

Vibeke Blaker Strand

# Algoritmer, kunstig intelligens og diskriminering

En analyse av likestillings- og  
diskrimineringslovens  
muligheter og begrensninger

Utredning skrevet på oppdrag fra  
Likestillings- og diskrimineringsombudet

# Algoritmer, kunstig intelligens og diskriminering

En analyse av likestillings- og diskrimineringslovens muligheter og begrensninger

**Vibeke Blaker Strand**

professor ved Institutt for offentlig rett, Universitet i Oslo

Utredning skrevet på oppdrag fra Likestillings- og diskrimineringsombudet

Levert 21. mars 2024

Likestillings- og diskrimineringsombudet  
LDO 2024

Forsidefoto: Incept-G  
Utforming: Munch design

ISBN  
978-82-8320-032-4 (trykt utgave)  
978-82-8320-033-1 (elektronisk utgave)

# Innhold

## Sammendrag 7

### Kapittel 1

#### Innledning 9

### Kapittel 2

#### Utredningens nærmere innretning 10

- 2.1 Siktemål 10
- 2.2 Begrepsbruk 10
- 2.3 Avgrensninger 11
- 2.4 Rettskilder 12
- 2.5 Struktur 12

### Kapittel 3

#### Linjer fra forbudet mot diskriminering til regulering av kunstig intelligens og personvern 13

- 3.1 Innledning 13
- 3.2 KI-forordningen 13
- 3.3 Personvernforordningen (GDPR) 14
- 3.4 Sammenfatninger og refleksjoner om veien videre 15

### Kapittel 4

#### Forbudet mot direkte og indirekte diskriminering 17

- 4.1 Innledning 17
- 4.2 Noen utgangspunkter knyttet til forbudet mot direkte diskriminering 17
- 4.3 Noen utgangspunkter knyttet til forbudet mot indirekte diskriminering 18
- 4.4 Forskjellsbehandling 19
- 4.5 Sammenheng mellom forskjellsbehandlingen og et diskrimineringsgrunnlag 20
- 4.6 Adgangen til lovlig forskjellsbehandling 22
- 4.7 Tidspunkt for diskriminering – særlig om innholdsbasert diskriminering 24
- 4.8 Den diskrimineringsrettslige bevisbyrderegelen 24
- 4.9 Sammenfatninger og refleksjoner om veien videre 27

### Kapittel 5

#### Aktørene: hvem er vernet mot diskriminering, og hvem kan holdes ansvarlig? 29

- 5.1 Innledning 29
- 5.2 Diskrimineringsvernedede grupper 29
- 5.3 Hvem skal påstanden om diskriminering rettes mot? 30
- 5.4 Kan forbudene mot instruksjon og medvirkning utvide kretsen av hvilke aktører som kan holdes ansvarlig? 31
- 5.5 Sammenfatninger og refleksjoner om veien videre 32

## **Kapittel 6**

### **Håndheving og sanksjoner** 34

2.1 Innledning 34

6.2 Organisasjoners mulighet til å få håndhevet spørsmål om algoritmisk diskriminering 34

6.3 Diskrimineringsnemndas og ordinære domstolars sanksjonskompetanse 35

6.4 Sammenfatninger og refleksjoner om veien videre 36

### **Litteratur** 39

### **Kilder** 41





# SAMMENDRAG

**For å sikre et robust vern** mot diskriminering i møte med ny teknologi, er det viktig at lovverket utvikles slik at det kan fange opp nye problemstillinger. Denne utredningen foretar en gjennomgang og analyse av sentrale bestemmelser og vilkår i likestillings- og diskrimineringsloven, med hovedvekt på forbudet mot diskriminering. Utgangspunktet er at dagens lovverk kommer til anvendelse på algoritmisk diskriminering. Samtidig viser utredningen en rekke utfordringer ved regelverket.

Samlet sett synliggjør utredningen at det er behov for en gjennomgang av likestillings- og diskrimineringsloven med tanke på å innføre særregulering som kan sikre at algoritmisk diskriminering fanges opp på en bedre og tydeligere måte enn i dag. Særregulering rettet mot algoritmisk diskriminering vil være viktig for å tydeliggjøre muligheten til å få håndhevet forbudet mot diskriminering i enkeltsaker. Slik regulering vil imidlertid også kunne ha en preventiv effekt, særlig overfor offentlige og private aktører som bruker algoritmiske systemer som del av sin virksomhet.

KAPITTEL 3 viser noen sentrale koblinger fra diskrimineringsforbudet i likestillings- og diskrimineringsloven – til EUs KI-forordning og EUs personvernforordning (GDPR).

- KI-forordningen presiserer at EU-rettens diskrimineringsregler ikke berøres av forordningen. Dette understreker viktigheten av å analysere og utvikle innholdet i det individuelle forbudet mot diskriminering, på diskrimineringsregelverkets egne premisser.
  - KI-forordningen artikkel 10 nr. 5 binder sammen de tre regelsettene (diskrimineringsforbudet, KI-regulering og GDPR). Bestemmelsen gir, på nærmere definerte vilkår, en åpning for å behandle særlige kategorier av personopplysninger dersom formålet er å forhindre at det finner sted diskriminering gjennom bruk av KI-systemer. I tiden fremover vil denne bestemmelsen kunne bli et viktig virkemiddel for å forhindre at diskriminering finner sted.
  - Kapitlet løfter frem muligheten for å innføre en egen bestemmelse i likestillings- og diskrimineringsloven om aktivitets- og redegjørelsesplikt for aktører som tar i bruk automatiserte og beslutningsstøttende algoritmer og KI-systemer i sin virksomhet.
- KAPITTEL 4 gir en analyse av likestillings- og diskrimineringslovens forbud mot direkte og indirekte diskriminering.
- Kapitlet viser at grenseoppgangen mellom direkte og indirekte forskjellsbehandling kan være krevende i møte med komplekse algoritmiske systemer og kunstig intelligens. Diskriminering kan skje gjennom systemers vektlegging av stedfortrederfaktorer («proxy discrimination»). Dette utløser nye problemstillinger, særlig i møte med forbudet mot indirekte diskriminering. Videre har EU-domstolen gjennom sin praksis utvidet rammene for hva som skal anses å utgjøre direkte forskjellsbehandling. I lys av disse utviklingstrekkene pekes det på muligheten for å innføre egne definisjoner av direkte og indirekte algoritmisk forskjellsbehandling i likestillings- og diskrimineringsloven.

- Det kan også vurderes å lage en egen unntaksbestemmelse for situasjoner med algoritmisk forskjellsbehandling. Dagens unntaksbestemmelse i likestillings- og diskrimineringsloven § 9 bygger på at vernet mot diskriminering er særlig sterkt i arbeidslivet, og grensen mellom direkte og indirekte forskjellsbehandling har betydning for unntaksadgangens rekkevidde. Utredningen viser at algoritmisk forskjellsbehandling utfordrer begge disse utgangspunktene. Algoritmiske systemer brukes på tvers av alle samfunnsområder og bidrar på ulike måter til avgjørelser som kan ha svært stor betydning for individer. Dette gjør at utgangspunktet om at arbeidslivsfeltet er særlig viktig, bør nyanseres. Videre gjør kompleksiteten i algoritmiske systemer at det kan være vanskelig å trekke en klar grense mellom direkte og indirekte forskjellsbehandling (også kommentert i strekpunktet over). I lys av disse forholdene presenteres muligheten for å innføre en egen og mer konkretisert unntaksbestemmelse for algoritmisk forskjellsbehandling i loven.
- Kapittelet viser at den diskrimineringsrettslige bevisbyrderegelen kan utgjøre en nøkkel i møte med algoritmisk diskriminering, særlig slik regelen er presisert i det nye lønnstransparensdirektivet (direktiv 2023/970) artikkel 18. Dette har sammenheng med at saker om likelønn kjennetegnes av manglende transparens om hva som ligger til grunn for lønnsfastsettelse. Her er det dermed en tydelig parallell til spørsmål om algoritmisk diskriminering.

#### KAPITTEL 5 retter oppmerksomheten mot diskrimineringsvernets aktører.

- Kapittelet tar opp likestilling- og diskrimineringslovens uttømmende angivelse av hvilke grupper som er vernet mot diskriminering. Algoritmiske systemer kan imidlertid føre til at også andre grupper utsettes for forskjellsbehandling. For eksempel faller faktorer som sosial status, helse og utdanningsnivå utenfor dagens opplisting av diskrimineringsgrunnlag. Algoritmisk diskriminering synliggjør et behov for å diskutere om det bør innføres en fleksibilitet i lovens angivelse av diskrimineringsgrunnlag. Dette kan for eksempel gjøres gjennom en særbestemmelse som forbyr algoritmisk diskriminering. Bestemmelsen kan inneholde en positiv opplisting av diskrimineringsgrunnlag (jf. ldl. § 6), men også ivareta behovet for fleksibilitet gjennom å inkludere en åpen sekkekategori.

- Kapittelet synliggjør at et individ som mener seg diskriminert skal rette en påstand om diskriminering mot den private eller offentlige aktøren som har fattet en beslutning, utformet et vedtak mv. overfor individet, og som – som ledd i sin virksomhet – har brukt et algoritmiske system.
- Kapittelet viser at likestillings- og diskrimineringslovens forbud mot instruks og mot medvirkning kan bidra til å utvide kretsen av hvilke aktører som kan holdes ansvarlig for algoritmisk diskriminering, for eksempel slik at også utviklere i noen tilfeller kan holdes ansvarlig. Innholdet i de to forbudene er imidlertid lite utviklet.

#### KAPITTEL 6 tar opp håndheving og sanksjonering av algoritmisk diskriminering.

- Kapittelet viser at organisasjoners mulighet til å få håndhevet spørsmål om algoritmisk diskriminering kan vise seg å bli viktig i tiden fremover. Dette må ses i sammenheng med at det ikke nødvendigvis vil finnes enkeltindivider som ønsker å bringe en sak om algoritmisk diskriminering inn for Diskrimineringsnemnda eller for ordinære domstoler.
- De EU/EØS-rettslige diskrimineringsdirektivene stiller krav om effektiv sanksjonering av forbudet mot diskriminering. Dette kravet gjelder også i tilfeller der det er en organisasjon som er klager eller saksøker.
- Nemndas begrensede sanksjonskompetanse i saker utenfor arbeidslivet gjør at det kan bli særlig aktuelt å bringe saker om algoritmisk diskriminering inn for ordinære domstoler. Dette synliggjør at det kan være behov for en diskusjon om hva som skal være nemndas rolle på dette feltet.
- Det vil bli viktig med bevissthet om LDOs og ulike interesseorganisasjoners rolle i å sikre at spørsmål om algoritmisk diskriminering håndheves og kontrolleres i praksis. Dette kan handle om en bevisst strategi når det gjelder muligheten til å bringe saker inn for domstolene eller for Diskrimineringsnemnda. Det kan også handle om arbeid med å sikre at det i praksis finner sted tilsyn og kontroll med algoritmiske systemer og slike systemers virkninger – sett opp mot diskrimineringsforbudet.

Arbeidet videre bør ta utgangspunkt i det materielle forbudet mot diskriminering, slik det i dag er regulert i ldl. §§ 6-9 – og slik det eventuelt kan bli justert i lys av de nye utfordringene som bruken av algoritmiske systemer fører med seg.



