir-Fortuna G. (2022), “Automated decision-making under the GDPR: practical cases from courts and data protection authorities”, Future of Privacy Forum. Tillgänglig på: <https://fpf.org/blog/fpf-report-automated-decision-making-under-the-gdpr-a-comprehensive-case-law-analysis/> (hämtad 1 oktober 2025).

Big Brother Watch (2025), “Suspicion by design”. Tillgänglig på: <https://bigbrother-watch.org.uk/campaigns/welfare-data-watch/> (hämtad 10 november 2025).

Bird K. A., Castleman B. L. och Song Y. (2024), “Are algorithms biased in education? Exploring racial bias in predicting community college student success”, Journal of Policy Analysis and Management, Vol. 44 (2), Tillgänglig på: <https://onlinelibrary.wiley.com/doi/full/10.1002/pam.22569> (hämtad 17 oktober 2025).

Boscarol F. (2022), “Costly birthplace: discriminating insurance practice”, AlgorithmWatch. Tillgänglig på: <https://algorithmwatch.org/en/discriminating-insurance/> (hämtad 13 oktober 2025). Ceross A. (2025), Understanding artificial intelligence: concepts, limitations, and risks, Equinet. Tillgänglig på: <https://equineteurope.org/publications/understanding-artificial-intelligence-concepts-limitations-and-risks/> (hämtad 1 oktober 2025).

College voor de Rechten van de Mens (2023), “Oordeelnummer 2023-111”. Tillgänglig på: <https://oordelen.mensenrechten.nl/oordeel/2023-111#kop> (hämtad 14 oktober 2025).

College voor de Rechten van de Mens (CRM) (2025), “Toetsingskader risicoprofilering”. Tillgänglig på: www.mensenrechten.nl/actueel/nieuws/2025/01/28/nieuw-toetsingskader-tegen-discriminatie-door-risicoprofilering (hämtad 29 september 2025).

Council of Europe and European Union Agency for Fundamental Rights (2018), Handbook on European non-discrimination law, 2nd edn, Publications Office of the European Union. ISBN: 978-92-871-9851-8. Council of Europe (2025a), “Legal protection against algorithmic discrimination in Europe: current frameworks and remaining gaps”. Tillgänglig på: <https://rm.coe.int/legal-protection-eng-web/48802a38d7> (hämtad 4 juni 2026).

Council of Europe (2025b), “European policy guidelines on AI and algorithm-driven discrimination for equality bodies and other national human rights structures”.

Tillgänglig på: <https://rm.coe.int/policy-guidelines-eng-web/48802a38a5> (hämtad 4 juni 2026).

Council of Europe (2025c), "Methodology for the risk and impact assessment of artificial intelligence systems from the point of view of human rights, democracy and the rule of Law (HUDERIA Methodology)". Tillgänglig på: <https://rm.coe.int/prems-002726-gbr-2006-huderia-texte-web-a4/48802ba7b1> (hämtad 19 juni 2025).

Council of Europe (2026), "Model for the risk and impact assessment of artificial intelligence systems from the point of view of human rights, democracy and the rule of law (HUDERIA Methodology)". Tillgänglig på: <https://rm.coe.int/prems-002726-gbr-2006-huderia-texte-web-a4/48802ba7b1> (hämtad 19 juni 2025).

Crenshaw K. (1991), "Mapping the margins: intersectionality, identity politics, and violence against women of color", *Stanford Law Review*, 43(6), s. 1241-1299. Tillgänglig på: www.jstor.org/stable/1229039 (hämtad 1 oktober 2025).

Crenshaw K. (2015), "Why intersectionality can't wait", *The Washington Post*. Tillgänglig på: www.washingtonpost.com/news/in-theory/wp/2015/09/24/why-intersectionality-cant-wait/ (hämtad 1 oktober 2025).

Crowley N. (2023), Informing the policy agenda: equality bodies making recommendations, *Equinet*. Tillgänglig på: <https://equineteurope.org/informing-the-policy-agenda-equality-bodies-making-recommendations/> (hämtad 1 oktober 2025).

Dastin J. (2018), "Amazon scraps secret AI recruiting tool that showed bias against women", *Reuters*. Tillgänglig på: www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G (hämtad 10 november 2025).

Dimitrova D. och De Hert P. (2024), "The LED's right of access to one's data: loopholes on proper access, legality review and data protection authority accountability", *New Journal of European Criminal Law*, 15(1), s. 12-32. Tillgänglig på: <https://doi.org/10.1177/20322844231214484> (hämtad 17 oktober 2025).

Equinet (2017), Strategic litigation handbook. Tillgänglig på: <https://equineteurope.org/publications/strategic-litigation-handbook/> (hämtad 29 september 2025).

Equinet (2022), Equality Bodies working on cases without an identifiable victim: *Actio popularis*. Tillgänglig på: <https://equineteurope.org/equality-bodies-working-on-cases-without-an-identifiable-victim-actio-popularis/> (hämtad 29 september 2025).

Equinet (2025), Webinar series: equality by design, deliberation and oversight. Tillgänglig på: <https://equineteurope.org/webinar-series-equality-by-design-empowering-equality-bodies-and-civil-society-on-the-interface-between-technical-standards-and-legal-requirements/> (hämtad 1 oktober 2025).

Erdogan I. (2023), "Transparency in the face of new technologies: information rights under the law enforcement directive", *Jean Monnet Network on EU Law Enforcement Working Paper Series*. Tillgänglig på: <https://jmn-eulen.nl/papers/events-on-ai/> (hämtad 17 oktober 2025).

Erdogan I. (2024), "Diving into the iceberg: establishing transparency in AI for law enforcement", *European Papers*, 9(3), s. 956-977. Tillgänglig på: www.europeanpapers.eu/e-journal/diving-into-iceberg-establishing-transparency-ai-for-law-enforcement (hämtad 17 oktober 2025).

Eticas Foundation (2024), Automating (in)justice? An adversarial audit of riscanvi. Tillgänglig på: <https://www.eticasfoundation.org/automating-injustice-an-adversarial-audit-of-riscanvi/> (hämtad 13 oktober 2025).

European Center for Not-for-Profit Law and Danish Institute for Human Rights (2025), A guide to fundamental rights impact assessments (FRIA). Tillgänglig på: <https://ecnl.org/publications/guide-fundamental-rights-impact-assessments-fria> (hämtad 31 april 2026).

European Commission (2019), Equality law review: European network of legal experts in gender equality and non-discrimination, Publications Office of the European Union. Tillgänglig på: <https://op.europa.eu/s/AiAN> (hämtad 19 juni 2026).

European Commission (2021), Guidelines on improving the collection and use of equality data, Publications Office of the European Union. Tillgänglig på: <https://data.europa.eu/doi/10.2838/9725> (hämtad 17 oktober 2025).

European Commission (2025), "Supporting the implementation of the AI ACT with clear guidelines". Tillgänglig på: <https://digital-strategy.ec.europa.eu/en/news/supporting-implementation-ai-act-clear-guidelines> (hämtad 1 maj 2025).

European Court of Human Rights (2025), "Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention". Tillgänglig på: https://ks.echr.coe.int/documents/d/echr-ks/guide_art_14_art_1_protocol_12_eng (hämtad 10 november 2025).

Farkas L. och O'Farrell O. (2015), Reversing the burden of proof: practical dilemmas at the European and national level, Publications Office of the European Union. Tillgänglig på: <https://op.europa.eu/en/publication-detail/-/publication/a763ee82-b93c-4df9-ab8c-626a660c9da8/language-en> (hämtad 10 november 2025).

FRA (2018), Handbook on European non-discrimination law – 2018 edition. Tillgänglig på: <https://fra.europa.eu/en/publication/2018/handbook-european-non-discrimination-law-2018-edition#> (hämtad 1 juli 2026).

FRA (2025), "Assessing high-risk artificial intelligence: fundamental rights risks". Tillgänglig på: Assessing High-risk Artificial Intelligence: Fundamental Rights Risks – European Union Agency for Fundamental Rights (hämtad 4 juni 2026).

Fredman S. (2016), Intersectional discrimination in EU gender equality and non-discrimination law, European Equality Law Network, Publications Office of the European Union. Tillgänglig på: <https://op.europa.eu/en/publication-detail/-/publication/d73a9221-b7c3-40f6-8414-8a48a2157a2f/language-en> (hämtad 17 oktober 2025).

González Fuster G. (2020), "Artificial intelligence and law enforcement. Impact on fundamental rights". European Parliament. Tillgänglig på: <https://op.europa.eu/s/AiyD> (hämtad 12 november 2025).

Howard E. (2024), "Intersectional discrimination and EU law: time to revisit Parris", Sage Journals. Tillgänglig på: <https://journals.sagepub.com/doi/10.1177/13582291241285336> (hämtad 1 maj 2026). Ilieva S. (2024), Handbook on identifying and using equality data in legal casework, Equinet. Tillgänglig på: <https://equineteurope.org/publications/handbook-on-identifying-and-using-equality-data-in-legal-casework/> (hämtad 29 september 2025).

Information Commissioner's Office and Equality and Human Rights Commission (2023), Memorandum of understanding between the Information Commissioner and the Equality and Human Rights Commission. Tillgänglig på: <https://ico.org.uk/about-the-ico/our-information/working-with-other-bodies/> (hämtad 1 oktober 2025).

JuLIA Project (2024), Handbook: artificial intelligence, judicial decision-making and fundamental rights, JuLIA 101046631, European Commission (DG JUST – JUST-2021-JTRA). Tillgänglig på: <https://www.julia-project.eu/resources> (hämtad 1 oktober 2025).

Lighthouse Reports (2025), "How we investigated Amsterdam's attempt to build a 'fair' fraud detection model". Tillgänglig på: www.lighthousereports.com/methodology/amsterdam-fairness/ (hämtad 19 juni 2026).

Lighthouse Reports (2024), "How we investigated Sweden's suspicion machine". Tillgänglig på: www.lighthousereports.com/methodology/sweden-ai-methodology/ (hämtad 20 oktober 2025).

Lighthouse Reports (2023), "How we investigated France's mass profiling machine". Tillgänglig på www.lighthousereports.com/methodology/how-we-investigated-frances-mass-profiling-machine/ (hämtad 19 oktober 2025).

Mahieu R. L., Asghari H. och van Eeten M. (2018), "Collectively exercising the right of access: individual effort, societal effect", Internet Policy Review, 7(3). Tillgänglig på: <https://doi.org/10.14763/2018.3.927> (hämtad 10 november 2025).

Makkonen T. (2007), Measuring discrimination: data collection and EU equality law, Publications Office of the European Union. Tillgänglig på: <https://op.europa.eu/en/publication-detail/-/publication/7d20295d-212c-4acb-bd9f-6f67f4c7ce67> (hämtad 10 november 2025).

Makkonen T. (2016), European handbook on equality data: 2016 revision, Publications Office of the European Union. Tillgänglig på: https://ec.europa.eu/newsroom/just/document.cfm%3Faction%3Ddisplay%26doc_id%3D43205&ved=2ahUKEwiysrmm-eeQAxXiWAIHHUBOH9MQFnoECBgQAQ&usq=AOvVaw1aG86yvjJ51QDEGbA_FR7L (hämtad 10 november 2025).

Mantelero A. (2024), "The fundamental rights impact assessment (FRIA) in the AI Act: roots, legal obligations and key elements for a model template", Computer Law & Security Review, 54, s. 106020. Tillgänglig på: www.sciencedirect.com/science/article/pii/S0267364924000864 (hämtad:10 november 2025).

Mittelstadt B., Wachter, S. and Russell, C. (2023), The Unfairness of Fair Machine Learning: Levelling down and strict egalitarianism by default. Available at: <https://arxiv.org/abs/2302.02404> (hämtad 17 juli 2026).

Mittelstadt B. (2024), How to use the Artificial Intelligence Act to investigate AI bias and discrimination: a guide for equality bodies, Equinet. Tillgänglig på: <https://equineteurope.org/publications/how-to-use-the-artificial-intelligence-act-to-investigate-ai-bias-and-discrimination-a-guide-for-equality-bodies/> (hämtad 1 oktober 2025).

Norwegian Data Protection Authority (2023), "Evaluation of the Norwegian Data Protection Authority's regulatory sandbox for artificial intelligence". Tillgänglig på: <https://www.datatilsynet.no/en/news/aktuelle-nyheter-2023/the-sandbox-has-been-evaluated/> (hämtad 17 oktober 2025).

Parviainen H. (2024), "Challenges of direct discrimination in algorithmic recruitment: insuperable or not?*", International Journal of Comparative Labour Law and Industrial Relations, 40, 4 december 2024, s. 437-466. Tillgänglig på: <https://doi.org/10.54648/ijcl2024017> (hämtad 17 oktober 2025).

Purificato I. (2021), "Behind the scenes of Deliveroo's algorithm: in the blindness of 'frank' its discriminatory potential", Italian Labour Law e-Journal, 14(1), s. 169–194. Tillgänglig på: <https://doi.org/10.6092/issn.1561-8048/12990> (hämtad 22 juni 2026).

PWC (2024), Onderzoek misbruik uitwonendenbeurs. Tillgänglig på: www.tweedekamer.nl/kamerstukken/detail?id=2024D07564&did=2024D07564 (hämtad 12 november 2025). Research Institute – Digital Human Rights Center (2025), "Enlightenment 4.0 – human understanding of AI decision-making: report on the theoretical basis of the right to explanation of individual decision-making under Article 86 AI Act", Federal Ministry of Labour and Social Affairs, Health Care and Consumer Protection Republic of Austria. Tillgänglig på: <https://researchinstitute.at/en/enlightenment-4-0-project-results-published-in-english/>; Website of the Ministry of Social Affairs (hämtad 29 september 2025).

Schot E. och Davidson D. (2025), "Onbetrouwbaar algoritme bestempelt jongeren als toekomstig crimineel, Follow the money". Tillgänglig på: www.ftm.nl/artikelen/algoritme-bestempelt-jongeren-onterecht-als-toekomstig-crimineel?share=Uom-mUbfeHZLpB371b2sHbWwyodKB4uUhl7QNw2Jm6lIWXJmeq7VOeP7i73ixLPA%3D (hämtad 13 oktober 2025).

Skiotytė G. och Sadauskaitė A. (2026), "A Digital Omnibus: identifying interlinks and possible overlaps between different legal acts in the field of digital legislation to streamline tech rules", European Parliament, Policy Department for Economy and Growth, Brussels. Tillgänglig på: [https://www.europarl.europa.eu/thinktank/en/document/ECTI_STU\(2026\)772641](https://www.europarl.europa.eu/thinktank/en/document/ECTI_STU(2026)772641) (hämtad 31 april 2026).

Smith E. och Vogell H. (2022), "How your shadow credit score could decide whether you get an apartment", ProPublica. Tillgänglig på: www.propublica.org/article/how-your-shadow-credit-score-could-decide-whether-you-get-an-apartment (hämtad 17 oktober 2025).

Soler Garrido J., De Nigris S., Bassani E., Sanchez I., Tatjana E., Antoine-Alexandre A. och Boulangé T. (2024), Harmonised standards for the European AI Act, Joint Research Centre, European Commission Tillgänglig på: <https://publications.jrc.ec.europa.eu/repository/handle/JRC139430> (hämtad 17 oktober 2025).

The Guardian (2021), "What happened when a 'wildly irrational' algorithm made crucial healthcare decisions". Tillgänglig på: www.theguardian.com/us-news/2021/jul/02/algorithm-crucial-healthcare-decisions (hämtad 17 oktober 2025).

United Nations (2011), "Guiding principles on business and human rights: implementing the United Nations 'protect, respect and remedy' framework", United Nations, New York and Geneva. Tillgänglig på: <https://digitallibrary.un.org/record/720245?v=pdf> (hämtad 1 oktober 2025).

Vogiatzoglou P. och Marquiene T. (2022), Assessment of the implementation of the Law Enforcement Directive, European Parliament. Tillgänglig på: [https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU\(2022\)740209](https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2022)740209) (hämtad 17 oktober 2025).

Wachter, Mittelstadt och Russell, 2020, Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security Review, Volume 41, juli 2021. Tillgänglig på: <https://www.sciencedirect.com/science/article/abs/pii/S0267364921000406> (hämtad 1 juli 2026).

Weerts H., Xenidis R., Tarissan F., Olsen H. P. och Pechenizkiy M. (2023), "Algorithmic unfairness through the lens of EU non-discrimination law: or why the law is not a decision tree", Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, s. 805-816.

Weerts H., Kelly-Lyth A., Binns R. och Adams-Prassl J. (2024), "Unlawful proxy discrimination: a framework for challenging inherently discriminatory algorithms", The 2024 ACM Conference on Fairness, Accountability, and Transparency, s. 1850-1860. doi:10.1145/3630106.3659010.

Xenidis R. (2021), "Tuning EU equality law to algorithmic discrimination: three pathways to resilience", Maastricht Journal of European and Comparative Law, 27(6), s. 736-758.

Zuiderveen Borgesius F. (2018), "Discrimination, artificial intelligence and algorithmic decision making", Directorate General of Democracy, Council of Europe. Tillgänglig på: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> (hämtad 10 november 2025).

Lagstiftning

Rådets beslut 2007/533/RIF av den 12 juni 2007 om inrättande, drift och användning av andra generationens Schengens informationssystem (SIS II), Europeiska unionens officiella tidning, L 205.

Rådets direktiv 76/207/EEG av den 9 februari 1976 om genomförandet av principen om likabehandling av kvinnor och män i fråga om tillgång till anställning, yrkesutbildning och befordran samt arbetsvillkor, Europeiska gemenskapernas officiella tidning, L 39, 14 februari 1976, s. 40-42.

Rådets direktiv 97/80/EG av den 15 december 1997 om bevisbördan i mål om könsdiskriminering, Europeiska gemenskapernas officiella tidning, L 14, 20 januari 1998, s. 6-8.

Rådets direktiv 2000/43/EG av den 29 juni 2000 om genomförandet av principen om likabehandling av personer oavsett deras ras eller etniska ursprung, Europeiska gemenskapernas officiella tidning, L 180, 19 juli 2000, s. 22–26.

Rådets direktiv 2000/78/EG av den 27 november 2000 om inrättande av en allmän ram för likabehandling i arbetslivet, Europeiska gemenskapernas officiella tidning, L 303, 2 december 2000, s. 16–22.

Rådets direktiv 2004/113/EG av den 13 december 2004 om genomförande av principen om likabehandling av kvinnor och män när det gäller tillgång till och tillhandahållande av varor och tjänster, Europeiska unionens officiella tidning, L 373, 21 december 2004, s. 37–43.

Rådets direktiv (EU) 2024/1499 av den 7 maj 2024 om normer för EB på området likabehandling av personer oavsett deras ras eller etniska ursprung, likabehandling i fråga om anställning och yrke oavsett religion eller övertygelse, funktionsnedsättning, ålder eller sexuell läggning samt likabehandling av kvinnor och män i fråga om social trygghet och tillgång till samt tillhandahållande av varor och tjänster, och om ändring av direktiven 2000/43/EG och 2004/113/EG, Europeiska unionens officiella tidning, L 2024/1499, 29 maj 2024.

Europarådet (1950), Europeiska konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna, Rom, den 4 november 1950, Europeiska fördragsserien nr 5.

Europarådet (2000), Protokoll nr 12 till konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna, Rom, den 4 november 2000, Europeiska fördragsserien nr 177.

Europarådet (2024), Ramkonvention om artificiell intelligens, mänskliga rättigheter, demokrati och rättsstatsprincipen, Vilnius, den 5 september 2024, Europarådets fördragsserie nr 225.

Europaparlamentets och rådets direktiv 2006/54/EG av den 5 juli 2006 om genomförandet av principen om lika möjligheter och likabehandling av kvinnor och män i arbetslivet (omarbetning), Europeiska unionens officiella tidning, L 204, 26 juli 2006, s. 23–36.

Europaparlamentets och rådets direktiv (EU) 2023/970 av den 10 maj 2023 om stärkt tillämpning av principen om lika lön för kvinnor och män för lika eller likvärdigt arbete genom insyn i lönesättningen och verkställighetsmekanismer, Europeiska unionens officiella tidning, L 132, 17 maj 2023, s. 21–44.

Europaparlamentets och rådets direktiv (EU) 2024/1385 av den 14 maj 2024 om bekämpning av våld mot kvinnor och våld i nära relationer, Europeiska unionens officiella tidning, L 2024/1385, 24 maj 2024.

Europaparlamentets och rådets direktiv (EU) 2024/1500 av den 14 maj 2024 om normer för EB, Europeiska unionens officiella tidning, L 2024/1500, 29 maj 2024.

Europeiska kommissionen, Riktlinjer om förbjudna metoder för artificiell intelligens enligt förordning (EU) 2024/1689 (AI-förordningen), C(2025) 5052 slutlig, 29 juli 2025.

Europaparlamentets och rådets förordning (EU) 2026/1744 av den 8 juli 2026 om ändring av förordningarna (EU) 2024/1689, (EU) 2018/1139 och (EU) 2023/1230 vad gäller förenkling av genomförandet av de harmoniserade reglerna om artificiell intelligens (digital omnibusförordning om AI) Tillgänglig på: <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng> (hämtad den 24 juli 2026).

Europaparlamentets och rådets förordning (EU) 2016/679 av den 27 april 2016 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter och om upphävande av direktiv 95/46/EG (allmän dataskyddsförordning), *Europeiska unionens officiella tidning*, L 119, 4 maj 2016, s. 1–88.

Förslag till Europaparlamentets och rådets förordning om ändring av förordningarna (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725 och (EU) 2023/2854 samt direktiven 2002/58/EG, (EU) 2022/2555 och (EU) 2022/2557 vad gäller en förenkling av den digitala lagstiftningsramen samt om upphävande av förordningarna (EU) 2018/1807, (EU) 2019/1150 och (EU) 2022/868 samt direktiv (EU) 2019/1024 (Digital Omnibus), COM/2025/837 slutlig, [EUR-Lex – 52025PC0837 - EN - EUR-Lex](#).

Europaparlamentets och rådets förordning (EU) 2019/1020 av den 20 juni 2019 om marknadskontroll och överensstämmelse för produkter, *Europeiska unionens officiella tidning*, L 169, 25 juni 2019, s. 1–44.

Europaparlamentets och rådets förordning (EU) 2023/988 av den 10 maj 2023 om allmän produktsäkerhet, om ändring av Europaparlamentets och rådets förordning (EU) nr 1025/2012 och direktiv (EU) 2020/1828 samt om upphävande av Europaparlamentets och rådets direktiv 2001/95/EG och rådets direktiv 87/357/EEG, *Europeiska unionens officiella tidning*, L 135, 23 maj 2023, s. 1–51.

Europaparlamentets och rådets förordning (EU) 2024/1689 av den 13 juni 2024 om harmoniserade regler för artificiell intelligens (AI-förordningen), *Europeiska unionens officiella tidning*, L 2024/1689, 12 juli 2024.

Europaparlamentets och rådets förordning (EG) nr 767/2008 av den 9 juli 2008 om informationssystemet för viseringar (VIS) och utbytet mellan medlemsstaterna av uppgifter om viseringar för kortare vistelse (VIS-förordningen), *Europeiska unionens officiella tidning*, L 218, 13 augusti 2008, s. 60–81.

Europaparlamentets och rådets förordning (EU) nr 603/2013 av den 26 juni 2013 om inrättande av "Eurodac" för jämförelse av fingeravtryck för en effektiv tillämpning av förordning (EU) nr 604/2013 om kriterier och mekanismer för att avgöra vilken medlemsstat som är ansvarig för att pröva en ansökan om internationellt skydd som lämnats in i någon medlemsstat av en tredjelandsmedborgare eller en statslös person, om medlemsstaternas brottsbekämpande myndigheters och Europols begäran om jämförelse med Eurodacuppgifter för brottsbekämpande ändamål samt om ändring av förordning (EU) nr 1077/2011 om inrättande av en europeisk byrå för den operativa förvaltningen av stora it-system inom området frihet, säkerhet och rättvisa (omarbetning), *Europeiska unionens officiella tidning*, L 180, 29 juni 2013, s. 1–30.

Svenska diskrimineringslagen (2008:567).

Rättspraxis

Andersen v. Region Syddanmark (2010), Europeiska unionens domstol, mål C-499/08. Tillgänglig på: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=oj:JOC_2010_346_R_0007_01 (hämtad den 1 juli 2026).