# KAPITTEL 1

## Innledning

Likestillings- og diskrimineringsloven<sup>1</sup> er som utgangspunkt innrettet for å håndtere spørsmål om usaklig forskjellsbehandling i mellommenneskelige relasjoner, enten forskjellsbehandlingen finner sted på en arbeidsplass, ved en utdanningsinstitusjon, i forbindelse med leie av bolig, eller i utelivet. Bruk av algoritmiske systemer og kunstig intelligens (KI) utfordrer det opprinnelige bildet av hvordan diskriminering finner sted.

Det er vel dokumentert at algoritmer og KI kan føre til krenkninger av enkeltindividers grunnleggende rettigheter, herunder retten til ikke å bli diskriminert. Når offentlige og private virksomheter tar i bruk ny teknologi, og når spørsmål om diskriminering oppstår som følge av algoritmers bidrag til handlinger og beslutninger som retter seg mot enkeltindivider, blir likestillings- og diskrimineringslovens regler og kategorier utfordret.

Forbudet mot diskriminering er relativt fleksibelt i sin utforming. Dette har sammenheng med at forbudet gjelder på tvers av alle livs- og rettsområder. Det har vist seg hensiktsmessig med bredt formulerte vilkår for diskriminering, som ved håndheving har gjort det mulig å trekke inn kontekstuelle forhold.<sup>2</sup> At diskrimineringsforbudet har en innebygget fleksibilitet, har bidratt til at vernets innhold har kunnet utvikle seg over tid. Slik har det kunnet respondere på utfordringer som etter hvert har kommet til syne. EU-domstolen har spilt en sentral rolle i denne sammenheng, for

eksempel ved å bidra til å utvikle vernet mot indirekte diskriminering, vernet mot tilknytningsdiskriminering og en egen diskrimineringsrettslig bevisbyrde-regel.

Samtidig pågår det diskusjoner om hvordan forbudet mot diskriminering skal tolkes. For eksempel kan grenseoppgangen mellom direkte og indirekte diskriminering være krevende. Det er også diskusjon om hvordan diskrimineringsvernet bør utvikles videre. Både i Norge og i EU diskuteres for eksempel spørsmålet om hvilke grupper som skal inkluderes i opplistingen av diskrimineringsgrunnlag. I en rapport om algoritmer, diskriminering og EUs diskrimineringsvernbestemmelser skriver Gerards og Xenidis at «algorithmic discrimination shines a new light on many of the ‘traditional’ problems and critiques of EU gender equality and non-discrimination law.»<sup>3</sup>

Dette er en treffende beskrivelse. Ny teknologi møter et regelverk som i årtier har blitt utviklet og brukt på forskjellsbehandling som har funnet sted på helt andre arenaer, med særlig fokus på diskriminering i arbeidslivet. Kompleksiteten som omkranser bruken av algoritmiske systemer – herunder KI-systemer – gjør at regelverket om diskriminering konfronteres med helt nye problemstillinger. Dermed må de grunnleggende diskrimineringsrettslige bestemmelsene og vilkårene vurderes i lys av en ny kontekst.

Denne utredningen belyser møtetpunktet mellom likestillings- og diskrimineringslovens forbud mot diskriminering og diskrimineringsproblematikk utløst gjennom bruk av algoritmer og kunstig intelligens.

<sup>1</sup> Lov 16. juni 2017 nr. 51 om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven).

<sup>2</sup> Diskrimineringsforbudet kan ses som utslag av såkalt «contextual equality», se Wachter, Mittelstadt og Russell (2021).

<sup>3</sup> Gerards og Xenidis (2020) s. 9.

# KAPITTEL 2

## Utredningens nærmere innretning

### 2.1 Siktemål

Utredningen søker å identifisere og synliggjøre hvilke muligheter og begrensninger som ligger i likestillings- og diskrimineringslovens forbud mot diskriminering – sett opp mot diskrimineringsutfordringer utløst ved bruk av algoritmiske systemer.

Videre søker utredningen å peke på hvilket rom som kan finnes for å utvikle diskrimineringsvernet videre, enten gjennom tolkning eller gjennom regelverksutvikling. Ved dette søker utredningen å gi et bidrag til fremtidig diskusjon om hva som skal til for å sikre et robust diskrimineringsvern i møte med den raske teknologiske utviklingen.

I utredningen knyttes det løpende kommentarer til det nærmere forholdet mellom den norske likestillings- og diskrimineringsloven og EU/EØS-rettens vern mot diskriminering. Dette er av interesse, fordi det på noen punkter er forskjeller mellom regelsettene som kan være relevante i møte med spørsmål om algoritmer, kunstig intelligens og diskriminering.

Utredningen legger til grunn at det individuelle forbudet mot diskriminering utgjør et sentralt utgangspunkt for håndtering av algoritmisk diskriminering. Dette vil si at kunnskap om diskrimineringsforbudets innhold og rekkevidde ikke bare har betydning ved vurderinger av om diskriminering har funnet sted når et system først er satt i drift. Innsikt i diskrimineringsvernets innhold og rekkevidde er også av betydning for fasene der algoritmiske systemer (herunder systemer som tar i bruk KI) utvikles og testes.<sup>4</sup>

---

4 At innholdet i det individuelle vernet må ha betydning også i fasene forut for at et system settes i drift, er lagt til grunn og utviklet i doktorgraden til Mathias Karlsen Hauglid, se Hauglid (2023) s. 14–15.

### 2.2 Begrepsbruk

En algoritme er en samling instruksjoner som beskriver steg for steg hva som skal gjøres for å løse et problem eller oppnå et bestemt resultat.<sup>5</sup> Algoritmer kan brukes til mange formål.

Det kan trekkes opp et skille mellom faste algoritmer og lærende algoritmer. Mens faste algoritmer kjennetegnes av å være regelbundne og basert på en bestemt instruks om fremgangsmåte slik at det er forutsigbart hva systemet vil komme frem til, er lærende algoritmer noe som utvikler seg dynamisk. Både faste og lærende algoritmer er av relevans i møte med spørsmål om diskriminering.

Uttrykket kunstig intelligens (KI) er vidt, idet det rommer mange former for algoritmer, maskinlæring, modeller og statistiske metoder.<sup>6</sup> 13. mars 2024 ble KI-forordningen (AI Act) vedtatt i EU Parlamentet.<sup>7</sup> Uttrykket KI-system er definert på følgende måte i forordningen artikkel 3 (1):

*‘AI system’ is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.*

---

5 Strümke (2023) s. 18.

6 Digdir, «Hva er kunstig intelligens?» (nettressurs).

7 *European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act)*, P9\_TA(2024)0138.

Det fremkommer av definisjonen at KI-systemer dreier seg om systemer som «may exhibit adaptiveness after deployment». Dette vil si at systemet skal kunne ha evnen til å endre seg eller tilpasse seg etter at det er tatt i bruk. I fortalet til KI-forordningen kommer det frem at denne definisjonen er basert på «key characteristics of AI systems», nemlig «their capability to infer». Definisjonen skal ikke omfatte «simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations.»<sup>8</sup> Ved dette er det trukket opp en grense mellom faste algoritmer og lærende algoritmer, der det kun er lærende algoritmer som inkluderes i forordningens definisjon av å være et «KI-system».<sup>9</sup> Videre i utredningen bruker jeg uttrykket «KI-system» på samme måte som i KI-forordningen.

Det finnes ulike former for lærende algoritmer. I diskusjoner om KI og diskriminering rettes det ofte oppmerksomhet mot bruken av maskinlæringssystemer. Slike systemer kjennetegnes av algoritmer som læres opp til å finne mønstre i store datamengder og til autonomt å tilpasse seg, utvikle seg og forbedre seg på bakgrunn av dataene. En annen form for lærende algoritme er såkalte «deep learning» algoritmer. Det finnes også kunstig intelligente systemer som evner å koble sammen ulike teknologier og algoritmer.<sup>10</sup>

I denne utredningen omfatter uttrykket «algoritmisk diskriminering» hele bredden av systemer, det vil si at både faste algoritmer og lærende algoritmer (herunder KI og maskinlæringssystemer) er omfattet.<sup>11</sup> Jeg bruker også uttrykket «algoritmiske systemer». Å kun snakke om «KI-systemer» ville vært for snevert, fordi faste algoritmer da ville falt utenfor. Uttrykket «KI-systemer» brukes der tematikken som tas opp er særlig aktuell for slike.

Når algoritmer har betydning for avgjørelser som tas overfor enkeltpersoner, vil graden av menneskelig påvirkning variere. På den ene side finnes fullautomatiserte beslutninger som kan komme fra både faste algoritmer og lærende algoritmer.<sup>12</sup> På den annen side finnes situasjoner der algoritmiske systemer brukes

som støtte for et menneske (for eksempel en lege, en arbeidsgiver eller en offentlig saksbehandler), som tar den endelige avgjørelsen. Både bruk av fullautomatiserte og beslutningsstøttende algoritmer og KI-systemer kan utløse spørsmål om diskriminering – og begge deler omfattes av denne utredningen.<sup>13</sup>

### 2.3 Avgrensninger

Videre i utredningen vil jeg i liten grad forholde meg til de tekniske sidene knyttet til bruk av algoritmer og KI-systemer. Jeg går heller ikke nærmere inn på typetilfeller som viser hvordan algoritmisk diskriminering kan finne sted. Eksempler på situasjoner som kan utløse spørsmål om diskriminering er for eksempel gitt av Borgesius. Han trekker frem systemer som omhandler bekjempelse av kriminalitet, seleksjon av jobbsøkere og studenter, tilpasset reklame, prisdiskriminering ved netthandel, diskriminerende bildesøkningsprogrammer og bildegjenkjenningssystemer, og automatiserte språkoversettingssystemer.<sup>14</sup>

Algoritmer og KI-systemer utvikles og settes i drift gjennom flere faser, og diskrimineringsrelevante spørsmål oppstår i alle fasene. Likestillings- og diskrimineringsombudets veileder om KI og diskriminering opererer med fem faser: planlegging, treningsdata, modellutvikling, testing, implementering og tilsyn.<sup>15</sup> I denne utredningen går jeg ikke eksplisitt inn i hvilke utfordringer som foreligger i de ulike fasene. Dette er det redegjort for både i LDOs veileder og i andre rapporter om algoritme- og KI-relatert diskriminering.<sup>16</sup> Jeg konsentrerer fremstillingen om innholdet i forbudet mot diskriminering. Som ledd i analysen av forbudet mot diskriminering, vil det imidlertid i noen grad trekkes linjer til for eksempel arbeid med å teste og utvikle et KI-system.

Utredningen fokuserer på diskrimineringsvernets individuelle side. Den strukturelle siden, herunder

<sup>8</sup> *Artificial Intelligence Act*, språkversjon vedtatt av EU Parlamentet 13. mars 2024. Se fortalet punkt 12.

<sup>9</sup> Haram (2021) s. 18.

<sup>10</sup> I sin fremstilling av algoritmer skiller Gerards og Xenidis mellom «rule-based algorithms», «machine-learning algorithms», «deep learning» og «enabling technologies and combining algorithms», se Gerards og Xenidis (2020) s. 32–36.

<sup>11</sup> Dette uttrykket brukes f.eks. også av Gerards og Xenidis (2020).

<sup>12</sup> Gerards og Xenidis (2020) s. 39–40.