Assessment of creditworthiness, authority, direct multiple discrimination, gender, language, age, place of residence, financial reasons, conditional fine (2018), Finlands diskriminerings- och jämställdhetsnämnd, plenum (omröstning), ärende 216/2017. Tillgänglig på: <https://www.yvtltk.fi/en/decisions/multiple-discrimination-in-the-assessment-of-creditworthiness/> (hämtad den 13 oktober 2025).

Association Belge des Consommateurs Test-Achats ASBL and Others v. Conseil des ministres (2011), Europeiska unionens domstol, mål C-236/09, ECLI:EU:C:2011:100. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62009CJ0236> (hämtad den 12 november 2025).

Beghal v. the United Kingdom, ansökan nr 4755/16 (22 augusti 2016), Europadomstolen. För en närmare redogörelse för de krav som ställs på nationell lagstiftning för att en ifrågasatt åtgärd ska anses vara "i enlighet med lag", se: <https://hudoc.echr.coe.int/eng/?i=001-166724>.

CHEZ Razpredelenie Bulgaria AD v. Komisia za zashtita ot diskriminatsia (2015), Europeiska unionens domstol, mål C-83/14, ECLI:EU:C:2015:480. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/ALL/?uri=CELEX:62014CJ0083> (hämtad den 10 november 2025).

CK v. Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien (2025), Europeiska unionens domstol, mål C-203/22, ECLI:EU:C:2025:117. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-203/22> (hämtad den 7 oktober 2025).

David L. Parris v. Trinity College Dublin and Others (2016), Europeiska unionens domstol, mål C-232/09, ECLI:C:2010:674. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62015CJ0443> (hämtad den 1 maj 2026).

Dekker v. Stichting Vormingscentrum voor Jong Volwassenen (VJV-Centrum) Plus (1990), Europeiska unionens domstol, mål C-177/88, ECLI:EU:C:1990:383. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:61988CJ0177> (hämtad den 7 oktober 2025).

D.H. and Others v. The Czech Republic (2007), Europadomstolen, ansökan nr 57325/00. Tillgänglig på: <https://hudoc.echr.coe.int/fre#%7B%22itemid%22%3A%22001-83256%22%7D> (hämtad den 1 oktober 2025).

Dita Danosa v. LKB Lizings SIA (2010), Europeiska unionens domstol, mål C-232/09, ECLI:EU:C:2010:674. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:62009CJ0232> (hämtad den 20 oktober 2025).

Enderby v. Frenchay Health Authority and Secretary of State for Health (1993), Europeiska unionens domstol, mål C-127/92, ECLI:EU:C:1993:313. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-127/92> (hämtad den 1 oktober 2025).

Filcams CGIL Bologna and Others v. (...) S.R.L. (2020), Domstolen i Bologna, beslut nr 2949/2019. Tillgänglig på: <https://apps.eurofound.europa.eu/platformeconomydb/italian-court-rules-that-the-deliveroo-algorithm-discriminates-workers-106121>.

Frédéric Hay v. Crédit Agricole mutuel de Charente Maritime et des Deux-Sèvres (2013), Europeiska unionens domstol, mål C-267/12, ECLI:EU:C:2013:823. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:62012CJ0267> (hämtad den 12 november 2025).

Galina Meister v. Speech Design Carrier Systems GmbH (2012), Europeiska unionens domstol, mål C-415/10, ECLI:EU:C:2012:217. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:62010CJ0415> (hämtad den 12 november 2025).

Glukhin v. Russia (2023), Europadomstolen, ansökan nr 11519/20. Tillgänglig på: <https://hudoc.echr.coe.int/eng?i=001-225655> (hämtad den 7 oktober 2025).

Handels- og Kontorfunktionærernes Forbund i Danmark v. Dansk Arbejdsgiverforening, acting on behalf of Danfoss (1989), Europeiska unionens domstol, mål C-109/88, ECLI:EU:C:1989:383. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:61988CJ0109> (hämtad den 12 november 2025).

Helga Nimz v. Freie und Hansestadt Hamburg (1991), Europeiska unionens domstol, mål C-184/89, ECLI:EU:C:1991:50. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:61989CJ0184> (hämtad den 12 november 2025).

Hoogendijk v. The Netherlands (2005), Europadomstolen, ansökan nr 58641/00. Tillgänglig på: <https://hudoc.echr.coe.int/eng?i=001-68064> (hämtad den 1 oktober 2025).

Inge Nolte v. Landesversicherungsanstalt Hannover (1995), Europeiska unionens domstol, mål C-317/93, ECLI:EU:C:1995:438. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:61993CJ0317> (hämtad den 12 november 2025).

Ingeniørforeningen i Danmark v. Region Syddanmark (2010), Europeiska unionens domstol, mål C-499/08, ECLI:EU:C:2010:600. Tillgänglig på: <https://curia.europa.eu/juris/documents.jsf?critereEcli=ECLI:EU:C:2010:600> (hämtad den 1 oktober 2025).

Ingrid Rinner-Kühn v. FWW Spezial-Gebäudereinigung GmbH & Co. KG (1989), Europeiska unionens domstol, mål C-171/88, ECLI:EU:C:1989:328. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:61988CJ0171> (hämtad den 12 november 2025).

IX v. WABE eV och MH Müller Handels GmbH v. MJ (2021), Europeiska unionens domstol, förenade målen C-804/18 och C-341/19, ECLI:EU:C:2021:144. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/nl/TXT/?uri=CELEX:62019CC0341> (hämtad den 7 oktober 2025).

James v. Eastleigh Borough Council (1990), House of Lords, [1990] 2 AC 751. Tillgänglig på: <https://www.bailii.org/uk/cases/UKHL/1990/6.html> (hämtad den 12 november 2025).

Jyske Finans v. Ligebehandlingsnævnet (2017), Europeiska unionens domstol, mål C-668/15, ECLI:EU:C:2017:278. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-668/15> (hämtad den 10 november 2025).

Kelly v. National University of Ireland (University College Dublin) (2011), Europeiska unionens domstol, mål C-104/10, ECLI:EU:C:2011:506. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/nl/TXT/?uri=CELEX:62010CJ0104> (hämtad den 10 november 2025).

La Quadrature du Net m.fl. mot Premier ministre och Ministère de la Culture (2024), Europeiska unionens domstol, mål C-470/21, ECLI:EU:C:2024:370. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62021CJ0470> (hämtad den 14 november 2025).

(Loredas) HJ mot US och MU (2024), Europeiska unionens domstol, mål C-531/23, ECLI:EU:C:2024:1050. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/nl/TXT/?uri=CELEX:62023CJ0531> (hämtad den 12 november 2025).

M. A. De Weerd, född Roks, m.fl. mot Bestuur van de Bedrijfsvereniging voor de Gezondheid, Geestelijke en Maatschappelijke Belangen m.fl. (1994), Europeiska unionens domstol, mål C-343/92, ECLI:EU:C:1994:71. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:61992CJ0343> (hämtad den 12 november 2025).

Manjang mot Uber Eats UK Ltd m.fl. (2021), Arbetsdomstolen, mål nr 3206212/2021. Tillgänglig på: www.gov.uk/employment-tribunal-decisions/mr-p-e-manjang-v-uber-eats-uk-ltd-and-others-3206212-slash-2021 (hämtad den 7 oktober 2025).

Maria Kowalska mot Freie und Hansestadt Hamburg (1990), Europeiska unionens domstol, mål C-33/89, ECLI:EU:C:1990:265. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:61989CJ0033> (hämtad den 12 november 2025).

Minoo Schuch-Ghannadan mot Medizinische Universität Wien (2019), Europeiska unionens domstol, mål C-274/18, ECLI:EU:C:2019:828. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:62018CJ0274> (hämtad den 10 november 2025).

N.B. mot Slovakien (2010), Europadomstolen, ansökan nr 29518/10, dom meddelad den 12 juni 2012. Tillgänglig på: <https://hudoc.echr.coe.int/eng?i=001-111427> (hämtad den 12 november 2025).

Oršuš m.fl. mot Kroatien (2010), Europadomstolen, ansökan nr 15766/03, dom meddelad den 16 mars 2010. Tillgänglig på: <https://hudoc.echr.coe.int/eng?i=001-97689> (hämtad den 12 november 2025).

PublicResourceOrg, Inc. och Right to Know CLG mot Europeiska kommissionen (2024), Europeiska unionens domstol, mål C-588/21, ECLI:EU:C:2024:201. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62021CJ0588> (hämtad den 17 oktober 2025).

R (on the application of Bridges) mot Chief Constable of South Wales Police (2020), Appellationsdomstolen, mål EWCA Civ 1058. Tillgänglig på: www.bailii.org/ew/cases/EWCA/Civ/2020/1058.html (hämtad den 1 oktober 2025).

R mot Secretary of State for Employment, ex parte Seymour-Smith and Perez (1999), Europeiska unionens domstol, mål C-167/97, ECLI:EU:C:1999:60. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A61997CJ0167> (hämtad den 1 oktober 2025).

Rechtbank Overijssel (2024), *Indirekte discriminatie door DUO bij selectie van uitwonende studenten voor controle. Het verkregen bewijs wordt uitgesloten*, mål ECLI:NL:RBOVE:2024:5627. Tillgänglig på: ECLI:NL:RBOVE:2024:5627, Rechtbank Overijssel, mål AK_23_1340 och AK_23_1341 (hämtad den 1 juli 2026).

(SCHUFA Holding) OQ mot Land Hessen (2023), Europeiska unionens domstol, mål C-634/21, ECLI:EU:C:2023:957. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-634/21> (hämtad den 7 oktober 2025).

(Generaladvokat Pikamäes förslag till avgörande i SCHUFA Holding) Förslag till avgörande av generaladvokat Pikamäe i mål OQ mot Land Hessen (2023), Europeiska unionens domstol, mål C-634/21, ECLI:EU:C:2023:220. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-634/21> (hämtad den 7 oktober 2025).

S. and Marper v. United Kingdom [GC] (2008), Europadomstolen för de mänskliga rättigheterna, ansökningar nr 30562/04 och 30566/04, dom meddelad den 4 december 2008. Tillgänglig på: <https://hudoc.echr.coe.int/fre?i=001-90051> (hämtad den 12 november 2025).

Sylvie Beghal v. United Kingdom (2016), Europadomstolen för de mänskliga rättigheterna, ansökan nr 4755/16, dom meddelad den 14 januari 2016. Tillgänglig på: [https://hudoc.echr.coe.int/eng/#%22itemid%22:\[%22001-166724%22\]}](https://hudoc.echr.coe.int/eng/#%22itemid%22:[%22001-166724%22]}) (hämtad den 12 november 2025).

Tadao Maruko v. Versorgungsanstalt der deutschen Bühnen (2008), Europeiska unionens domstol, mål C-267/06, ECLI:EU:C:2008:179. Tillgänglig på: <https://curia.europa.eu/juris/liste.jsf?num=C-267/06> (hämtad den 12 november 2025).

Verli Hariev Belov v. CHEZ Elektro Bulgaria AD and others. (2013), Europeiska unionens domstol, mål C-394/11, ECLI:EU:C:2013:48. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62011CJ0394> (hämtad den 17 oktober 2025).

VL v. Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie (2021), Europeiska unionens domstol, mål C-16/19, ECLI:EU:C:2021:64, punkt 48. Tillgänglig på: <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:62019CJ0016> (hämtad den 7 oktober 2025)

Bilaga A

Mall för begäran om tillgång till personuppgifter (DSAR)

Observera: En begäran om tillgång till personuppgifter (DSAR) kan utformas antingen i breda termer eller på ett mer avgränsat sätt. Båda tillvägagångssätten har sina egna fördelar och risker.

Specifika begäranden kan öka sannolikheten för att få den exakta information som efterfrågas, vilket minskar behovet av ytterligare korrespondens och påskyndar processen. I praktiken lämnar dock organisationer som har fått en sådan begäran ofta ett individualiserat svar först efter en kompletterande begäran. Snävt avgränsade begäranden i ett tidigt skede kan innebära en belastning för den registrerade, som kanske inte känner till relevanta behandlingsaktiviteter eller den terminologi som används. Detta ökar risken för att information utelämnas och möjliggör restriktiva tolkningar från personuppgiftsansvarigas sida.

Enligt förslaget till Digital Omnibus-förordning⁶¹ kommer detta att formaliseras. Där anges att begäranden om tillgång till personuppgifter ska vara "så specifika som möjligt", samtidigt som "alltför breda och odifferentierade begäranden också bör betraktas som överdrivna", vilket innebär att allmänna och odifferentierade begäranden kan avslås som överdrivna. This template serves as an example only and should be adapted to the specific case.

Denna mall är endast ett exempel och bör anpassas till det specifika ärendet.

Den registrerades namn och adress]

[Valfritt: den registrerades telefonnummer och e-postadress]

[Datum för begäran]

[Personuppgiftsansvarigs uppgifter – företagets adress]⁶²

61. Skäl 35, artikel 3, förslag till Europaparlamentets och rådets förordning om ändring av förordningarna (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 och direktiven 2002/58/EG, (EU) 2022/2555 och (EU) 2022/2557 vad gäller förenkling av det digitala regelverket, och om upphävande av förordningarna (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868 och direktiv (EU) 2019/1024 (Digital Omnibus)

62. Skicka till personuppgiftsansvarig: deras kontaktuppgifter finns vanligtvis i organisationens integritetspolicy eller juridiska meddelande på dess webbplats. I fall som rör ett AI-system kommer den personuppgiftsansvarige typiskt att vara användaren av systemet (den organisation som använder det, det vill säga den enhet som den registrerade interagerade med under den relevanta processen, såsom en bank, försäkringsbolag, universitet eller arbetsgivare) Den ideella organisationen Datenanfragen.de e.V., som driver webbplatsen datarequests.org (tillgänglig även på flera andra språk) underhåller en [foretagsdatabas](https://foretagsdatabas.org) med relevanta kontaktuppgifter för integritetsrelaterade förfrågningar för många företag.

Begäran om tillgång till personuppgifter enligt artikel 15 i GDPR

Till berörd part,

Jag utövar härmed mina rättigheter enligt **artikel 15 i dataskyddsförordningen (GDPR)** och begär bekräftelse på om ni behandlar personuppgifter som rör mig eller inte.

Om ni behandlar mina personuppgifter, ber jag er att, i enlighet med artikel 15 i GDPR, tillhandahålla mig följande information.

1. Tillgång till mina personuppgifter avseende [X] ändamål

- ▶ En kopia av alla personuppgifter som rör mig och som har använts eller beaktats för följande ändamål:

[Infoga här de ändamål som är viktiga för ditt ärende; till exempel att öppna eller administrera ett bankkonto i mitt namn/att ingå, förnya eller förlänga mitt mobiltelefonabonnemang/behandla eller bedöma min rätt till [ange specifik förmån]/behandla eller bedöma min jobbsökansökan/fastställa min tillgång till onlinekonton, appar eller digitala abonnemang/tilldela mig körningar eller leveransbeställningar på plattformen osv.]

2. Information om ytterligare behandlingsändamål

- ▶ Information om andra ändamål för vilka mina personuppgifter har eller behandlas, utöver de ändamål som anges ovan.
- ▶ En beskrivning av kategorierna av berörda personuppgifter.

3. Utlämnanden till tredje part

- ▶ Mottagarna eller kategorierna av mottagare till vilka mina personuppgifter har utlämnats, tillsammans med kopior av de delade uppgifterna och syftet med delningen, samt datum för utlämnandet.

4. Automatiserat beslutsfattande och profilering

- ▶ Information om huruvida automatiserat beslutsfattande, som avses i artikel 22.1 och 4 i GDPR, har genomförts i förhållande till mina personuppgifter.
- ▶ Om så är fallet, vänligen tillhandahåll meningsfull information om den logik som tillämpas, inklusive kategorierna av personuppgifter, de förfaranden och principer som används för att erhålla ett specifikt resultat samt de planerade konsekvenserna av sådan behandling för mig.

5. Profileringsdetaljer

- ▶ Information om huruvida någon profilering har genomförts med användning av mina personuppgifter.
- ▶ Om så är fallet, var god och tillhandahåll mig innehållet i denna profil och förklara hur den skapades.

Om ni överför mina personuppgifter till ett tredjeland eller en internationell organisation, begär jag att bli informerad om de ändamålsenliga skyddsrättsmedlen enligt artikel 46 i GDPR beträffande överföringen.

Vänligen gör de personuppgifter som angår mig, som jag har begärt, tillgängliga för mig i ett strukturerat, allmänt använt och maskinläsbart format som fastställs i artikel 20.1 i GDPR.

Min begäran inkluderar uttryckligen alla andra tjänster och bolag för vilka ni är personuppgiftsansvarig som definieras i artikel 4.7 i GDPR.

Enligt artikel 12.3 i GDPR måste ni tillhandahålla den begärda informationen till mig utan onödigt dröjsmål och i varje fall inom en månad från mottagandet av begäran. Om en förlängning med upp till två ytterligare månader är nödvändig, informera mig i enlighet med detta.

Enligt artikel 15.3 i GDPR måste ni besvara denna begäran utan att det innebär en kostnad för mig.

[Som bevis för min identitet bifogar jag en kopia av mitt pass/körkort m.m.]
eller

[Ange dina identifikationsuppgifter här. Detta inkluderar ofta uppgifter som ditt namn, ditt födelsedatum, din adress, din e-postadress och så vidare.]

Om ni inte besvarar min begäran inom den angivna perioden förbehåller jag mig rätten att vidta rättsliga åtgärder mot er och att lämna in ett klagomål till den ansvariga tillsynsmyndigheten.

Tack på förhand.

Med vänliga hälsningar,

[Ange ditt namn här]

Bilaga B

Mall för begäran om förklaring enligt artikel 86 i AI-förordningen

Följande mall innehåller en version utvecklad av Research Institute – Digital Human Rights Center (2025b), som publicerats på [website of the Austrian Ministry of Social Affairs](#).

[Namn, adress till den berörda personen]

[Frivilligt: telefonnummer och e-postadress till den berörda personen]

[Datum för begäran]

[Uppgifter om användaren av AI-systemet⁶³]

Begäran om förklaring enligt artikel 86 i AI-förordningen

Till den det må beröra,

Jag hänvisar till beslutet som berör mig [beskrivning av beslutet, till exempel avslag på en låneansökan, mottagen riskbedömning, tilldelning av en kreditpoäng], som jag mottog den [datum för mottagande av beslutet].

Detta beslut [] får rättsliga konsekvenser för mig, eftersom det [kortfattat förklara beslutets inverkan, till exempel avslag på en låneansökan, nekande av tillgång till vissa tjänster, ökning av försäkringspremier m.m.]

[eller] påverkar mig betydligt och negativt, eftersom det [ange skäl för försämringen, såsom diskriminering på grund av differentierande kriterier, riktad reklam för sportspel till personer med spelberoende].

Jag gör härmed gällande min rätt till förklaring enligt artikel 86 i förordning (EU) 2024/1689 (AI-förordningen) och begär en tydlig och meningsfull förklaring av AI-systemets roll i beslutsprocessen och de huvudsakliga elementen i det fattade beslutet.

63. Skicka till användaren av AI-systemet: den organisation som använder det – det vill säga den enhet som den registrerade interagerade med under den relevanta processen, såsom en bank, försäkringsbolag, universitet eller arbetsgivare. Om det är en offentlig myndighet – eller en enhet som agerar å deras vägnar – hitta deras kontaktuppgifter i del C i EU-databasen. I alla andra fall, leta efter användarens kontaktuppgifter i AI-systemets gränssnitt, organisationens integritetspolicy eller juridiska meddelande på dess webbplats. Om en användare inte uttryckligen anges, använd kontaktuppgifterna till den personuppgiftsansvarige, som typiskt sett också kommer att vara användaren av systemet.

För detta ändamål begär jag följande information.

1. AI-systemets roll i beslutsprocessen

- ▶ I vilken utsträckning och på vilket sätt bidrog det använda AI-systemet till beslutsprocessen?

2. De huvudsakliga elementen i det fattade beslutet

- ▶ Vilka var de huvudsakliga skälen till beslutet, inklusive huvudsakliga påverkande kriterier, indata och preliminära beslut som påverkade det fattade beslutet?
- ▶ Vilka andra konsekvenser har beslutet för mig, och hur bedömdes dessa konsekvenser?
- ▶ Hur viktades de kriterier som påverkade beslutet och varför?

3. Mänsklig granskning och grund [ej obligatoriskt]

- ▶ Fattades beslutet av AI-systemet på ett helt automatiserat sätt, eller granskades det av en mänsklig beslutsfattare? Om det senare är fallet, hur såg denna granskning exakt ut och vem/vilken avdelning var ansvarig?
- ▶ [Frivilligt om tillämpligt] På vilken avtalsmässig eller rättslig grund baseras beslutet?

[Som bevis för min identitet bifogar jag en kopia av mitt pass/körkort m.m.]

eller

[Ange dina identifikationsuppgifter här. Dessa inkluderar ofta uppgifter som ditt namn, ditt födelsedatum, din adress, din e-postadress och så vidare.]

Jag begär formellt att ni tillhandahåller mig den begärda informationen inom [x månad/månader] [skriftligen], kostnadsfritt och i en begriplig form. Om ni har några frågor, tveka inte att kontakta mig [för pappersinlämningar, använd kontaktuppgifterna som anges ovan].

Tack på förhand för hanteringen av min begäran och jag ser fram emot att få ert snabba svar.

Med vänliga hälsningar,

[Namn på den berörda personen]

Bilaga C

Exempel på användning av algoritmer och AI

Denna bilaga innehåller exempel där algoritmer och AI-system vanligtvis används. Den är ordnad per sektor, så att man lätt kan fastställa sannolik algoritmisk inblandning. För denna förteckning har vi använt oss av Allen och Masters (2020) och AlgorithmWatch (2022), vars närmare detaljer återfinns i referensförteckningen.

Sektor	Exempel
Bankverksamhet	Kreditvärdering
	Automatisering av kundgodkännandeprocessen
	Upptäckt av bedrägerier eller penningtvätt
Barn- och ungdomsvården	Att förutsäga det framtida behovet av socialtjänst för ett barn
Boende	Automatiserad bakgrundskontroll av potentiella hyresgäster
	Onlineformulär
	Kreditprövning
	Lägenhetsannonser på onlineplattformar: vilka annonser visas för vem? (Automatisk placering)
Försäkring	Onlineformulär
	Prissättningsalgoritmer
	Automatisering av kundgodkännandeprocessen
	Automatisk skadehantering
	Upptäckt av bedrägerier
Hälso- och sjukvård	Att förutsäga risken för sjukdom
	Diagnostiksystem
	Att förutsäga antalet äldre som behöver extra vård
Invandring och gränsförvaltning	Ansiktsgenkänning av personer på en bevakningslista
	Automatisering vid mottagandet av visumansökningar
	Automatiserad granskning av resedokument

Sektor	Exempel
Polisen	Att förutsäga förekomsten av brott i vissa stadsdelar eller områden (geografisk prediktiv brottsbekämpning)
	Att förutsäga kaotiskt eller våldsamt beteende
	Verktyg för riskbedömning eller profilering som används för att förutsäga enskilda personers framtida risk att begå brott, inklusive återfall, eller mer generellt för att identifiera "farliga" eller "relevanta" personer (individinriktad prediktiv polisverksamhet)
	Att bedöma eller uppskatta sannolikheten för att en person är gärningsmannen bakom ett begånget brott (individinriktad prediktiv polisverksamhet)
	Kameraövervakning som upptäcker brottslig verksamhet
Rättsprocessen	Medverka vid utarbetandet av lagar
Rättvisa	Databaserad rekommendation om straffpåföljd
Skatter	Att förutse bedrägerier eller misstag
Sociala förmåner	Onlineformulär
	Att förutse bedrägerier eller misstag
	Automatisering av handläggningen av förmåner
Sysselsättning	Urval av jobbansökningar eller urval av CV:n
	Arbetsförmedlingar som förutsäger sannolikheten för att hitta ett arbete
	Tilldelning av jobberbjudanden till arbetssökande
	Att hänvisa arbetssökande till utbildning eller coaching
	Analys av (avgångs-)intervjuer
Utbildning	Antagning
	Automatisering av fördelningen av elever mellan skolor, universitet eller utbildningsprogram
	Automatiserad övervakning
	Att förutsäga eller upptäcka skolk och avhopp från skolan i ett tidigt skede
	Att upptäcka inlärningsssvårigheter