<sup>13</sup> Slik også Borgesius (2018) s. 11. Se også Schartum (2019) s. 25 der Schartum trekker opp et skille mellom «rettslige beslutningssystemer» og «beslutningsstøttesystemer». Rettslige beslutningssystemer kjennetegnes ved at «resultatet fra systemet legges til grunn for vedtak», mens et beslutningsstøttesystem kjennetegnes av at datasystemene utgjør en støtte for saksbehandling, men at det er et menneske som har kontroll over det endelige innholdet av et vedtak.

<sup>14</sup> Borgesius (2018) s. 23 flg.

<sup>15</sup> Likestillings- og diskrimineringsombudet, *Innebygd diskrimineringsvern* (2023).

<sup>16</sup> Se for eksempel Gerards og Xenidis (2020) s. 37 flg.; og Borgesius (2018) s. 15 flg.

linjen til likestillings- og diskrimineringslovens bestemmelser om aktivitets- og redegjøringsplikt (ARP) blir kun kort berørt.<sup>17</sup>

KI-systemer kan føre til diskriminering, men slike systemer kan også være et viktig redskap som kan tas i bruk for å avdekke systemisk ulikhet og mulig diskriminering. KI-systemer kan dermed utgjøre redskap for bekjempelse av diskriminering.<sup>18</sup> Mulighetene dette gir vil ikke bli nærmere behandlet i utredningen.

## 2.4 Rettskilder

Forbudet mot diskriminering i likestillings- og diskrimineringsloven er utgangspunktet for utredningen. Det finnes lite nasjonal og internasjonal håndhevingspraksis som direkte kaster lys over diskrimineringsforbudets innhold og betydning i møte med algoritme- og KI-relatert diskriminering.

På nasjonalt plan foreligger det ennå ikke litteratur som omhandler diskrimineringsforbudets betydning i møte med algoritmer og KI-systemer. Imidlertid har det vært rettet en del oppmerksomhet mot algoritmisk diskriminering i rapporter og litteratur som omhandler EU/EØS-rettens bestemmelser om diskrimineringsvern.<sup>19</sup> I og med at likestillings- og diskrimineringsloven gjennomfører EU/EØS-rettens diskrimineringsdirektiver i norsk rett, har disse arbeidene interesse for denne utredningen.<sup>20</sup>

Mathias Karlsen Hauglid leverte i november 2023 avhandlingen «Bias and Discrimination in Clinical Decision Support Systems Based on Artificial Intelligence» til bedømmelse for graden Ph.D. ved Universitetet i Tromsø. Den trekker inn EU/EØS-rettens bestemmelser om vern mot diskriminering med utgangspunkt i fasene før et KI-system settes i drift og med spissing inn mot helsesektoren. Avhandlingen har vært til stor inspirasjon i arbeidet med denne utredningen.

---

<sup>17</sup> Se utredningen punkt 3.4.

<sup>18</sup> Dette er også noe som er aktuelt for nasjonale organer som arbeider med spørsmål om likestilling og ikke-diskriminering («Equality bodies»), slik som LDO. Se *Proposal for a Council directive on standards for equality bodies in the field of equal treatment* [...]. COM(2022) 689 final, 7. desember 2022, se fortalen punkt 20; og *Proposal for a directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men* [...]. COM(2022) 688 final, 7. desember 2022, se fortalen punkt 19.

<sup>19</sup> Se for eksempel Borgesius (2018); Gerards og Xenidis (2020); Wachter, Mittelstadt og Russell (2021); Wachter (2022).

<sup>20</sup> Jeg vil i denne sammenheng opplyse om at jeg er medlem av komiteen som er oppnevnt for å bedømme avhandlingen.

I utredningen trekker jeg inn praksis fra EU-domstolen der det finnes saker som kan ha overføringsverdi fra en problemstilling og over til problemstillingen knyttet til algoritmisk diskriminering. Det samme gjelder praksis fra Diskrimineringsnemnda.<sup>21</sup>

Som ledd i å synliggjøre sammenhengen mellom diskrimineringsvern, personvernregulering og KI-regulering, trekker jeg kort inn EUs personvernforordning (GDPR) og EUs KI-forordning, samt rapporter, litteratur, pressemeldinger mv. som er egnet til å kaste lys over forordningenes innhold.

## 2.5 Struktur

Videre består utredningen av fire kapitler.

Kapittel 3 tar for seg forholdet mellom forbudet mot diskriminering og reguleringen av henholdsvis kunstig intelligens og personvern.

Kapittel 4 omhandler diskrimineringsvernets materielle innhold med utgangspunkt i lovens forbud mot direkte og indirekte diskriminering. Følgende problemstillinger drøftes: Hva betyr lovens vilkår for diskriminering i møte med algoritmisk diskriminering? På hvilket tidspunkt skal diskrimineringsvurderingen foretas, og hva med innholdsbasert diskriminering? Kan den diskrimineringsrettslige bevisbyrderegelen spille en rolle i saker om algoritme- og KI-relatert diskriminering?

Kapittel 5 tar for seg aktørene. Et første sentralt spørsmål omhandler diskrimineringsvernets subjekter: Hvilke personer og grupper er omfattet av forbudet mot diskriminering? Et neste spørsmål handler om hvem som kan holdes ansvarlig dersom det foreligger en situasjon med diskriminering. Det er gjerne flere ulike aktører involvert når algoritme- og KI-systemer utvikles og tas i bruk. Er det utvikleren eller den som bruker systemet som kan holdes ansvarlig, eller begge? Kan forbudene mot instruksjon og medvirkning utvide kretsen av hvilke aktører som kan holdes ansvarlig?

I kapittel 6 er temaet håndheving og sanksjonering. Kan organisasjoner få håndhevet spørsmål om algoritmisk diskriminering? Og hvilke sanksjoner er aktuelle? Er det forskjeller på Diskrimineringsnemndas sanksjonskompetanse og kompetansen til ordinære domstoler?

Til slutt i hvert kapittel gis en sammenfatning over sentrale temaer som er identifisert underveis, samt refleksjoner over hvilke muligheter som kan finnes til eventuelt å utvikle et mer robust vern mot algoritmisk diskriminering. Litteratur- og kildeoversikt er inntatt helt til sist i utredningen.

---

<sup>21</sup> Det vises kun til én sak fra nemnda, se punkt 4.8.

# KAPITTEL 3

## Linjer fra forbudet mot diskriminering til regulering av kunstig intelligens og personvern

### 3.1 Innledning

Diskriminering som oppstår gjennom bruk av algoritmer og KI-systemer er ikke eksplisitt tatt opp verken i lovteksten i likestillings- og diskrimineringsloven eller i lovens forarbeider. Det er imidlertid slått fast i ldl. § 2 (1) at: «Loven gjelder på alle samfunnsområder». Utgangspunktet er derfor at forbudet mot diskriminering kommer til anvendelse på spørsmål om algoritmisk diskriminering. Det er også lagt til grunn at EU/EØS-rettens diskrimineringsbestemmelser og ulike menneskerettskonvensjoners diskrimineringsbestemmelser kommer til anvendelse på spørsmål om KI-relatert diskriminering.<sup>22</sup>

Likestillings- og diskrimineringsloven etablerer et diskrimineringsvern for enkeltpersoner. Det dreier seg om lovbestemmelser som trekker opp grensen mellom lovlig og ulovlig forskjellsbehandling. Forskjellsbehandling som ikke har en tilstrekkelig begrunnelse, utgjør diskriminering og er ulovlig.

---

<sup>22</sup> Borgesius (2018) s. 31 flg. En slik forståelse ligger også til grunn i Gerards og Xenidis (2020). LDOs veileder om kunstig intelligens slår fast samme utgangspunkt: «Likestillings- og diskrimineringsloven er teknologinøytral. Det betyr at forbudet mot diskriminering gjelder uavhengig av om den diskriminerende praksisen eller avgjørelsen er tatt av en person eller følger av et kunstig intelligent system.» Se Likestillings- og diskrimineringsombudet, *Innebygd diskrimineringsvern* (2023).

Likestillings- og diskrimineringsloven befinner seg imidlertid ikke i et rettslig vakuum, og algoritmers og KI-systemers virkninger i møte med enkeltindividers rettigheter er et sentralt tema også i andre regelsett. I det følgende beskriver jeg overordnede trekk ved KI-forordningen og ved personvernforordningen (GDPR), som er de to sentrale regelverk som på ulike måter har en berøringsflate med vernet mot diskriminering.

### 3.2 KI-forordningen

Når EUs KI-forordning trer i kraft, blir den verdens første lov som spesifikt regulerer KI.<sup>23</sup> KI-forordningen er primært en produktansvarslov. Den har en risikobasert tilnærming til ulike KI-systemer, og omfanget av regulering og plikter har sammenheng med hvilken kategori et KI-system anses å falle inn under. De fleste forpliktelsene ligger på tilbydere av KI-systemer, og gjennom dette når KI-forordningen en aktør som ikke like lett kan holdes ansvarlig etter regelverket om diskriminering.<sup>24</sup> Også brukere av KI-systemer, for

---

<sup>23</sup> European Parliament, *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI* (2023).

<sup>24</sup> Se imidlertid utredningens punkt 5.4 som tar opp om forbudene mot instruksjon og mot medvirkning kan utvide kretsen av hvilke aktører som kan holdes ansvarlig.

eksempel virksomheter som bruker KI-systemer i sin daglige drift, omfattes av noen forpliktelser.<sup>25</sup> Disse vil også gjerne være pliktsubjekt etter likestillings- og diskrimineringsloven.<sup>26</sup>

KI-forordningen etablerer fire hovedkategorier for KI-systemers risiko: uakseptabel risiko (slike systemer er forbudte), høyrisiko, begrenset risiko og minimal risiko. KI-systemer som «pose a high risk of harm to the health and safety or the fundamental rights of persons», herunder forbudet mot diskriminering, skal som utgangspunkt klassifiseres som høyrisiko-systemer og vil omfattes av strenge produktkrav.<sup>27</sup> Dette innebærer for eksempel en plikt til å foreta såkalte «fundamental rights impact assessments». Jeg ser ikke nærmere på dette.

Når det gjelder forholdet mellom KI-forordningen og annen EU-rettslig regulering, herunder regulering av vernet mot diskriminering, heter det i fortalen punkt 45:

*«Practices that are prohibited by Union legislation, including data protection law, non-discrimination law, consumer protection law, and competition law, should not be affected by this Regulation.»*

Dette er et viktig utgangspunkt, idet det markerer at innholdet i forbudet mot diskriminering – slik dette er å forstå på bakgrunn av de EU-rettslige diskrimineringsdirektivene – skal legges til grunn i møte med KI-reguleringen. Punktet synliggjør at det er viktig med selvstendige analyser av diskrimineringsforbudets innhold og betydning.

AI Act skal tre i kraft stegvis slik at hele forordningen er implementert i løpet av 2026. Implementering i

norsk lov er ikke avklart, men det er grunn til å tro at dette vil skje.

### 3.3 Personvernforordningen (GDPR)

EUs personvernforordning (GDPR) inneholder et omfattende regelverk om personvern og er gjennomført i norsk rett gjennom personopplysningsloven.<sup>28</sup> GDPR inneholder blant annet krav om at det skal foretas en rekke vurderinger i forkant av at det behandles personopplysninger. Hovedpliktsubjektet etter GDPR er aktøren som «bestemmer formålet med behandlingen av personopplysninger og hvilke midler som skal benyttes», nærmere definert som den «behandlingsansvarlige», jf. GDPR artikkel 4 nr. 7. Dette vil i mange tilfeller kunne være det samme pliktsubjektet som etter likestillings- og diskrimineringsloven, for eksempel en offentlig eller privat arbeidsgiver eller en offentlig myndighet. I mange tilfeller vil dette også kunne være en aktør som har forpliktelser etter KI-forordningen.