Bilaga D

Exempel på algoritmisk diskriminering eller risk för diskriminering

Bransch eller teknik	Beskrivning av algoritmen	Beskrivning av den diskriminerande praxisen	Källa
Bankväsendet	"Poängsättning" eller bedömning av kreditvärdighet	ADM-systemet hade regler som ledde till att det fattade beslutet att en "40-årig kvinna" inte ansågs kreditvärdig.	Algorithm-Watch (2022) (s. 6) (Ett liknande fall i Finland, finns här) AlgorithmWatch (2022) (s. 6) (Ett liknande fall i Finland, finns här)
Utbildning	Algoritm för att förutsäga om en kurs eller examen kommer att fullföljas	Underrepresenterade minoritetsgrupper tilldelades genomgående en högre risknivå, samtidigt som deras noggrannhetsbetyg var lägre.	Bird, Castleman och Song (2024)
Teknik för ansiktsigenkänning	Teknik som kan identifiera eller jämföra ansiktsbilder	Det finns skillnader i noggrannhet mellan olika demografiska grupper, såsom ålder, kön och hudfärg. Dessa skillnader kan leda till fall av diskriminering. Den nederländska likabehandlingsmyndigheten har fastställt att organisationer måste bevisa att deras teknik inte diskriminerar, vilket innebär att bevisbördan vänds.	College voor de Rechten van de Mens (2023)
Hälsö- och sjukvård	Algoritm för fördelning av "vårdbehov" bland patienter	Ett godtyckligt poängsystem som leder till olika fördelningar av timmar i liknande fall.	The Guardian (2021)

Bransch eller teknik	Beskrivning av algoritmen	Beskrivning av den diskriminerande praxisen	Källa
Boende	Algoritmer för hyresgäst-bedömning, tilldelning av riskpoäng	Uthyrningsprövning baserad på egenskaper som inte visat sig ha något samband med risk.	Smith and Vogell (2022)
Försäkring	Prissättnings-algoritm	Direkt diskriminering på grund av medborgarskap/ nationalitet påverkade priset på bilförsäkringen för annars identiska profiler.	Boscarol (2022)
Hantering av invandring och gräns-kontroll	Granskning av resedokument, teknik för ansikts-igenkänning	Direkt diskriminering, eftersom tekniken oftare gav ett felaktigt resultat för bilder med mörkare hudton än för bilder med ljusare hudton.	Ahmed (2020)
Rätts-väsendet/ barn- och ungdoms-vården	Riskbedömning av återfall	Vissa nationaliteter ledde till en högre riskpoäng, vilket ökade risken för åtal.	Schot and Davidson (2025)
Polis/brotts-bekämpande myndigheter	Riskbedömning av återfall	AI-systemet saknade transparens, förklarbarhet och robusthet. Inkonsekventa och högre riskpoäng för personer utan "faktorer som tyder på social stabilitet", främst yngre individer.	Eticas Foundation (2024)
Sociala förmåner	Upptäcka felaktiga eller bedrägliga ansökningar om barnomsorgs-bidrag	Rasprofilering, där ansökningar från personer som inte är nederländska medborgare genomgående tilldelades högre riskbetyg.	Amnesty International (2021)
Sociala förmåner	Social-försäkrings-systemet förvärrar utestängningen utestängningen Socialförsäkrings-systemet förvärrar utestängningen	Indirekt diskriminering genom ett till synes neutralt kriterium som missgynnar skyddade befolkningsgrupper.	Amnesty International (2023)

Bransch eller teknik	Beskrivning av algoritmen	Beskrivning av den diskriminerande praxisen	Källa
Upptäckt av bidrags-bedrägerier	Ett system för bedömning av bedrägeririsk som används för att identifiera mottagare av sociala förmåner för vidare granskning	Modellen flaggade i oproportionerligt hög utsträckning kvinnor, invandrare, låginkomsttagare och personer utan högskoleutbildning som misstänkta.	Lighthouse Reports (2024)
Upptäckt av bidrags-bedrägerier	System för bedömning av bedrägeririsk som används för att identifiera mottagare av sociala förmåner för vidare granskning (CNAF)	Systemet kritiserades för att det medförde direkta och indirekta diskriminerande konsekvenser för kvinnor, personer med funktionsnedsättning och personer som lever i fattigdom.	La Quadrature du Net (2024) Lighthouse Reports (2023)

Bilaga E

Proxyvariabler

Ur ett datavetenskapligt /dataanalytiskt-perspektiv ersätter proxyvariabler andra variabler, till exempel om dessa inte kan observeras eller mätas direkt. Proxyvariabler möjliggör slutsatser om vissa egenskaper utan att registrera dessa egenskaper direkt. Till exempel indikerar "30 års arbetslivserfarenhet" att personen måste vara åtminstone i mitten av 40-årsåldern (AlgorithmWatch 2022).

I diskrimineringslagstiftning och i algoritm-/AI-system avser "proxies" egenskaper, kännetecken eller kombinationer av sådana som i sig själva inte listas bland de uttryckligen förbjudna diskrimineringsgrunderna enligt det relevanta rättsliga ramverket, men som korrelerar med dessa grunder på ett sätt att deras användning kan leda till otillåten diskriminering.

Två typer av proxies kan identifieras i detta avseende.

- a) Statistiska proxies: egenskaper (eller kombinationer av sådana) som korrelerar med lagligt skyddade grunder och vars användning kan missgynna personer som tillhör en skyddad grupp oproportionerligt. Sådana proxies är typiskt förknippade med indirekt diskriminering. Till exempel finns det i de flesta europeiska länder en korrelation mellan utbildningsgrad och ras eller etniskt ursprung.
- b) Oskiljaktigt kopplade proxies: en egenskap (eller kombinationer av sådana) som är oskiljaktigt kopplade till en skyddad egenskap, så att deras användning återspeglar en inneboende diskriminerande praxis. Dessa proxies är typiskt förknippade med direkt diskriminering. Till exempel har graviditet ansetts vara oskiljaktigt kopplat till kön (se tabell 8 nedan).

Även om det fortfarande är oklart hur nära överensstämmelsen mellan en proxy och en skyddad grupp måste vara för att utgöra direkt diskriminering, kräver en oskiljaktig koppling inte full (100 %) överensstämmelse (Parviainen 2024). Adams-Prassl, Binns och Kelly-Lyth har ansett att en "tillräcklig grad av överensstämmelse" krävs (2023), medan Weerts m.fl. (2024) har föreslagit att proxyn och den skyddade grunden måste vara "åtminstone starkt associerade och möjligen till och med deterministiskt relaterade". Weerts m.fl. noterar relevansen av den kontrafaktiska "but for"-frågan: en klagande i ett land med könsspecifik pensionsålder kunde rimligen argumentera för att hen skulle ha uppnått pensionsåldern om det inte varit för hens kön; ett samkönat par som bor i ett land som förbjuder samkönade äktenskap kunde lätt säga att de skulle vara gifta om det inte varit för deras sexuella läggning; en kvinna kunde rimligen argumentera för att hon inte skulle vara gravid om det inte varit för hennes kön; och en muslimsk kvinna kunde argumentera för att hon inte skulle bära huvudduk om det inte varit för hennes religion.

Huruvida en oskiljaktig koppling finns i ett givet fall kräver en fall-för-fall-bedömning av domstol, där kontextuella kriterier – snarare än exakta statistiska trösklar – ska ges särskild vikt, i linje med EU-domstolens tillvägagångssätt i diskrimineringsmål (se till exempel IX mot Wabe eV och MH Müller Handels GmbH mot MJ, punkt 78, där

vikt lades vid typen av tecken som var involverade). Se tabellen nedan för exempel på olika typer av proxies för skyddade egenskaper.

Tabell 8: Exempel på kända proxies (som nämns i, bland annat, CRM 2025, Weerts m.fl. 2024, AlgorithmWatch 2022)

Egenskap	Proxyvariabel	Typ av proxyvariabel
30 års arbetslivserfarenhet (AlgorithmWatch 2022)	Personen är över 45 år	Oskiljaktig koppling
Graviditet (<i>Dekker v. Stichting voor Jong Volwassenen (VJV-Centrum) Plus</i> , 1990)	Kön	Oskiljaktig koppling
Civilstånd (<i>Tadao Maruko v. Versorgungsanstalt der deutschen Bühnen</i> , 2008; <i>Frédéric Hay v Crédit Agricole mutuel de Charente Maritime et des Deux-Sèvres</i> , 2012)	Sexuell läggning i länder där samkönade äktenskap inte är tillåtna	Oskiljaktigt kopplade proxies i länder där samkönade äktenskap inte är tillåtna
Könsbaserad pensionsålder (<i>James v. Eastleigh Borough Council</i> , 1990, §10)	Kön	Oskiljaktig koppling
Språk eller språkkunskapsnivå	Etnicitet	Statistisk indikator eller oskiljaktig koppling
Läs- och skrivkunnighet	Etnicitet (beroende på land kan detta även vara en indikator på ålder och kön)	Statistisk indikator
Utbildningsnivå	Etnicitet och etniskt ursprung	Statistisk indikator
Postnummer eller stadsdel	Etnicitet (även om bostadsområdena är segregerade)	Statistisk indikator
Inkomst	Etnicitet, kön	Statistisk indikator
Antal personer i hushållet eller den närmaste familjen	Etnicitet, kön	Statistisk indikator
Födelseort	Etnicitet	Statistisk indikator
Mottagande av nödvändiga ekonomiska förmåner	Etnicitet	Statistisk indikator
Religion	Etnicitet	Statistisk indikator

Egenskap	Proxyvariabel	Typ av proxyvariabel
Namn som innehåller, slutar på eller börjar med en viss bokstavskombination	Etnicitet	Statistisk indikator
Frånvaro från arbetet/sena avbokningar/begränsad tillgänglighet för arbete (<i>Filcams Cgil Bologna and Others v. Deliveroo Italia S.R.L., 2020</i>)	Övertygelse (fackförenings-medlemskap), kön eller hälsotillstånd	Statistisk indikator
Klädesplagg och huvudbonader (<i>till exempel att bära huvudduk</i>) (<i>IX v. WABE and MH Müller Handels, 2021</i>)	Religion och övertygelse	Statistisk indikator (som avser exempelvis tecken på politisk, filosofisk eller religiös övertygelse) eller oskiljaktigt kopplad (som uttryckligen avser exempelvis iögonfallande, stora tecken)

Bilaga F

Förteckning över AI- och algoritmregister i Europa

Geografisk tillämpnings-område	Titel	Tillämpnings-område	Länk	Uppdaterings-datum
Belgien	BoSA:s AI-observatorium	Förteckning över alla kända AI-tillämpningar och -projekt som används av federala institutioner och myndigheter	https://bosa.belgium.be/fr/themes/administration-numerique/cooperation-et-partage-des-connaissances/cooperation-nationale-5	
EU:s databas för AI-system med hög risk – under utveckling		AI-system med hög risk		
Varje EU-land – under utveckling		AI-system med hög risk som är kopplade till kritisk infrastruktur		
EU	Utvalda exempel på AI-tillämpningar inom den offentliga sektorn AI-tillämpningar inom den offentliga sektorn	AI inom den offentliga sektorn	Joint Research Centre Data Catalogue – Selected AI cases in the public sector (JRC129301), European Commission	31 december 2021

Geografisk tillämpnings-område	Titel	Tillämpnings-område	Länk	Uppdaterings-datum
EU	Teknik-bevakning inom den offentliga sektorn	Ett observations-centrum med uppgift att övervaka, analysera och sprida kunskap om användningen av ny teknik	https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch	
Tyskland			Use Case Platform – Applied AI Institute	
Helsingfors	Helsingfors stads system för artificiell intelligens	AI inom den offentliga sektorn	Tekoälyrekisteri – Helsingin kaupungin tekoälyrekisteri	Löpande uppdateringar
Nederländerna	Den nederländska regeringens algoritm-register	Algoritmiska system inom den offentliga sektorn	The Algorithm Register of the Dutch Government	Löpande uppdateringar
Storbritannien	Registrerings-central för standarder för algoritmisk transparens	Storbritanniens regering (frivillig standard)	Find out how algorithmic tools are used in public organisations – GOV.UK	Löpande uppdateringar
Storbritannien	Icke-statligt register: Public Law Project Tracking Automated Government (TAG)-registret		TAG Register	Löpande uppdateringar

Bilaga G

Direktivet om likabehandling, skyddade grunder och tillämpningsområde

Direktiv	Skyddade områden	Tillämpningsområde	Direktivet om standarder
2000/43/EG direktivet om likabehandling oavsett ras	Ras och etniskt ursprung – inte medborgarskap och utan att det påverkar den behandling som följer av den rättsliga ställningen för tredjelandsmedborgare och statslösa personer	Sysselsättning, yrkesutbildning, socialt skydd, utbildning, varor och tjänster, inklusive bostäder	2024/1499
EG- direktivet 2000/78/EC	Religion eller övertygelse, funktionsnedsättning, ålder, sexuell läggning funktionsnedsättning, ålder, sexuell läggning	Anställning och yrke	2024/1499
2004/113/EG-direktivet om jämställdhet i fråga om varor och tjänster	Kön	Varor och tjänster	2024/1499
2006/54/EG-direktivet om lika möjligheter och likabehandling av män och kvinnor (Omarbetning)	Kön (inklusive graviditet och moderskap)	Sysselsättning och yrke (inklusive socialförsäkrings-system)	2024/1500

Direktiv	Skyddade områden	Tillämpnings- område	Direktivet om standarder
2010/41/ EU-direktivet om likabehandling av kvinnor och män som är egenföretagare	Kön	Egenföretagare, makar eller sambor till egenföretagare	2024/1500
2023/970/EU- direktivet om lika lön, (ska införlivas senast i juni 2026)		Sysselsättning	2024/1500 (observera dock att detta gäller utan att det påverkar tillämpningen av mer specifika bestämmelser i själva direktivet)

Bilaga H

Andra potentiellt relevanta förordningar och riktlinjer

Europarådet

Likabehandling

Rekommendation CM/Rec(2019)1 från Ministerkommittén till medlemsstaterna om att förhindra och bekämpa sexism: <https://search.coe.int/cm?i=090000168093b26a>.