Kunstig intelligens er ikke nevnt eksplisitt i GDPR, men GDPR kommer til anvendelse på KI-systemers bruk av personopplysninger. Dette er blant annet utdypet i rapporten «The impact of the General Data Protection Regulation (GDPR) on artificial intelligence» utgitt av Panel for the Future of Science and Technology på vegne av EU Parlamentet.<sup>29</sup> Rapporten synliggjør hvordan KI-systemer kan brukes på måter som er i tråd med GDPR, men også GDPR sine begrensninger og at det er behov for å utvikle og konkretisere regelverket videre.

GDPR forbyr ikke eksplisitt diskriminering, men det er erkjent at krenkelser av personvernet også kan føre til diskriminering.<sup>30</sup> I slike tilfeller trekker reguleringen av personvern og reguleringen av diskrimineringsvern i samme retning. Det finnes mekanismer i GDPR som kan bidra til å fremme likebehandlingsprinsippet og vernet mot diskriminering.<sup>31</sup> Et eksempel på dette er den behandlingsansvarliges plikt til å foreta vurderinger av personvernkonsekvenser (Data Protection Impact Assessment, DPIA) i forkant av at personopplysninger behandles.

Det er stor grad av overlapp mellom det som utgjør særlige kategorier av personopplysninger etter GDPR

<sup>25</sup> Advokatfirmaet Thommessen ved Werring, Clausen og Mossige, AI Act: Fremtiden for kunstig intelligens i EU (2023).

<sup>26</sup> I tillegg kan disse være hovedpliktsubjekt etter GDPR, se utredningens punkt 3.3.

<sup>27</sup> *European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act)*, P9\_TA(2024)0138. Se fortalen punkt 52. Fortalen punkt 48 lister opp følgende eksempler på fundamentale rettigheter: «the right to human dignity, respect for private and family life, protection of personal data, freedom of expression and information, freedom of assembly and of association, and non-discrimination, right to education consumer protection, workers' rights, rights of persons with disabilities, gender equality, intellectual property rights, right to an effective remedy and to a fair trial, right of defence and the presumption of innocence, right to good administration.»

<sup>28</sup> Lov 15. juni 2018 nr. 38 om behandling av personopplysninger (personopplysningsloven).

<sup>29</sup> European Parliament. Directorate General for Parliamentary Research Services, rapport ved Sartor og Lagioia (2020).

<sup>30</sup> EUs personforordning (GDPR), fortalen punkt 75.

<sup>31</sup> Naudts (2019).

og vernede diskrimineringsgrunnlag etter likestillings- og diskrimineringsloven. GDPR artikkel 9 nr. 1 lyder:

*Behandling av personopplysninger om rasemessig eller etnisk opprinnelse, politisk oppfatning, religion, filosofisk overbevisning eller fagforeningsmedlemskap, samt behandling av genetiske opplysninger og biometriske opplysninger med det formål å entydig identifisere en fysisk person, helseopplysninger eller opplysninger om en fysisk persons seksuelle forhold eller seksuelle orientering, er forbudt.*

Artikkel 9 nr. 2 gir imidlertid åpning for at slike opplysninger kan behandles på nærmere definerte vilkår.<sup>32</sup>

I juridisk litteratur er det diskutert om forbudet mot behandling av særlige kategorier av personopplysninger etter GDPR artikkel 9 utgjør en hindring i møte med KI-relatert diskriminering. Bekkum og Borgesius beskriver problemstillingen slik:

*To assess whether its AI system harms people with a certain ethnicity, the organisation needs to know the ethnicity of its job applicants. In Europe, however, the organisation may not know the ethnicity because, in principle, the GDPR prohibits the use of 'special categories of data' (article 9). Special categories of data include data on ethnicity, religion, health, and sexual preferences. Hence it appears that the organisation is not allowed to infer, collect, or use the ethnicity of the applicants.<sup>33</sup>*

Forfatterne diskuterer blant annet om det bør innføres en egen hjemmel i GDPR som åpner for å behandle særlige kategorier av personopplysninger dersom formålet er å forhindre KI-relatert diskriminering. De gir også eksempler på at slike unntak er innført i enkelte land, for eksempel i England.

Tematikken illustrerer at personvernregulering og diskrimineringsvernsregulering ikke nødvendigvis alltid trekker i samme retning.

KI-forordningen inneholder en bestemmelse som løser opp i den beskrevne spenningen. Artikkel 10 nr. 5 slår fast at særlige kategorier av personopplysninger slik dette blant annet er definert i GDPR artikkel 9 nr. 1 kan behandles med det formål å sikre «bias detection» og «correction in relation to the high-risk AI systems», dersom dette er «strictly necessary» og «subject to appropriate safeguards for the fundamental rights and freedoms of natural persons.» Bestemmelsen

inneholder en rekke krav som må være oppfylt for at slik behandling av særlige kategorier av personopplysninger skal være tillatt. Formålet med bestemmelsen er å beskytte personer mot «the discrimination that might result from the bias in AI systems.»<sup>34</sup> Gjennom denne bestemmelsen blir diskrimineringsregelverket, GDPR og KI-forordningen eksplisitt knyttet sammen.

Etter mitt syn vil det i tiden som kommer være viktig å følge med på diskusjoner som omhandler KI-forordningen artikkel 10 nr. 5 om adgang til innhenting av særlige kategorier av personopplysninger – og på hvilken måte dette kan brukes som et middel for å forhindre diskriminering. I denne sammenheng bør det påses at det legges til grunn en korrekt forståelse av diskrimineringsforbudets materielle innhold.

### 3.4 Sammenfatninger og refleksjoner om veien videre

Viktige dimensjoner som er tatt opp i kapittel 3:

- Både KI-forordningen og GDPR forholder seg til spørsmål om diskriminering, og begge regelsett inneholder krav om vurderinger som skal foretas før virksomhet kan igangsettes på en lovlig måte. De har ved dette et prosessuelt hovedfokus. Likestillings- og diskrimineringslovens forbud mot diskriminering skiller seg fra dette ved at lovens omdreiningspunkt er et materielt vern mot diskriminering.
- KI-forordningen presiserer at EU-rettens diskrimineringsregler ikke berøres av forordningen. Dette understreker viktigheten av å analysere og utvikle innholdet i det individuelle forbudet mot diskriminering, på diskrimineringsregelverkets egne premisser.
- Det er KI-forordningen som klartest legger plikter på tilbydere (herunder utviklere) av KI-systemer. Pliktsubjektet etter GDPR og etter likestillings- og diskrimineringsloven vil som hovedregel være aktøren som tar et KI-system i bruk, for eksempel en offentlig eller privat arbeidsgiver eller en offentlig myndighet.<sup>35</sup>

<sup>32</sup> Disse ser jeg ikke nærmere på her.

<sup>33</sup> van Bekkum og Borgesius (2023). Se også Borgesius (2018) s. 45.

<sup>34</sup> *Artificial Intelligence Act*, språkversjon vedtatt av EU Parlamentet 13. mars 2024. Se fortalen punkt 70.

<sup>35</sup> Se imidlertid utredningens punkt 5.4 om forbudene mot instruksjon og mot medvirkning.

- Det foreligger en spenning mellom GDPR artikkel 9 nr. 1 om særlige kategorier av personopplysninger og behovet for å behandle opplysninger om etnisitet, religion mv. som ledd i arbeid med å forhindre KI-relatert diskriminering. Denne spenningen er søkt håndtert gjennom KI-forordningen artikkel 10 nr. 5, som gir en åpning for å behandle særlige kategorier av personopplysninger i forbindelse med «bias detection». Dette koblingspunktet mellom diskrimineringsforbudet, personvernregulering og KI-regulering er det god grunn til å følge med på i tiden fremover.

*Refleksjoner om veien videre:*

- I tillegg til bestemmelser om det individuelle vernet mot diskriminering, inneholder likestillings- og diskrimineringsloven bestemmelser om aktivt likestillingsarbeid. Lovens §§ 24, 26 og 26a omhandler offentlige myndigheters og arbeidsgiveres aktivitets- og redegjørelsesplikt (ARP). ARP gir plikt til å arbeide «aktivt, målrettet og planmessig for å fremme likestilling og hindre diskriminering» og til å redegjøre for arbeidet som gjøres. Inspirert av de strukturelle tiltakene som følger av GDPR og av KI-forordningen, kan det vurderes om det kan være hensiktsmessig å innføre ARP-lignende strukturelle forpliktelser for aktører som tar i bruk automatiserte og beslutningsstøttende algoritmer og KI-systemer. Dette ville kunne sikre at diskrimineringsproblemstikken ikke tapes av syne ved at systemenes virkninger overfor diskrimineringsvernede grupper regelmessig tas opp og evalueres innen rammen av diskrimineringslovgivningen.



# KAPITTEL 4

## Forbudet mot direkte og indirekte diskriminering

### 4.1 Innledning

Som utgangspunkt kommer både forbudet mot direkte og indirekte diskriminering til anvendelse på algoritrisk diskriminering.

Likestillings- og diskrimineringsloven §§ 6-9 regulerer diskrimineringsforbudets innhold. For at diskriminering skal kunne slås fast må det må foreligge:

1. direkte eller indirekte forskjellsbehandling,
2. sammenheng mellom forskjellsbehandlingen og et/flere diskrimineringsgrunnlag,
3. forskjellsbehandlingen må ikke ha en begrunnelse som likevel gjør den lovlig.

Disse tre vilkårene kommer til uttrykk også på internasjonalt plan, for eksempel i de EU/EØS-rettslige diskrimineringsdirektivene. Det er ikke et krav for diskriminering at noen har hatt til hensikt å diskriminere.

I det følgende behandler jeg først overordnede trekk ved forbudene mot henholdsvis direkte og indirekte diskriminering sett opp mot problemstillinger som har sammenheng med bruk av algoritmer og KI-systemer. Videre belyses innholdet i de tre vilkårene for diskriminering, spørsmål knyttet til innholdsbasert diskriminering og hvilken betydning den diskrimineringsrettslige bevisbyrderregelen kan ha i møte med algoritmisk diskriminering.

### 4.2 Noen utgangspunkter knyttet til forbudet mot direkte diskriminering

Ldl. § 7 definerer direkte forskjellsbehandling som at «en person behandles dårligere enn andre blir, har blitt eller ville blitt behandlet i en tilsvarende situasjon, på grunn av forhold som nevnt i § 6 første ledd.»

Forbudet retter seg mot situasjoner der personer direkte på grunn av sitt kjønn, sin alder, sin religion, sin etnisitet mv. behandles dårligere enn andre. Der som direkte forskjellsbehandling har en tungtveiende begrunnelse som tilfredsstillende vilkårene i ldl. § 9 om saklighet, nødvendighet og forholdsmessighet, vil den kunne være tillatt.<sup>36</sup> Utgangspunktet er imidlertid at det skal mye til for at direkte forskjellsbehandling er tillatt. Direkte forskjellsbehandling som ikke har en tilstrekkelig begrunnelse, utgjør direkte diskriminering og er forbudt.<sup>37</sup>

Forbudet mot direkte diskriminering retter seg mot forskjellsbehandling som direkte er begrunnet i et ønske om å behandle en gruppe dårligere enn andre. For at forbudet mot direkte diskriminering skal kunne komme til anvendelse i en situasjon med bruk av algoritmer og KI, må direkte forskjellsbehandling finne sted på to nivåer. For det første må direkte forskjellsbehandling av personer fra bestemte diskrimineringsvernede grupper være noe algoritmen eller KI-systemet programmeres til å gjøre, eller selv lærer

<sup>36</sup> På arbeidslivets område er det en særlig snever adgang til direkte forskjellsbehandling, se ldl. § 9 (2). Se også lovens § 10 om graviditets- og permisjonsdiskriminering.