Rekommendation CM/Rec(2020)1 från Ministerkommittén till medlemsstaterna om algoritmsystems inverkan på mänskliga rättigheter (antagen av Ministerkommittén den 8 april 2020 vid det 1373:e mötet för ministrarnas ställföreträdare): <https://search.coe.int/cm?i=09000016809e1154>

Rekommendation CM/Rec(2026)1 från Ministerkommittén om jämlikhet och artificiell intelligens, GEC/ADI-AI, under GEC och CDADI: <https://search.coe.int/cm?i=09125948802acd7f>.

Rapporter

Rapport om att förhindra diskriminering orsakad av användningen av artificiell intelligens – Europarådets parlamentariska församling, utskottet för jämlikhet och icke-diskriminering; rapportör: Christophe Lacroix, (Belgien, SOC), 29 september 2020: <https://pace.coe.int/en/files/28807>

Shrishak, Kris och Pénicaud, Soizic (2025) Legal protection against algorithmic discrimination in Europe: current frameworks and remaining gaps, Council of Europe. (även på franska, nederländska, portugisiska, finska, svenska). Tillgänglig på: <https://www.coe.int/en/web/inclusion-and-antidiscrimination/upholding-equality-and-non-discrimination-by-equality-bodies-regarding-the-use-of-artificial-intelligence-ai-in-public-administrations1> (hämtad 1 juli 2026)

Study on discrimination, artificial intelligence and algorithmic decision making – ECRI, november 2018: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>.

Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination – GEC/CDADI, augusti 2023: <https://rm.coe.int/prems-112923-gbr-2530-etude-sur-l-impact-de-ai-web-a5-1-2788-3289-7544/1680ac7936>.

Xenidis, Raphaële (2025) European policy guidelines on AI and algorithm-driven discrimination for equality bodies and other national human rights structures (även på franska, nederländska, portugisiska, finska, svenska), Council of Europe. Tillgänglig på: <https://www.coe.int/en/web/inclusion-and-antidiscrimination/upholding-equality-and-non-discrimination-by-equality-bodies-regarding-the-use-of-artificial-intelligence-ai-in-public-administrations1> (hämtad 1 juli 2026)

Kön

Rekommendation CM/Rec(2026)2 från Ministerkommittén om ansvarsskyldighet för teknikunderlättat våld mot kvinnor och flickor: <https://search.coe.int/cm?i=09125948802acd83>.

Rättvisa och tillämpning

Assessment tool for the operationalisation of the European Ethical Charter on the use of artificial intelligence in judicial systems and their environment – CEPEJ(2023)16 Final: <https://rm.coe.int/cepej-2023-16final-operationalisation-ai-ethical-charter-en/1680adcc9c>.

European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment – CEPEJ(2018)14: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

Guidelines on facial recognition, Consultative Committee of the Convention for the Protection of Individuals with regard to automatic processing of personal data, Convention 108, 28 January 2021: <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3>.

Välfärd och social skyddsnät

Decl(17/03/2021)2 – Declaration by the Committee of Ministers on the risks of computer-assisted or artificial-intelligence-enabled decision making in the field of the social safety net (antagen av Ministerkommittén den 17 mars 2021 vid det 1399:e mötet för ministrarnas ställföreträdare) <https://search.coe.int/cm?i=0900001680a1cb98>.

Utbildning

Havinga, Beth, Holmes, Wayne och Persson Jen (2024), Regulating the use of Artificial Intelligence Systems in Education: Preparatory study on the development of a legal instrument. Tillgänglig på: <https://www.coe.int/en/web/education/-/regulating-the-use-of-artificial-intelligence-systems-in-education>

Recommendation on the impact of artificial intelligence (AI) on the education sector, antagen av Conference of International Non-governmental Organisations – CONF-AG(2023)REC4. Tillgänglig på: <https://rm.coe.int/conf-ag-2023-rec4-draft-recommendation-on-the-impact-of-artificial-int/1680ac8e25>

Migration, invandring och asyl

Recommendation CM/Rec(2022)17 från Ministerkommittén till medlemsstaterna om att skydda rättigheterna för migrant-, flykting- och asylsökande kvinnor och flickor (punkterna 5–8 och 22–25 i synnerhet): <https://rm.coe.int/prems-092222-gbr-2573-recommendation-cm-rec-2022-17-a5-bat-web-1-/1680a6ef9a>.

Se även: IT Flows policy brief: <https://bura.brunel.ac.uk/handle/2438/27982>.

Bilaga I

AI-system med hög risk

Högrisk-område	Användningsfall
1. Biometri	System för biometrisk identifiering på distans, exklusive verifiering
	Biometrisk klassificering av känsliga eller skyddade attribut, eller slutsatser om egenskaper utifrån dessa attribut
	Känsloigenkänning
2. Kritisk infrastruktur	Säkerhetskomponenter vid förvaltning och drift av kritisk infrastruktur
3. Utbildning och yrkesutbildning	Beslut om tillträde eller antagning, eller placering av elever på olika utbildningsnivåer
	Utvärdering av lärandemål och/eller styrning av inlärningsprocessen inom utbildning eller yrkesutbildning
	Bedömning av lämplig utbildningsnivå
	Övervakning eller upptäckt av förbjudet beteende hos elever under prov
4. Sysselsättning, personalhantering och möjligheter till egenföretagande	Rekrytering eller urval av fysiska personer, riktade jobbannonser eller bedömning av kandidater/ansökningar
	Beslutsfattande rörande arbetsrelaterade förhållanden (t.ex. befordran), fördelning av arbetsuppgifter samt uppföljning eller utvärdering av prestationer
5. Tillgång till grundläggande tjänster	Bedömning av rätten till offentliga förmåner och tjänster
	Bedömning av kreditvärdighet eller fastställande av kreditbetyg, med undantag för ekonomiskt bedrägeri
	Riskbedömning och prissättning inom liv- och sjukförsäkring
	Att bedöma och/eller klassificera nödsamtal inför uttryckning, eller fastställa prioriteringsordning vid uttryckning av första hjälpen-personal eller vid triagering inom hälso- och sjukvården

Högrisk-område	Användningsfall
6. Brottsbekämpning	Bedömning av risken att bli utsatt för ett brott
	Polygraf eller liknande verktyg
	Bedömning av bevisens tillförlitlighet
	Bedömning av risken för (åter)brottslighet, eller bedömning av personlighetsdrag eller kriminellt beteende
	Bistå vid upptäckt, utredning eller lagföring av brott
7. Hantering av migration, asyl och gränskontroll	Polygraf eller liknande verktyg
	Bedömning av de risker som inresa till en medlemsstats territorium medför när det gäller säkerhet, irreguljär migration eller hälsa
	Prövning av ansökningar om asyl, visum, uppehållstillstånd eller överklaganden av rätten att ansöka om status
	Att upptäcka, känna igen eller identifiera fysiska personer; undantag gäller för kontroll (av resehandlingar)
8. Rättskipning och demokratiska processer	Att bistå rättsliga myndigheter vid utredning och tolkning av sakförhållanden och lagstiftningen samt vid tillämpningen av lagstiftningen på konkreta sakförhållanden
	System som är avsedda att påverka röstbeteendet vid val och/eller folkomröstningar

Observera: kommissionen har publicerat utkast till riktlinjer om klassificering av högrisk-AI-system, som är under konsultation fram till juni 2026.⁶⁴

64. Se mer här: Kommissionen efterfrågar synpunkter på utkastet till riktlinjer för klassificering av högrisk-system för artificiell intelligens; Shaping Europe's digital future, hämtad den 5 juni 2026.

Bilaga J

Förbjudna AI-praxis

Tabell 9. Förteckning över förbjudna användningsområden för AI-system

Förbjuden praxis	Artikel i AI-förordningen	Motiveringen till AI-förordningen
Att medvetet manipulera eller vilseleda, väsentligt förvränga ett beteende eller försämrå förmågan att fatta ett välgrundat beslut, vilket med rimlig sannolikhet kan orsaka betydande skada, inklusive framställning av sexuellt och intimt innehåll utan samtycke eller material som visar sexuella övergrepp mot barn (CSAM).	5.1 a, 5.1 ba & bb, 5.1a, 5.1b	Skäl 29
Att utnyttja en sårbarhet som beror på ålder, funktionsnedsättning eller socioekonomisk ställning, i syfte att eller med följden att väsentligt påverka ett beteende, vilket rimligen kan förväntas orsaka betydande skada.	5.1 b	Skäl 29
Att klassificera eller bedöma fysiska personer under en viss tidsperiod utifrån deras sociala beteende eller personliga egenskaper med hjälp av ett socialt betyg, vilket leder till skadlig eller ofördelaktig behandling i antingen: ▶ ett sammanhang som inte har något samband med det ursprungliga insamlingssyftet; ▶ på ett omotiverat eller oproportionerligt sätt.	5.1 c	Skäl 31
Riskbedömning eller förutsägelse av att ett brott kommer att begås, uteslutande baserad på profilering av en fysisk person.	5.1 d	Skäl 42
Icke-riktad insamling av ansiktsbilder från övervakningskameror eller internet, i syfte att utöka eller skapa en databas för ansiktigenkänning.	5.1 e	Skäl 43
Känsligenkänning på arbetsplatsen; undantag gäller av medicinska och säkerhetsmässiga skäl.	5.1 f	Skäl 44
Biometrisk klassificering av enskilda personer i syfte att fastställa deras etniska tillhörighet, politiska åsikter, fackföreningstillhörighet, religiösa eller filosofiska övertygelser, sexliv eller sexuell läggning.	5.1 g	Skäl 30

Förbjuden praxis	Artikel i AI-förordningen	Motiveringen till AI-förordningen
Biometrisk identifiering i "realtid" på platser som är tillgängliga för allmänheten i brottsbekämpningssyfte; undantag gäller.	5.1 h	Skäl 32, 33, 38, 39, 40 och 41

Tabell 10. Prohibited AI practices can be related to high-risk AI

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 iAI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
1) Biometri a) system för biometrisk identifiering på distans (RBI)	(h) Biometrisk identifiering i realtid på avstånd i offentliga utrymmen för brottsbekämpande ändamål, såvida den inte används för: i. sökandet efter en person som blivit bortförd; ii. att avvärja ett konkret, allvarligt och överhängande hot; iii. att identifiera en person i samband med vissa brott (som anges i bilaga II).	Viktiga aspekter: användningen av RBI i: i) "realtid" ii) offentliga utrymmen iii) av, eller på uppdrag av, brottsbekämpande myndigheter I skäl 32 anges att användningen av RBI-system i realtid "väcker en känsla av ständig övervakning och avskräcker indirekt från att utöva mötesfriheten och andra grundläggande rättigheter" Exempel: polisen använder övervakningskameror med ansiktsigenkänning i realtid för att identifiera och gripa en person av intresse. Efter att ha identifierat en fredlig demonstrant använde polisen i Moskva (Glukhin mot Ryssland, 2024) övervakningsbilder för att identifiera och lokalisera en person med hjälp av biometrisk identifiering i "realtid".