<sup>37</sup> For mer informasjon om forbudet mot direkte diskriminering, se Ballangrud og Søbstad (2021) s. 150 flg.; og Hellum og Strand (2022) kapittel 6.

seg opp til. Den direkte forskjellsbehandlingen må altså følge av algoritmen eller KI-systemet selv – ved at gruppebaserte kjennetegn (for eksempel etnisitet, kjønn, religion) er kriterier som systemet legger vekt på til ugunst. For det andre må dette ha som utslag at systemet kommer med anbefalinger/handlinger som gjør at enkeltpersoner utsettes for direkte forskjellsbehandling nettopp «på grunn av» et vernet diskrimineringsgrunnlag (for eksempel etnisitet, kjønn, religion), jf. ldl. § 7.

Forbudet mot direkte diskriminering vil for eksempel kunne komme til anvendelse dersom en algoritme legger opp til at det skal finne sted negativ forskjellsbehandling av bestemte diskrimineringsvernedede grupper og der dette fører til at enkeltpersoner fra disse gruppene stilles dårligere enn andre. Forbudet mot direkte diskriminering kan også være aktuelt dersom negative fordommer mot bestemte grupper har funnet sin vei inn i treningsdata som er brukt til å utvikle et KI-system, og der dette – når systemet tas i bruk – resulterer i at personer fra diskrimineringsvernedede grupper utsettes for direkte forskjellsbehandling.<sup>38</sup>

Det er ikke helt klart hvor langt forbudet mot direkte forskjellsbehandling strekker seg. Det antas at forbudet ikke bare omfatter tilfeller der forskjellsbehandlingen direkte har sammenheng med et vernet gruppekjennetegn, men også der forskjellsbehandlingen er resultat av at det anvendes kriterier som er uløselig forbundet med vernede gruppekjennetegn. Den nærmere grenseoppgangen mellom direkte og indirekte forskjellsbehandling kommer jeg tilbake til nedenfor i punkt 4.5.

### **4.3 Noen utgangspunkter knyttet til forbudet mot indirekte diskriminering**

Ikke all gruppebasert forskjellsbehandling passer inn i den strukturen som er utmeislet gjennom forbudet mot direkte diskriminering. Dette er bakgrunnen for etableringen av et forbud mot indirekte diskriminering.

I ldl. § 8 defineres indirekte forskjellsbehandling som «enhver tilsynelatende nøytral bestemmelse, betingelse, praksis, handling eller unnlattelse som vil stille personer dårligere enn andre, på grunn av forhold som nevnt i § 6 første ledd.» Gjennom dette forbudet erkjennes det at regler, praksiser mv. som tilfredsstillter kravene til formell likhet, likevel kan ha som konsekvens at enkelte grupper kommer dårligere ut enn andre. I forbudet mot indirekte diskriminering

er det de forskjellsbehandlende *virkningene* som står i sentrum.

Ikke all indirekte forskjellsbehandling er som utgangspunkt forbudt. Indirekte forskjellsbehandling som tilfredsstillter kriteriene i ldl. § 9 om saklighet, nødvendighet og forholdsmessighet til kunne være tillatt. Dersom den indirekte forskjellsbehandlingen ikke har en tilstrekkelig begrunnelse, vil den utgjøre indirekte diskriminering og være forbudt.

I motsetning til forbudet mot direkte diskriminering, krever ikke forbudet mot indirekte diskriminering at gruppebaserte karakteristikk i seg selv utgjør en begrunnelse for å forskjellsbehandle. Forbudet mot indirekte diskriminering retter oppmerksomheten mot hvordan gruppebasert forskjellsbehandling kan være *resultat* av tilsynelatende nøytrale handlinger/praksiser/bestemmelser. For at forbudet mot indirekte diskriminering skal komme til anvendelse, er det tilstrekkelig at praksiser, handlinger mv. har resultert i (eller vil kunne resultere i) ulempe/forsjellsbehandling av en diskrimineringsvernet gruppe.

Det er flere trekk ved KI-systemer som gjør at forbudet mot indirekte diskriminering kan være enklere å anvende i praksis enn forbudet mot direkte diskriminering. For eksempel er det kjent at KI-systemer kan være trent opp på data som inneholder informasjon som reproducerer ulemper og stigma mot nærmere definerte grupper. Imidlertid er det ikke nødvendigvis så enkelt å identifisere nøyaktig hvilke data som er egnet til å skape en skjevhet, noe som kan gjøre det vanskelig å ta i bruk forbudet mot direkte diskriminering. Forbudet mot indirekte diskriminering forutsetter imidlertid ikke at forskjellsbehandling er begrunnet i nærmere definerte gruppekjennetegn. Det er resultatet av at KI-systemet brukes, som er det sentrale. Dersom systemet fører til at en/flere diskrimineringsvernedede grupper kommer dårligere ut enn andre, vil forbudet mot indirekte diskriminering kunne anvendes.

KI-systemers kompleksitet er en utfordring i møte med nasjonale regler som forbyr diskriminering. Noen systemer er så komplekst satt sammen at mennesker ikke fullt kan forstå hva det er som gjør at systemet handler slik det gjør. Dette omtales ofte som «black box»-problematikken. I møte med slike komplekse og uoversiktlige systemer, vil det være svært krevende å anvende forbudet mot direkte diskriminering, da det ikke er mulig å få innsyn i kriteriene

som systemet opererer med.<sup>39</sup> I slike tilfeller vil det kunne være enklere å anvende forbudet mot indirekte diskriminering.

I gjennomgangen jeg har gjort av litteratur og rapporter, uttrykkes det flere steder at forbudet mot indirekte diskriminering er mer aktuelt i møte med KI-relatert diskriminering enn forbudet mot direkte diskriminering.<sup>40</sup> Jeg er enig i dette utgangspunktet. Samtidig kan grenseoppgangen mellom direkte og indirekte diskriminering problematiseres nærmere, se punkt 4.5 nedenfor.

#### 4.4 Forskjellsbehandling

Vilkåret om forskjellsbehandling handler om at noen behandles dårligere enn andre. Dette utgangspunktet gjelder også i relasjon til algoritmisk diskriminering.

Forskjellsbehandling vil for eksempel foreligge dersom en person opplever «tap av fordeler, økonomisk tap eller færre muligheter sammenlignet med andre i en tilsvarende situasjon.»<sup>41</sup> Ved vurderingen av om det foreligger forskjellsbehandling skal det tas utgangspunkt i en objektiv vurdering av hva som vanligvis regnes som en negativ virkning.<sup>42</sup> For å kunne slå fast om det foreligger direkte eller indirekte forskjellsbehandling, må det foretas en sammenligning mellom situasjonen til personen/personer som mener seg diskriminert og situasjonen til andre personer.

For at indirekte forskjellsbehandling skal kunne slås fast, må det finne sted en sammenligning på gruppenivå ved at virkningene et KI-system har overfor en diskrimineringsvernet gruppe skal sammenlignes med hvilken virkning systemet har overfor andre. I saker om indirekte diskriminering er det ikke alltid så klart hva som skal til for å kunne slå fast at personer som tilhører en diskrimineringsvernet gruppe kommer dårligere ut enn andre. Etter norsk rett er det imidlertid ikke så mye som skal til. Det er tilstrekkelig at den dårligere behandlingen *særlig* rammer personer som tilhører en diskrimineringsvernet gruppe.<sup>43</sup> Nøyaktig hvor grensen går avgjøres gjerne gjennom håndheving – med utgangspunkt i faktum i

enkeltsaker. Det samme er karakteristisk for EU-domstolens praksis. Heller ikke her foreligger det en klar beskrivelse av hva slags grad av ulempe som kreves for at forskjellsbehandlingskriteriet er oppfylt. Wachter, Mittelstadt og Russell uttrykker det slik:

*Thresholds are ... flexible and set on a case-by-case basis, if they are explicitly set by the judiciary at all.*<sup>44</sup>

I praksis kan bevismessige utfordringer gjøre det krevende å underbygge en påstand om diskriminering.

I situasjoner med mer «tradisjonell» diskriminering, er det utviklet mekanismer som har gjort det enklere for klager/saksøkere å underbygge en påstand om forskjellsbehandling. For eksempel er det ikke et krav om at det må finnes faktiske sammenligningspersoner for å kunne underbygge en påstand om direkte forskjellsbehandling. Både etter norsk rett og etter EU/EØS-retten kan en påstand om direkte forskjellsbehandling underbygges gjennom både faktiske og hypotetiske sammenligningspersoner.<sup>45</sup> I saker om indirekte diskriminering er det vide rammer for hvordan en klager/saksøker skal kunne underbygge en påstand om forskjellsbehandling. For eksempel er det ikke noe krav om at klager/saksøker skal legge frem statistikk for å underbygge en påstand om indirekte forskjellsbehandling. Statistikk kan brukes dersom dette eksisterer, men indirekte forskjellsbehandling skal også kunne underbygges på andre måter.<sup>46</sup>

Hvordan direkte eller indirekte forskjellsbehandling forårsaket av algoritmer eller KI-systemer skal kunne underbygges i enkelttilfeller, byr imidlertid på utfordringer. For hvordan skal enkeltpersoner kunne få innsikt i de kriteriene som systemene opererer med eller i systemenes virkninger, slik at forskjellsbehandling kan slås fast?

Gerards og Xenidis uttaler følgende med hensyn til forbudet mot direkte diskriminering:

*Categorising algorithmic discrimination as direct discrimination is therefore likely to be a challenge given the opacity of particular algorithms, especially in light of the need to establish a comparator under EU law. Indeed, if the lack of transparency of the functioning of an algorithm prevents the gathering of evidence on how the algorithm has treated or would have treated a group that does not share the protected characteristic*

<sup>39</sup> Muligens vil utvikling av metoder for såkalt «explainable AI» kunne bidra her.

<sup>40</sup> Se for eksempel Gerards og Xenidis (2020) s. 67; Borgesius (2018) s. 36; og Wachter, Mittelstadt og Russell (2021) s. 15.

<sup>41</sup> Prop. 81 L (2016–2017) punkt 12.2.4.1 s. 101.

<sup>42</sup> Prop. 81 L (2016–2017) punkt 12.2.4.2 s. 102.

<sup>43</sup> I de EU/EØS-rettslige diskrimineringsdirektivene brukes uttrykket «særlig ufordelaktig». Se nærmere om dette i Hellum og Strand (2022) s. 236–238.

<sup>44</sup> Wachter, Mittelstadt og Russell (2021) s. 29.

<sup>45</sup> Hellum og Strand (2022) s. 381 flg.

<sup>46</sup> Hellum og Strand (2022) s. 383.

at stake (the comparator group), then a finding of direct discrimination might be precluded altogether.<sup>47</sup>

Hvor terskelen skal settes for hva som innebærer forskjellsbehandling i en kontekst med bruk av algoritmer og KI-systemer, er noe som bør vies oppmerksomhet fremover. Den diskrimineringsrettslige bevisbyrderegelen kan imidlertid utgjøre en viktig nøkkel, noe jeg kommer til i punkt 4.8.

Vi har sett at algoritme- og KI-relatert ulempe gir noen utfordringer i møte med vilkåret om forskjellsbehandling. Det dreier seg dels om at forskjellsbehandlingsvilkårets materielle innhold allerede i utgangspunktet er vagt. Dels handler det om bevismessige utfordringer knyttet til å få innsikt i og kunnskap om algoritmers og KI-systemers kriterier og virkninger, slik at det blir mulig å konstatere at forskjellsbehandling har funnet sted.

#### 4.5 Sammenheng mellom forskjellsbehandlingen og et diskrimineringsgrunnlag

For at diskriminering skal kunne konstateres, må det foreligge en sammenheng mellom forskjellsbehandlingen og et eller flere vernede diskrimineringsgrunnlag.<sup>48</sup>

I møte med algoritme- og KI-relatert diskriminering, vil spørsmål om sammenheng mellom forskjellsbehandling og et vernet diskrimineringsgrunnlag oppstå som en mer kompleks tematikk enn ellers. Det er nemlig ikke bare systemers negative vektlegging av diskrimineringsvernede gruppekjennetegn som kan føre til (direkte) diskriminering. I mange tilfeller bruker ikke KI-systemer vernede gruppekjennetegn i sin behandling av data. Isteden brukes faktorer som har så stor grad av nærhet til de vernede gruppekjennetegnene, slik som adresse, økonomi og utdanning, at resultatet likevel blir at personer utsettes for forskjellsbehandling som har sammenheng med et eller flere vernede gruppekjennetegn. Dette omtales gjerne som «proxy discrimination». Datta m.fl. definerer ordet «proxy» som «a feature correlated with a protected class whose use in a decision procedure can result in indirect discrimination.»<sup>49</sup>

På norsk kan vi kanskje snakke om diskriminering ved bruk av «stedfortrederfaktorer», altså at systemet tar en slags omvei om faktorer med nær forbindelse til vernede gruppekjennetegn. I slike tilfeller vil det ikke alltid være lett å få oversikt over hvilke stedfortrederfaktorer som var avgjørende for forskjellsbehandlingen som har oppstått overfor bestemte diskrimineringsvernede grupper. Forbudet mot indirekte diskriminering, med dets fokus på hvordan gruppebasert forskjellsbehandling kan være resultat av tilsynelatende nøytrale handlinger/praksiser/bestemmelser, gir dermed en åpning for å håndtere slike spørsmål. Gerards og Xenidis mener at forbudet mot indirekte diskriminering «provides a safety net for tackling proxy discrimination when there is doubt as to whether the link between a given proxy and a given protected ground is direct enough for direct discrimination to arise.»<sup>50</sup>

I sin doktorgradsavhandling ser Hauglid nærmere på EU-domstolens praksis og grenseoppgangen mellom direkte og indirekte diskriminering i tilfeller der forskjellsbehandling ikke begrunnes direkte i diskrimineringsvernede gruppekjennetegn, men også i faktorer som er uløselig forbundet med vernede gruppekjennetegn. Han er av den oppfatning at direkte diskriminering kan foreligge dersom et AI-system inkorporerer «any protected characteristics or inextricably linked features variables».<sup>51</sup> Ved hjelp av en analyse av praksis fra EU-domstolen synliggjør han hvordan domstolen i en del saker vurderer at et faktum handler om direkte forskjellsbehandling fordi det er operert med faktorer som er uløselig forbundet med vernede gruppekjennetegn – og dette utelukkende er til ulempe for en diskrimineringsvernet gruppe, eller det innebærer at alle personer fra en vernet gruppe utelukkes fra å oppnå en fordel.<sup>52</sup>

I den sammenslåtte saken *Wabe og Müller* uttalte EU-domstolen om rekkevidden av forbudet mot direkte diskriminering at «unequal treatment resulting from a rule or practice which is based on a criterion that is inextricably linked to a protected ground, in the present case religion or belief, must be regarded as being directly based on that ground.»<sup>53</sup> Saken handlet om forskjellsbehandling i arbeidslivet på grunn av bruk av religiøse plagg og symboler. I tilfeller der et regelverk er utformet slik at kun én eller noen nærmere

<sup>47</sup> Gerards og Xenidis (2020) s. 69.

<sup>48</sup> For en nærmere diskusjon av hva som ligger i dette vilkåret, se Hellum og Strand (2022) s. 249 flg.

<sup>49</sup> Datta, Fredrikson, Ko, Mardziel og Sen (2017) s. 1.

<sup>50</sup> Gerards og Xenidis (2020) s. 71.

<sup>51</sup> Hauglid (2023) s. 214.

<sup>52</sup> Hauglid (2023) s. 221-228.

<sup>53</sup> EU-domstolen, *C-804/18 og C-341/19 (sammenslåtte saker, WABE og Müller)* (2021), premiss 73.

bestemte religioner og religiøse uttrykk rammes av et forbud mens andre religiøse uttrykk ikke rammes, vil det være snakk om direkte forskjellsbehandling på grunn av religion. Saken viser dermed at direkte forskjellsbehandling også kan konstateres for en gruppe innad i en større diskrimineringsvernet gruppe. I møte med algoritmisk diskriminering vil dette for eksempel kunne bety at direkte forskjellsbehandling kan anses å foreligge der systemet har en innretning som resulterer i at utelukkende en eller noen få etniske minoritetsgrupper utsettes for forskjellsbehandling, mens andre etniske minoritetsgrupper ikke rammes.<sup>54</sup>

I tillegg til å omfatte eksisterende forhold, omfatter diskrimineringsforbudet også tilknytningsdiskriminering, samt diskriminering på grunn av «antatte, tidligere eller fremtidige forhold», jf. ldl. § 6 (2) og (3). Dette er med på å vide ut personkretsen som diskrimineringsforbudet er aktuelt for – også i møte med algoritmisk diskriminering. Det vil for eksempel si at feilkategoriseringer av personer på måter som gjør at de utsettes for diskriminering, også vil omfattes av diskrimineringsforbudet. Dette er nok særlig aktuelt dersom personer utsettes for direkte forskjellsbehandling.<sup>55</sup>

At likestillings- og diskrimineringsloven § 6 inneholder et forbud mot sammensatt diskriminering, gjør at vernet mot diskriminering kan anvendes i tilfeller der mer enn ett diskrimineringsgrunnlag er involvert. EU/EØS-retten tar utgangspunkt i ett og ett diskrimineringsgrunnlag og mangler et forbud mot sammensatt diskriminering. Dette er av Xenidis ansett som en svakhet som «casts doubt on the adequateness of EU non-discrimination law when it comes to redressing intersectional manifestations of algorithmic discrimination».<sup>56</sup>

Algoritmisk diskriminering utfordrer vilkåret om hva som skal til for å etablere sammenheng mellom forskjellsbehandlingen og et eller flere vernede diskrimineringsgrunnlag. Algoritmer og KI-systemer gjør spørsmål om «stedfortrederdiskriminering»/«proxy discrimination» mer aktuelt. I tillegg har EU-domstolen gjennom sin praksis utvidet rammene for hva som skal anses å utgjøre direkte forskjellsbehandling. I lys av disse utviklingstrekkene kan det vurderes å innføre egne definisjoner av direkte og indirekte algoritmisk forskjellsbehandling i likestillings- og diskrimineringsloven. Her kan det blant annet presiseres at stedfortrederdiskriminering omfattes av forbudet mot algoritmisk diskriminering.

Likestillings- og diskrimineringslovens forbud mot sammensatt diskriminering kan være et viktig redskap i tilfeller der det er en kombinasjon av ulike diskrimineringsgrunnlag som fører til forskjellsbehandling. Sammensatt diskriminering har et sterkere vern etter norsk rett enn etter EU/EØS-retten.

Forbudet mot diskriminering omfatter ikke bare eksisterende forhold. Det omfatter også tilknytningsdiskriminering, samt diskriminering på grunn av «antatte, tidligere eller fremtidige forhold» ved en person. Dette gjør at personkretsen utvides.

Dette har også en side til det grunnleggende spørsmålet om hvilke grupper som skal være inkludert i listen over diskrimineringsgrunnlag, noe jeg kommer til i kapittel 5, se punkt 5.2.

<sup>54</sup> Se også EU-domstolen, *C-148/22 (OP)* (2023) premiss 25 der det heter: «an internal rule decreed by an employer which prohibits in the workplace only the wearing of conspicuous, large-scale signs of beliefs – philosophical or religious in particular – may constitute direct discrimination on the grounds of religion or belief within the meaning of Article 2(2)(a) of Directive 2000/78 where that criterion is inextricably linked to one or more specific religions or beliefs [...]»

<sup>55</sup> Gerards og Xenidis (2020) s. 68.

<sup>56</sup> Xenidis (2020), se særlig s. 741-745.

#### 4.6 Adgangen til lovlig forskjellsbehandling

Ikke all forskjellsbehandling som har sammenheng med et vernet diskrimineringsgrunnlag er ulovlig. Noen ganger kan forskjellsbehandlingen ha en begrunnelse som gjør at den – etter nærmere vurdering – anses som tillatt. Likestillings- og diskrimineringslovens hovedbestemmelse om tillatt forskjellsbehandling er inntatt i lovens § 9 (1). Her slås det fast at forskjellsbehandling ikke er i strid med forbudet mot diskriminering når den:

- a. har et saklig formål
- b. er nødvendig for å oppnå formålet og
- c. ikke er uforholdsmessig inngripende overfor den eller de som forskjellsbehandles

Dette dreier seg om en bred forholdsmessighetsvurdering. De tre vilkårene for tillatt forskjellsbehandling har en vag utforming. Særlig vilkåret om at forskjellsbehandlingen må være nødvendig for å oppnå formålet og at den ikke må være uforholdsmessig inngripende, åpner opp for en rekke ulike vurderinger. Med utgangspunkt i adgangen til tillatt indirekte forskjellsbehandling etter EU-retten uttaler Gerards og Xenidis:

*If cases of algorithmic discrimination fall ‘by default’ into the indirect discrimination category, leading to an open pool of possible justifications, legal certainty for potential victims, developers and users of algorithmic systems will decrease as the appreciation of the validity of potential justifications would exclusively be bestowed upon courts. In particular, the application of the ‘necessity’ part of the objective justification test by courts poses questions in light of the trade-off between accuracy and performance on the one hand and non-discrimination on the other that might arise in cases of algorithmic discrimination.<sup>57</sup>*

Som påpekt av Gerards og Xenidis, gir vage kriterier for lovlig forskjellsbehandling stor frihet for domstolene i forbindelse med håndheving. Dette kan gå på bekostning av de ulike aktørenes behov for forutsigbarhet. Manglende forutsigbarhet kan i neste omgang utgjøre en hindring for at diskrimineringsforbudet kan få en reell betydning i praksis på KI-feltet.