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
1a) System för biometrisk identifiering på distans (RBI)	(e) Icke-riktad insamling av ansiktsbilder i syfte att skapa eller utöka en databas	Viktiga aspekter: metoden att skapa eller utöka en igenkänningsdatabas genom icke-riktad datainsamling i stället för att identifiera sig själv. Detta är förbjudet eftersom icke-riktad datainsamling fungerar som en "dammsugare", enligt artikel 5.1 e, och samlar in ansiktsbilder urskillningslöst utan att rikta in sig på specifika individer. Skäl 43 kompletterar motiveringen med hänvisning till "känslan av massövervakning" och risken för kränkning av grundläggande rättigheter. Exempel: gränsmyndigheter som bygger upp en databas för ansiktsgenkänning genom massinsamling Det kan förekomma situationer där gränsmyndigheter samlar in ansiktsbilder från sociala medier, nyhetsartiklar och övervakningskameror utan att rikta in sig på specifika individer. Om dessa sedan läggs till i en databas för biometrisk identifiering skulle detta vara förbjudet enligt artikel 5.1 e.
1b) Biometrisk klassificering utifrån skyddade egenskaper (eller slutsatser om sådana)	(g) Biometrisk klassificering som ger upphov till slutsatser om ras, politiska åsikter, fackförenings-tillhörighet, religiösa eller filosofiska övertygelser, sexliv eller sexuell läggning	Viktiga aspekter: det specifika syftet att utläsa de angivna skyddade egenskaperna (ras, politiska åsikter, fackföreningstillhörighet, religion, sexliv eller sexuell läggning) utifrån biometriska uppgifter.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
		Biometrisk kategorisering betraktas redan omedelbart som en hög risk, och artikel 5.1 g förbjuder helt och hållet användningen av biometriska uppgifter för att dra slutsatser om dessa specifika egenskaper. Exempel: ett AI-system som analyserar en persons bilder på sociala medier för att utifrån hens biometriska uppgifter dra slutsatser om hens sexuella läggning i syfte att skicka riktad reklam.
1c) Känsligenkänning	(f) Känsligenkänning på arbetsplatser och utbildningsinstitutioner (undantag: medicinska och säkerhetsmässiga ändamål)	Viktiga aspekter: att dra slutsatser om känslor eller avsikter i arbets- och utbildningssammanhang där undantag av medicinska eller säkerhetsmässiga skäl inte gäller. Artikel 5.1 f förbjuder användning av högriskmetoder för känsligenkänning i dessa specifika sammanhang. I skäl 44 framhålls vidare farhågorna kring den vetenskapliga grunden för sådana AI-system som, särskilt i kombination med maktobalanser i arbets- eller utbildningssammanhang, kan få skadliga konsekvenser. Observera: att identifiera fysiska tillstånd, till exempel om en bussförare är trött eller inte, är inte förbjudet här. Exempel: ett AI-system som avläser om anställda lämnar sin arbetsplats "glada", "arga" eller "ledsna". En arbetsgivare som försöker bedöma känslorna hos sina restauranganställda utifrån bilder från övervakningskameror.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
3) Utbildning och yrkesutbildning a) fastställande av tillträde/antagning	(a) Manipulation, vilseledning eller påverkan som hindrar ett välgrundat beslut och som orsakar skada	Viktiga aspekter: användning av en manipulativ eller vilseledande teknik som väsentligt påverkar (eller med rimlig sannolikhet kommer att påverka) en persons beteende. Artikel 5.1 a förbjuder tekniker som har till syfte eller till följd att påverka beteendet genom metoder som hindrar ett välgrundat beslutsfattande och orsakar betydande skada. Observera: uppsåt krävs inte, konsekvensen är tillräcklig. Exempel: ett universitet som använder en chattbot vid antagningsbesluten, med nedräkningar eller påhittade meddelanden. En chattbot som använder tekniker som även kallas "dark patterns", till exempel "Endast 3 platser kvar!", skulle vara förbjuden enligt denna artikel.
3) Utbildning och yrkesutbildning b) utvärdering av läranderesultat och styrning av inlärningsprocessen c) bedömning av (lämplig) utbildningsnivå d) övervakning/upptäckt av förbjudet beteende hos elever	(f) Känslögenkänning på arbetsplatser och utbildningsinstitutioner (undantag: medicinska och säkerhetsmässiga ändamål)	Viktiga aspekter: att dra slutsatser om studenternas känslomässiga tillstånd eller avsikter, och hur dessa påverkar deras inlärningsprocess. AI-system får användas för att bedöma studenter; det är dock förbjudet att identifiera deras känslor eller avsikter. I skäl 44 framhålls vidare farhågor om den vetenskapliga grunden för sådana AI-system, vilka – särskilt i kombination med maktobalanser i ett arbets- eller utbildningssammanhang – kan få skadliga konsekvenser.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
		Exempel: Ett universitet använder en personaliserad inlärningsplattform som tar hänsyn till studenternas känslor, som registreras via deras webbkameror under online föreläsningar, för att styra inlärningsprocessen.
4) Anställning och personalhantering a) rekrytering (riktad annonsering) eller urval (bedömning av ansökningar)	a) Manipulation, vilseledning eller påverkan av ett välgrundat beslut, vilket orsakar skada (b) Utnyttjande av en sårbarhet	Viktiga aspekter: användning av en manipulativ eller vilseledande teknik som väsentligt snedvrider (eller med rimlig sannolikhet kommer att snedvrida) en individs beteende. Artikel 5.1 a förbjuder tekniker som har till syfte eller till följd att snedvrida beteendet genom metoder som hindrar ett välgrundat beslutsfattande och orsakar betydande skada. Observera: uppsåt krävs inte, konsekvensen är tillräcklig. Exempel: en AI-baserad rekryteringsplattform som skapar en påhittad brådska och använder vilseledande manipulation för att få någon att underteckna ett avtal. En rekryteringsbyrå som använder ett AI-system med överdrivna löne- och förmånsuppgifter eller skapar en brådska som hindrar ett välgrundat beslut skulle vara förbjuden enligt denna artikel.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
4) Anställning och personalhantering b) beslut som påverkar arbetsförhållandena, såsom uppsägning eller arbetsfördelning	(a) Manipulation, vilseledning eller påverkan av ett välgrundat beslut, vilket orsakar skada (b) Utnyttjande av en sårbarhet	Viktiga aspekter: användning av en manipulativ eller vilseledande teknik som väsentligt påverkar (eller med rimlig sannolikhet kommer att påverka) en persons beteende. Artikel 5.1 a förbjuder tekniker som har till syfte eller till följd att påverka beteendet genom metoder som hindrar ett välgrundat beslutsfattande och orsakar betydande skada. Observera: uppsåt krävs inte, konsekvensen är tillräcklig. Exempel: ett logistikföretag som använder AI för att fördela leveransuppdrag och övervaka förarnas prestationer. Det finns ett tänkbart scenario där förarnas beteende genom ett sådant system skulle kunna påverkas så att de arbetar längre än de lagstadgade arbetstiderna, vilket påverkar hälsa och säkerhet och orsakar betydande skada. Ett sådant system skulle därför vara förbjudet.
4) Anställning och personalhantering b) beslut som påverkar arbetsförhållanden, såsom uppsägning eller arbetsfördelning	(f) Identifiering av känslor och/eller avsikter på arbetsplatser och utbildningsinstitutioner	Viktig aspekt: att dra slutsatser om anställdas känslomässiga tillstånd eller avsikter, samt hur dessa påverkar deras arbete. AI-system får användas för att underlätta beslutsfattandet. Det är dock förbjudet att basera sådant stöd på användning av känsloligenkänning. I skäl 44 framhålls vidare farhågor om den vetenskapliga grunden för sådana AI-system, vilka – särskilt i kombination med maktobalanser i arbets- eller utbildningssammanhang – kan få skadliga konsekvenser.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
		Exempel: ett callcenter som använder AI för att övervaka anställda utnyttjar deras känslomässiga tillstånd för prestationsbedömning. I exemplet skulle det vara en förbjuden praxis om ett AI-system skulle dra slutsatser om den anställdes känslomässiga tillstånd utifrån röstinspelningar och beakta detta vid prestationsbedömningen.
5) Tillgång till nödvändiga tjänster och förmåner a) bedömning av en fysisk persons berättigande	(c) Social bedömning som leder till ofördelaktig behandling: i) i ett annat sammanhang än det för vilken uppgifterna samlats in och/eller ii) omotiverad eller oproportionerlig	Viktiga aspekter: ofördelaktig behandling som grundar sig på deras sociala beteende i ett sammanhang som inte har något samband med ärendet och/eller som har en oproportionerlig konsekvens. Om ett AI-system används för att bedöma berättigande är det inte säkert att det hämtar data från ett annat sammanhang. Exempel: myndigheter använder ett AI-system för att tilldela förmåner som bygger på data som inte har något samband med ärendet. I dessa förbjudna fall är det viktigt att sammanhanget där uppgifterna samlas in väsentligt skiljer sig från det sammanhang där de används.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
5b) Bedömning av kreditvärdighet	(c) Social bedömning som leder till ofördelaktig behandling: i) i ett annat sammanhang än de insamlade uppgifterna och/eller ii) omotiverad eller oproportionerlig	Viktiga aspekter: ofördelaktig behandling som grundar sig på personens sociala beteende i ett sammanhang som saknar relevans och/eller som har en oproportionerlig inverkan. Om ett AI-system används för att bedöma en persons kreditvärdighet får detta baseras på dennes ekonomiska historik, men inte på andra, icke-relaterade sammanhang. Exempel: finansinstitut använder ett AI-system för att bedöma låneansökningar, men har även tillgång till andra, icke-finansiella databaser. Om en bank dessutom skulle analysera kontakter på sociala medier eller beakta personens adress vid beslutet om huruvida sökandens lån ska beviljas eller inte, skulle detta kunna vara förbjudet enligt artikel 5.1 c.
5c) Riskbedömning i samband med prissättningen av sjuk- eller livförsäkringar	(c) Social bedömning som leder till ofördelaktig behandling: i) i ett annat sammanhang än det för vilken uppgifterna samlats in och/eller ii) omotiverad eller oproportionerlig	Viktiga aspekter: ofördelaktig behandling som grundar sig på en persons sociala beteende i ett icke-relaterat sammanhang och/eller som har en oproportionerlig konsekvens. Om ett AI-system används för prissättning av livförsäkringar får det baseras på sjukdomshistoria, men inte på andra, icke-relaterade sammanhang. Exempel: ett försäkringsbolag använder ett AI-system för att fastställa priser utifrån information från ett icke-relaterat sammanhang.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
		Om ett försäkringsbolag skulle analysera kontakter och beteende på sociala medier skulle detta kunna vara förbjudet enligt artikel 5.1 c.
6) Brottsbekämpning a) bedömning av risken att bli brottsoffer b) lögn-detektortester c) bedömning av bevisens tillförlitlighet d) bedömning av risken för (åter) brottslighet eller bedömning av personliga egenskaper e) profilering under upptäckt, utredning eller åtal	(d) Bedömning av brottsrisk som uteslutande grundar sig på profilering	Viktiga inslag: bedömningen baseras på personlighetsdrag och egenskaper och inte på objektiva fakta. Observera: om ett system stödjer ett mänskligt beslut kan ett undantag gälla. Artikel 5.1 d förbjuder riskbedömningar av brott som uteslutande baseras på profilering av personlighetsdrag utan objektiva, verifierbara fakta som kopplar dessa till brottslig verksamhet. Exempel: system för prediktiv polisverksamhet, där individer profileras utifrån demografiska uppgifter. Om polisen skulle förutsäga sannolikheten för brottslig verksamhet utifrån egenskaper såsom ålder, nationalitet eller bostadsområde, skulle detta utgöra ett förbjudet system enligt artikel 5.1 d.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 iAI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
6) Brottsbekämpning e) profilering i samband med upptäckt, utredning eller lagföring	(h) Biometrisk identifiering i realtid på distans i offentliga utrymmen för brottsbekämpande ändamål, såvida den inte används för: i) sökandet efter en person som blivit bortförd; ii) att avvärja ett konkret, allvarligt och överhängande hot; iii) att identifiera en person i samband med specifika brott (som anges i bilaga II).	Viktiga aspekter: användningen av RBI i: i) "realtid" ii) offentliga utrymmen iii) av, eller på uppdrag av, brottsbekämpande myndigheter I skäl 32 anges att användningen av RBI-system i realtid "väcker en känsla av ständig övervakning och avskräcker indirekt från att utöva mötesfriheten och andra grundläggande rättigheter". Exempel: polisen använder övervakningskameror med ansiktsigenkänning i realtid för att identifiera och gripa en person av intresse. Efter att ha identifierat en fredlig demonstrant använde polisen i Moskva (Glukhin mot Ryssland, 2024) övervakningsbilder för att identifiera och lokalisera en person med hjälp av biometrisk identifiering i "realtid".
7) Migration, asyl och gränskontroll b) riskbedömning av säkerhet, oegentligheter eller hälsa c) bedömning av asylansökningar, visumansökningar och tillståndsansökningar, behörighet och klagomål	(d) Bedömning av brottsrisk som uteslutande grundar sig på profilering	Viktiga aspekter: bedömningen baseras på personlighetsdrag och inte på objektiva fakta. Observera: om ett system stödjer ett mänskligt beslut kan undantag gälla. Artikel 5.1 d förbjuder riskbedömningar av brott som uteslutande baseras på profilering av personlighetsdrag utan objektiva, verifierbara fakta som kopplar dessa till brottslig verksamhet. Exempel: gränsmyndigheter som använder ett AI-system som baserar riskprofileringen av individer på deras demografiska uppgifter.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 iAI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
		Om gränsmyndigheter skulle förutsäga risken för brottslighet utifrån egenskaper såsom nationalitet eller region, skulle det kunna vara ett förbjudet system enligt artikel 5.1 d.
7) Migration, asyl och gränskontroll (b) riskbedömning av säkerhet, oegentligheter eller hälsa	(h) Biometrisk identifiering i realtid på distans i offentliga utrymmen för brottsbekämpande ändamål, såvida den inte används för: i) sökande efter en person som blivit bortförd; ii) att avvärja ett konkret, allvarligt och överhängande hot; iii) att identifiera en person i samband med specifika brott (som anges i bilaga II).	Viktiga aspekter: användningen av RBI i: i) "realtid" ii) offentliga utrymmen iii) av, eller på uppdrag av, brottsbekämpande myndigheter I skäl 32 anges att användningen av RBI-system i realtid "väcker en känsla av ständig övervakning och avskräcker indirekt från att utöva mötesfriheten och andra grundläggande rättigheter". Exempel: användning av CCTV-kameror med ansiktsigenkänning i realtid vid gränsen för att identifiera personer av intresse.