Ldl. § 9 åpner opp for både tillatt direkte og indirekte forskjellsbehandling, men adgangen til tillatt direkte forskjellsbehandling anses å være snevrere enn adgangen til tillatt indirekte forskjellsbehandling.<sup>58</sup>

Reguleringsmessig kommer dette til uttrykk i ldl. §§ 9 (2) og 10 som etablerer en svært snever adgang til direkte forskjellsbehandling i arbeidsforhold og et absolutt forbud mot graviditetsdiskriminering. Likestillings- og diskrimineringsloven skal tolkes i tråd med EU/EØS-retten. De EU/EØS-rettslige diskrimineringsdirektivene er lagt opp slik at det er hjemlet en generell adgang til lovlig indirekte forskjellsbehandling (forutsatt at den aktuelle problemstillingen omhandler et spørsmål som faller innenfor direktivenes virkeområde). Direkte forskjellsbehandling er imidlertid som utgangspunkt forbudt, med mindre direktivene eksplisitt gir en åpning for lovlig direkte forskjellsbehandling på nærmere definerte vilkår. Gjennom EU/EØS-direktivene er det for eksempel etablert en svært snever adgang til direkte forskjellsbehandling i arbeidslivet, og dette er implementert på nasjonalt plan gjennom ldl. § 9 (2).

Det er altså to faktorer som legger ytre rammer for vurderingen av adgangen til lovlig forskjellsbehandling. Hvilket samfunnsområde forskjellsbehandlingen finner sted på – er den ene. Hvorvidt det dreier seg om direkte eller indirekte forskjellsbehandling – er den andre.

*Hvilke samfunnsområder er av særlig betydning?*

At vernet mot diskriminering er særlig sterkt innen arbeidslivet, har sin bakgrunn i at lik rett til arbeid er svært viktig for enkeltindividers stilling. Det er her lovgivningen om vern mot diskriminering har store deler av sitt opphav, og det er på arbeidslivets område at vernet særlig har utviklet seg.

Med en omfattende teknologisk utvikling har imidlertid algoritmer og KI-systemer blitt tatt i bruk på tvers av mange samfunnsområder. Systemene brukes på arbeidslivets område, for eksempel i rekrutteringssammenheng, men anvendes også på en rekke andre arenaer. Eksempler er diagnostisering og prioriteringer i helsevesenet, prioriteringer av hvem som skal kontrolleres av politi eller tollvesen, eller risikokartlegging som foretas av banker, forsikringsselskaper, utleiefirmaer mv. Et gjennomgående trekk er at algoritmenes og KI-systemenes analyser og vurderinger kan ha stor betydning for enkeltindividers stilling. Det dreier seg om analyser og beslutningsstøtte som kan ha betydning for individers tilgang til essensielle tjenester og rettigheter, slik som helsetjenester, bolig eller muligheten til å tegne forsikring. I lys av teknologiens brede nedslagsfelt på områder av stor betydning for enkeltindivider, bør utgangspunktet om at vernet mot diskriminering er særlig sterkt på arbeidslivets område revurderes. Det bør isteden vurderes å etablere felles kriterier for tillatt forskjellsbehandling

<sup>57</sup> Gerards og Xenidis (2020) s. 73.

<sup>58</sup> Hellum og Strand (2022) s. 255.

som gjelder all algoritme- og KI-utløst ulempe/forskjellsbehandling – uavhengig av hvilket samfunnsområde forskjellsbehandlingen finner sted på.

#### *Om grenseoppgangen mellom direkte og indirekte forskjellsbehandling*

Videre har vi sett at skillet mellom direkte og indirekte forskjellsbehandling har betydning for adgangen til lovlig forskjellsbehandling. Det skal mer til for å anse direkte forskjellsbehandling som lovlig enn indirekte forskjellsbehandling. Heller ikke dette utgangspunktet treffer så godt en sammensatt og kompleks teknologisk virkelighet. Som vi så over i punkt 4.5, trekker KI-systemer ofte inn andre faktorer enn de positivt opplistede diskrimineringsgrunnlagene. Samtidig kan dataene som brukes ha en nær eller nærmest uløselig kobling til de vernede diskrimineringsgrunnlagene. Den allerede krevende grenseoppgangen mellom direkte og indirekte forskjellsbehandling blir ytterligere utfordret i møte med algoritmiske systemer.

Det bør vurderes å innføre en egen unntaksbestemmelse for algoritmisk forskjellsbehandling i likestillings- og diskrimineringsloven. En slik bestemmelse bør ta høyde for kompleksiteten som blant annet oppstår når KI-systemer legger vekt på stedfortrederfaktorer. Samtidig bør en slik særbestemmelse være mer konkret enn hovedregelen i ldl. § 9 (1), og den bør ha som utgangspunkt at et særlig sterkt vern mot direkte algoritmisk diskriminering bør etableres på tvers av alle samfunnsområder, ikke bare i arbeidslivet.

#### *Refleksjoner om utformingen av en slik særlig unntaksbestemmelse*

Når det gjelder vilkåret om saklig formål, kan det slås fast at algoritmisk forskjellsbehandling for å være tillatt må utgjøre et ledd i å realisere det aktuelle datasystemets formål. Det kan også presiseres at systemene ikke kan ha formål som strider mot nasjonal og internasjonal lovgivning (herunder Norges menneskerettslige forpliktelser).

Videre kan det oppstilles et krav om at forskjellsbehandlingen må være egnet og nødvendig for at datasystemets formål skal kunne realiseres. I tilfeller der forskjellsbehandling er en følge av at systemet direkte legger vekt på diskrimineringsvernede gruppekjennetegn – eller på faktorer som uløselig er forbundet med vernede gruppekjennetegn – kan det oppstilles et krav om at forskjellsbehandlingen må ha *avgjørende betydning* for at datasystemets formål skal kunne realiseres. Ved dette heves terskelen for at direkte algoritmisk forskjellsbehandling skal være tillatt – på tvers av alle samfunnsområder. Denne formuleringen er tilpasset

utformingen av dagens ldl. § 9 (2) om tillatt direkte forskjellsbehandling i arbeidslivet.

I relasjon til vilkåret om at forskjellsbehandlingen ikke må være uforholdsmessig inngripende, kan det for eksempel presiseres at det i vurderingen skal legges vekt på om forskjellsbehandlingen gjelder individers tilgang til grunnleggende rettigheter og tjenester, slik som helse, utdanning, arbeid, bank- og forsikringstjenester mv.

Etter mitt syn vil ikke en slik konkretisert unntaksbestemmelse innebære at terskelen for lovlig direkte og indirekte algoritmisk forskjellsbehandling settes høyere i Norge enn det som må antas å gjelde ellers i Europa – på bakgrunn av et meget sammensatt rettskildemateriale om diskrimineringsvern og individers grunnleggende menneskerettigheter, som blant annet er forankret i EU Charteret. En egen unntaksbestemmelse innebærer bare at sentrale elementer i det omkringliggende regelverket blir synliggjort direkte i lovverket.

**Også vilkåret om adgang til lovlig forskjellsbehandling utfordres i situasjoner med direkte og indirekte algoritmisk forskjellsbehandling. Unntaksadgangen er utformet gjennom fleksible kriterier i ldl. § 9, noe som kan føre til utfordringer fordi diskrimineringsforbudets rekkevidde blir lite forutsigbart. Ldl. § 9 bygger på at vernet mot diskriminering er særlig sterkt i arbeidslivet, og grensen mellom direkte og indirekte forskjellsbehandling har betydning for unntaksadgangens rekkevidde.**

Algoritmiske systemer utfordrer de bakenforliggende premisene for dagens unntaksbestemmelse. Dette har sammenheng med at algoritmiske systemer brukes på tvers av alle samfunnsområder. Videre henger det sammen med at den allerede krevende grenseoppgangen mellom direkte og indirekte forskjellsbehandling blir ytterligere utfordret i møte med algoritmiske systemer.

Muligheten for eventuelt å innføre en egen mer konkret unntaksbestemmelse for algoritmisk forskjellsbehandling ble presentert.

#### 4.7 Tidspunkt for diskriminering – særlig om innholdsbasert diskriminering

Likestillings- og diskrimineringsloven krever ikke at diskriminerende virkninger skal ha inntruffet, for at forbudene mot direkte og indirekte diskriminering skal kunne komme til anvendelse.<sup>59</sup> I proposisjonen til likestillings- og diskrimineringsloven heter det i relasjon til forbudet mot direkte diskriminering at:

*lover, vedtak, vedtekter og avtalebetingelser vil kunne rammes av diskrimineringsforbudet til tross for at det ikke har inntruffet en konkret negativ virkning, fordi dette vil være bindende betingelser som etter sin art forutsettes å bli fulgt.*<sup>60</sup>

Det er selve forskjellsbehandlingen som ligger innbakt i det aktuelle regelverket, avtalene mv. som er avgjørende i møte med forbudt mot direkte diskriminering.

Heller ikke ved indirekte diskriminering er det et vilkår at forskjellsbehandling har inntruffet i praksis. Dette fremkommer eksplisitt av ordlyden i ldl. § 8, der indirekte forskjellsbehandling defineres slik at det omfatter «enhver tilsynelatende nøytral bestemmelse, betingelse, praksis, handling eller unnlattelse som vil stille personer dårligere enn andre».<sup>61</sup> Med dette er det slått fast at forbudet mot indirekte diskriminering også omfatter forhold som er egnet til å diskriminere. Forskjellsbehandlende virkninger vil for eksempel kunne ligge innbakt i ulike typer reglementer/regelverk. Den norske tilnærmingen i ldl. § 8 er i tråd med forbudet mot indirekte diskriminering etter EU/EØS-retten.

Det kan stilles spørsmål ved hvilken betydning forbudet mot innholdsbasert diskriminering har i møte med algoritmer og KI-systemer. Mange datasystemer går gjennom en lang testperiode der det anvendes treningsdata og der systemene justeres og utvikles i forkant av at de settes i drift. Det gir lite mening å skulle anvende forbudet mot diskriminering på algoritmer og KI-systemer mens de er i en testperiode, all den tid utviklere i denne perioden nettopp må foreta kartlegginger og justeringer av systemenes mulige negative effekter. Når et system først tas i bruk, vil imidlertid forbudet mot diskriminering gjelde, også før systemet har fått noen konkrete negative effekter.

I praksis vil det nok være en del utfordringer forbundet med å anvende forbudet mot innholdsbasert diskriminering på algoritmiske systemer (herunder KI-systemer), fordi kunnskap om systemenes

funksjonsmåte kan være vanskelig tilgjengelig. Forutsatt at kunnskap finnes om at et system anvender forskjellsbehandlende gruppebaserte kriterier, eller har en innretning som vil stille personer fra diskrimineringsvernede grupper dårligere enn andre, vil imidlertid forbudet mot innholdsbasert direkte og/eller indirekte diskriminering være aktuelt.

En slik situasjon kan for eksempel oppstå dersom enkeltpersoner mener seg utsatt for direkte eller indirekte gruppebasert forskjellsbehandling gjennom hvordan et hel- eller delautomatisert datasystem er anvendt overfor dem, samtidig som disse personene ikke selv ønsker å ta konkrete skritt for å bringe saken inn for Diskrimineringsnemnda eller for ordinære domstoler. I et slikt tilfelle vil kunnskap om systemets mulige diskriminering kunne tilflytte andre individer med samme gruppekjennetegn som de personer som allerede er forskjellsbehandlet, for eksempel personer som tilhører samme tros- eller livssynsretning eller samme etnisitet som personer som allerede er utsatt for forskjellsbehandling. Slik informasjon kan også tilflytte en interesseorganisasjon, se punkt 6.2. Dersom det er grunn til å tro at systemet vil handle på samme måte også overfor andre personer, vil det kunne være aktuelt å bringe en sak inn for Diskrimineringsnemnda eller for ordinære domstoler med påstand om at systemet har en innretning som strider med forbudet mot innholdsbasert diskriminering.