AI-system med hög risk (artikel 6.2, bilaga III till AI-förordningen)	Förbjudna metoder (artikel 5 i AI-förordningen)	Exempel på förbjudna och potentiellt förbjudna användningsfall samt förklaringar
7) Migration, asyl och gränskontroll (d) att upptäcka, känna igen eller identifiera fysiska personer (undantag för kontroll av resehandlingar)	(e) Icke-riktad insamling av ansiktsbilder	Viktiga inslag: metoden att skapa eller utöka en igenkänningsdatabas genom icke-riktad datainsamling i stället för att identifiera sig själv. Detta är förbjudet eftersom icke-riktad datainsamling fungerar som en "dammsugare", enligt artikel 5.1 e, och samlar in ansiktsbilder urskillningslöst utan att rikta in sig på specifika individer. Skäl 43 kompletterar motiveringen med hänvisning till "känslan av massövervakning" och risken för kränkning av grundläggande rättigheter. Exempel: gränsmyndigheter som bygger upp en databas för ansiktsgenkänning genom massinsamling. Det kan förekomma fall där gränsmyndigheter samlar in ansiktsbilder från sociala medier, nyhetsartiklar och övervakningskameror utan att rikta in sig på specifika individer. Om dessa sedan skulle läggas till i en databas för biometrisk identifiering vid gränsen skulle detta vara förbjudet enligt artikel 5.1 e.

Bilaga K

Dokumentationsblad för steg 3

Detta dokumentationsblad har utformats som ett exempel på hur en handläggare som följer stegen i avsnitt 3 kan dokumentera sina iakttagelser. Exempelen är fiktiva och syftar till att ge en bild av vilken typ av information som kan dokumenteras. Den faktiska dokumentationen förväntas vara betydligt mer detaljerad. Den sista kolumnen är avsedd att möjliggöra en förnyad bedömning av bevisningen efter det att svaranden har lämnat ut ytterligare information.

Steg		Antagande om diskriminering	Efter den svarandes upplysningar
3.1. Kända och misstänkta risker 3.1.1 Systemtyp	Ange om systemet är ett högrisk-system eller om det är avsett att användas inom någon av de sektorer eller områden som enligt AI-förordningen identifierats som medför risk för diskriminering.	Exempel Bransch: Användningsfall:	
3.1.2 Kända risker för diskriminering	Dokumentera det område där det föreligger en känd risk, vilken typ av teknik det rör sig om samt all annan relevant information och bevisning.	Exempel: Sektor: brott-sbekämpning Teknik: ansikts-igenkänning i realtid, biometri Källa: icke-statlig organisation eller akademisk forskning, EU:s AI-förordning, artikel 5 (förbjudna system), artikel 6 (hög risk) och bilaga III.	

Steg		Antagande om diskriminering	Efter den svarandes upplysningar
3.2 Protected characteristics and proxies	Dokumentera direkt användning av skyddade egenskaper. Var används de och på vilket sätt?	Skyddade egenskaper Användningsområde: indata, till exempel	
Undersök närmare systemets tekniska utformning och hur de (misstänkta) indikatorerna påverkar systemets resultat. Det faktum att en indikator används som indata räcker i sig inte som bevis för diskriminering. Nästa steg är att utvärdera bevisningen för att förstå om användningen av dessa indata leder till ojämlik behandling.	Dokumentera alla identifierade fall där skyddade egenskaper eller indikatorer används för att, eller tros, påverka systemets resultat. Dokumentera även för varje indikator varför EB anser att dessa utgör indikatorer.	Exempel Proxyvariabler: till exempel ett postnummer som proxyvariabel för etnicitet, graviditet som proxyvariabel för kön. Bevis och argument	
3.3 Riskbedömningar, inklusive FRIA	Dokumentera eventuella risker för diskriminering som identifierats av leverantören eller den som genomför åtgärden samt eventuella åtgärder som fastställts för att minska dessa risker.	Exempel Risk: rasdiskriminering till följd av historiska fördomar Var risken har identifierats: riskbedömning Åtgärd för att minska risken: teknik för att motverka fördomar	

Steg		Antagande om diskriminering	Efter den svarandes upplysningar
	Överväg och dokumentera om det finns några tydliga och uppenbara brister i riskbedömningen, DPIA eller FRIA.	Riskbedömningen avser diskriminering men inte ras. DPIA: tar inte hänsyn till ras inom ramen för särskilda kategorier av personuppgifter. FRIA: tar inte hänsyn till rasdiskriminering och artikel 14 i Europeiska konventionen om de mänskliga rättigheterna.	
	Beakta och dokumentera förekomsten och inverkan av eventuella mänskliga aktörer i processen. Vem är den berörda personen? Har den berörda personen den kompetens, utbildning och befogenhet som krävs för att åsidosätta ett beslut som fattats av ett AI-system? Har den berörda personen fått utbildning i att förstå och tolka systemets resultat samt att upptäcka var partiskhet kan uppstå och var ingripande kan krävas?	Roll: granskar beslut men granskar inte de uppgifter som låg till grund för beslutet. Fråga: Är detta tillräckligt?	

Steg		Antagande om diskriminering	Efter den svarandes upplysningar
	Finns det några riktlinjer eller protokoll som utfärdats till beslutsfattare för att hjälpa dem att förstå hur AI-systemet ska användas eller hur resultaten ska tolkas?		
3.4 Systemets syfte	Beskriv syftet med systemet och vad det är avsett att optimera.	Systemets syfte: Utformat för att optimera:	
	Registrera indikatorer och skyddade egenskaper som påverkar systemets resultat samt, i den mån sådana finns tillgängliga, mått på deras vikt/inflytande på resultatet.	Skyddade egenskaper som påverkar systemet: Proxyer som påverkar systemet: Mått på vikt/inflytande:	
3.5 Data från utbildning, testning och validering	Dokumentera alla allmänna problem som rör data, inklusive problem som gäller datakvaliteten.	Till exempel kan datamaterialet vara mycket begränsat, föråldrat eller avse ett annat geografiskt område.	
	Datafrågor som handlar om ojämlikhet och diskriminering.		
	Problem som handlar om bristfällig dokumentation.		
	Farhågor kring åtgärder för att begränsa konsekvenserna .		

Steg		Antagande om diskriminering	Efter den svarandes upplysningar
	Argument för hur frågor rörande bevisning och lindrande omständigheter skulle leda till ojämlik behandling i det aktuella fallet.		
	Dokumentera alla tecken på bristande efterlevnad av de bestämmelser i AI-förordningen som rör data.		
3.6 Testning och utvärdering	Brister i vårdutbudet eller behandlingen som framgår av statistiken.		
	Problem med den rapporterade statistiken (t.ex. fel mått eller problem med uppgifterna).		
	Brister i testningen har konstaterats.		
Slutsats	Granska uppgifterna och redogör för på vilka grunder man kan hävda att det föreligger ett fall av prima facie-diskriminering, eller dra slutsatsen att så inte är fallet.		
	Gå igenom alla steg efter ytterligare redogörelse för att avgöra om svaranden har lyckats motbevisa presumptionen eller motiverat diskrimineringen.		

Artificiell intelligens (AI) förändrar våra samhällen och påverkar de grundläggande rättigheterna och jämlikheten. Dessa system kan förstärka och skapa diskriminering eftersom de ofta är utformade för att skilja mellan individer: genom att klassificera, rangordna eller ha dem som måltavla. I kombination med historiska ojämlikheter och mänsklig partiskhet som är inbäddade i data, regler och designval i algoritmiska system kan diskriminerande mönster reproducera och förstärkas om de inte proaktivt identifieras och motverkas.

Denna metodologier steg-för-steg-vägledning till jämställdhetsorgan och andra människorättsorganisationer i Europa för att bedöma diskrimineringsärenden som berör AI eller algoritmiska system. Den är avsedd för handläggare, rådgivare och jurister vid organisationer som utreder anmälningar om diskriminering där AI eller automatiserat beslutsfattande är inblandat eller misstänks vara inblandat.

Metodologin bygger på det europeiska regelverket för AI, särskilt EU:s AI-lag och Europarådets ramkonvention om artificiell intelligens och mänskliga rättigheter, demokrati och rättsstatsprincipen, samt antidiskrimineringslagstiftning, och hänvisar vid behov till annan lagstiftning, såsom data-skyddslagstiftning, exempelfall och rättspraxis.



YHDENVERTAISUUSVALTUUTETTU
NON-DISCRIMINATION OMBUDSMAN



COMISSÃO PARA A CIDADANIA
E A IGUALDADE DE GÉNERO

Medlemsstaterna i Europeiska unionen har beslutat att förena sina kunskaper, resurser och öden. Tillsammans har de byggt upp en zon av stabilitet, demokrati och hållbar utveckling samtidigt som de har bevarat kulturell mångfald, tolerans och individers friheter. Europeiska unionen är fast besluten att dela med sig av sina framsteg och sina värderingar till länder och folk utanför sina gränser.

<http://europa.eu>

Europarådet är kontinentens ledande organisation för de mänskliga rättigheterna. Det består av 46 medlemsstater, inklusive samtliga medlemmar av Europeiska unionen. Alla medlemsstater i Europarådet har undertecknat den europeiska konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna (Europakonventionen), ett fördrag som syftar till att skydda de mänskliga rättigheterna, demokrati och rättsstatsprincipen. Europeiska domstolen för de mänskliga rättigheterna (Europadomstolen) övervakar genomförandet av Europakonventionen i medlemsstaterna.

www.coe.int

Co-funded
by the European Union



EUROPEAN UNION

COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

Co-funded and implemented
by the Council of Europe