#### 4.8 Den diskrimineringsrettslige bevisbyrderegelen

Den diskrimineringsrettslige bevisbyrderegelen i ldl. § 37 har som formål å bidra til å effektivisere vernet mot diskriminering.<sup>62</sup> Bestemmelsen gjennomfører EU/EØS-retten i norsk rett, samtidig som den gjelder på flere samfunnsområder og for flere grupper enn det som følger av EU/EØS-retten. I norsk sammenheng er det i dag faglig diskusjon om bevisbyrderegelens nærmere innhold. Den norske diskusjonen knyttes til beviskrav og bevisbyrde går jeg ikke nærmere inn på her.

I EU/EØS-rettslig sammenheng har den diskrimineringsrettslige bevisbyrderegelen en tett forbindelse til ivaretagelse av retten til likelønn. Om vi ser bakover i tid, handlet det om at en egen diskrimineringsrettslig bevisbyrderegel skulle bidra til at retten til likelønn

<sup>59</sup> Se Hellum og Strand (2022) s. 242–245.

<sup>60</sup> Prop. 81 L (2016–2017) punkt 12.9.2.2 s. 111.

<sup>61</sup> Min kursivering.

<sup>62</sup> Bestemmelsens første ledd lyder: «Diskriminering skal anses å ha skjedd hvis det foreligger omstendigheter som gir grunn til å tro at diskriminering har skjedd, og den ansvarlige ikke sannsynliggjør at diskriminering likevel ikke har skjedd.»



kunne bli realisert som individuell rettighet. I dag er bevisbyrdebestemmelsen sin betydning i møte med spørsmål om likelønn på nytt aktualisert gjennom vedtakelsen av et nytt EU-direktiv: direktiv 2023/970 om lønnstransparens (EU Equal Pay Transparency Directive). Dette direktivet er EØS-relevant, og det pågår arbeid med å gjennomføre direktivet i norsk rett.<sup>63</sup>

Både algoritmisk diskriminering og lønnsdiskriminering kjennetegnes av situasjoner der enkeltpersoner kan mangle innsikt i nøyaktig hva det er som gjør at de utsettes for forskjellsbehandling, fordi systemene som anvendes er ugjennomsiktige. Dette kan gjøre det vanskelig å underbygge at diskriminering har funnet sted. Etter mitt syn er det verdifull innsikt å hente i innholdet i den diskrimineringsrettslige bevisbyrderegelen på likelønnsområdet.<sup>64</sup>

Den diskrimineringsrettslige bevisbyrderegelen har sin opprinnelse i EU-domstolens håndhevingspraksis i saker om likelønn mellom kvinner og menn. Den ble så i 1997 formalisert i bevisbyrdedirektivet (direktiv 97/80/EF). Deretter ble regelen inntatt i de ulike direktivene på diskrimineringsfeltet.<sup>65</sup> EU-domstolens utvikling av en egen bevisbyrderregel skjedde gjennom fortolkning av likelønnsbestemmelsen i EU-traktaten artikkel 119 (i dag TEUV artikkel 157) og har også tett sammenheng med utviklingen av et vern mot indirekte kjønnsdiskriminering. Tre sentrale saker var *Danfoss-saken*, *Enderby-saken* og *Royal Copenhagen*.<sup>66</sup> Felles for disse sakene var at de omhandlet lønnsfastsettingssystemer som på overflaten var nøytralt utformet, men som hadde virkninger som gjorde at kvinner kom dårligere ut lønnsmessig enn menn. Det dreide seg dermed om bruk av diskrimineringsforbudet i møte med ugjennomsiktige systemer som resulterte i kjønnsrelatert forskjellsbehandling. For å ivareta enkeltindividets rettsstilling, la EU-domstolen i disse sakene vekt på hensynet til effektiv gjennomføring av diskrimineringsforbudet. Som en følge av dette valgte domstolen å la tvil om sakens faktum gå ut over arbeidsgiver.

<sup>63</sup> Regjeringen, EØS-notatbasen, «Likelønnsdirektivet» (2023). Gjennomføringsfristen oppgis å være 7. juni 2026.

<sup>64</sup> Dette er også trukket frem i Hauglid (2023) s. 328.

<sup>65</sup> Se rasediskrimineringsdirektivet (2000/43/EF) artikkel 8; rammedirektivet (2000/78/EF) artikkel 10; kjønnslikestillingsdirektivet (2006/54/EF) artikkel 19 nr. 1; kjønnslikestillingsdirektivet om varer og tjenester (2004/113/EF) artikkel 9.

<sup>66</sup> C-109/88 (*Danfoss*) (1989); C-127/92 (*Enderby*) (1993); C-400/93 (*Royal Copenhagen*) (1995).

*Danfoss-saken* omhandlet spørsmål om diskriminering i forbindelse med lønn i en virksomhet som opererte med et ugjennomsiktig lønnsfastsettingssystem. Saksøker hadde bare tilgang til gjennomsnittsinformasjon som synliggjorde at det forelå kjønnsbaserte forskjeller med hensyn til lønn. EU-domstolen beskrev situasjonen slik:

*the system of individual supplements applied to basic pay is implemented in such a way that a woman is unable to identify the reasons for a difference between her pay and that of a man doing the same work. Employees do not know what criteria in the matter of supplements are applied to them and how they are applied. They know only the amount of their supplemented pay without being able to determine the effect of the individual criteria. Those who are in a particular wage group are thus unable to compare the various components of their pay with those of the pay of their colleagues who are in the same wage group.*<sup>67</sup>

Det forelå ingen informasjon om begrunnelsen for lønnsfastsettelsen. EU-domstolen uttalte:

*where an undertaking applies a system of pay which is totally lacking in transparency, it is for the employer to prove that his practice in the matter of wages is not discriminatory, if a female worker establishes, in relation to a relatively large number of employees, that the average pay for women is less than that for men.*<sup>68</sup>

Danfoss-saken er eksplisitt omtalt i fortalen punkt 52 i lønnstransparensdirektivet (direktiv 2023/970):

*In accordance with the case-law of the Court of Justice, Directive 2006/54/EC establishes provisions to ensure that the burden of proof shifts to the respondent when there is a prima facie case of discrimination. Nevertheless, it is not always easy for victims and courts to know how to establish even that presumption. In case C-109/88, the Court of Justice held that when a system of pay is totally lacking in transparency, the burden of proof should be shifted to the respondent, irrespective of the worker showing a prima facie case of pay discrimination. Accordingly, the burden of proof should be shifted to the respondent where an employer does not comply with the pay transparency obligations set out in this Directive, for instance by refusing to provide information requested by the workers or not reporting*

<sup>67</sup> C-109/88 (*Danfoss*), premiss 10.

<sup>68</sup> C-109/88 (*Danfoss*), punkt 1 i domstolens konklusjon.

*on the gender pay gap, where relevant, save where the employer proves that such an infringement was manifestly unintentional and of a minor character.*

Her pekes det på to sentrale temaer. Det første handler om bevisbyrderegelens utgangspunkt om at det normalt påhviler klager/saksøker å etablere en presumsjon for diskriminering («*a prima facie case of discrimination*»).<sup>69</sup> Det andre handler om lønnsfastsettingssystemer som er «totally lacking in transparency», herunder situasjoner der en arbeidsgiver ikke er transparent med hensyn til lønnsfastsettelse på den måten som lønnstransparensdirektivet krever.<sup>70</sup> I slike tilfeller påhviler det ikke lenger klager/saksøker å etablere en presumsjon for diskriminering. Ansvar for å vise at det ikke foreligger lønnsdiskriminering overføres da med en gang til arbeidsgiver/den ansvarlige.

Disse to bevismessige situasjonene er eksplisitt omtalt i lønnstransparensdirektivet artikkel 18 nr. 1 og 2, som lyder:

- 1. Member States shall take the appropriate measures, in accordance with their national judicial systems, to ensure that, when workers who consider themselves wronged because the principle of equal pay has not been applied to them establish before a competent authority or national court facts from which it may be presumed that there has been direct or indirect discrimination, it shall be for the respondent to prove that there has been no direct or indirect discrimination in relation to pay.*
- 2. Member States shall ensure that, in administrative procedures or court proceedings regarding alleged direct or indirect discrimination in relation to pay, where an employer has not implemented the pay transparency obligations set out in Articles 5, 6, 7, 9 and 10, it is for the employer to prove that there has been no such discrimination.*

*The first subparagraph of this paragraph shall not apply where the employer proves that the infringement of the obligations set out in Articles 5, 6, 7, 9 and 10 was manifestly unintentional and of a minor character.*  
[...]

---

<sup>69</sup> I ldl. § 37 omtales dette som klagers ansvar for å vise at det er «grunn til å tro at diskriminering har skjedd».

<sup>70</sup> Direktivet artikkel 18 nr. 2 viser videre til direktivet artikkel 5 «Pay transparency prior to employment», artikkel 6 «Transparency of pay setting and pay progression policy», artikkel 7 «Right to information», artikkel 9 «Reporting on pay gap between female and male workers» og artikkel 10 «Joint pay assessment».

Punkt 52 i fortalen og direktivets artikkel 18 synliggjør måter å håndtere saker om diskriminering på – der forskjellsbehandling finner sted innen rammen av et ugjennomtrengelig system og der arbeidsgiver er pålagt ulike strukturelle forpliktelser for å sikre transparens om lønn. Dette kan ha overføringsverdi til spørsmål om algoritmisk diskriminering.

Lønnstransparensdirektivet synliggjør to muligheter med hensyn til bevisbyrde. Den ene muligheten er å legge til grunn at bevisbyrden overføres til den ansvarlige dersom det foreligger tilstrekkelig med dokumentasjon om at et datasystem resulterer i forskjellsbehandling.

Den andre muligheten er å gå ett skritt videre ved å trekke tilnærmingen i artikkel 18 nr. 2 over til algoritme- og KI-området. Etter mitt syn bidrar lønnstransparensdirektivet artikkel 18 nr. 2 til å videreutvikle og konkretisere Danfoss-tilnærmingen. Det handler om en slags prosessuell dreining og ansvarliggjøring, som i neste omgang spiller inn på spørsmålet om bevisbyrde. I tilfeller der krav til forhåndsvurderinger, sikkerhetsmekanismer, transparens og rapportering ikke er fulgt, er utgangspunktet etter lønnstransparensdirektivet at det foreligger lønnsdiskriminering, med mindre den ansvarlige kan bevise at diskriminering ikke har funnet sted.

Etter mitt syn kan det være grunnlag for å regel-feste en slik forståelse også i møte med KI-relatert diskriminering. I så fall kan det være aktuelt å trekke inn ansvar og plikter som brukere av algoritmer og KI-systemer er pålagt, både etter likestillings- og diskrimineringslovgivningen, men også etter andre regelsett – slik som etter KI-forordningen og GDPR. Det bør utredes nærmere hvilke strukturelle forpliktelser som kan være aktuelle. Hva som skal til for å motbevise en presumsjon om algoritmisk diskriminering, er også et interessant spørsmål. Systemer for KI-revisjon («AI auditing») og utviklingen av metoder som kan forklare hva som skjer inni KI-systemer («explainable AI methods») kan være relevant i denne sammenheng.<sup>71</sup>

---

<sup>71</sup> Strümke, Slavkovik og Stachl (2023).

#### 4.9 Sammenfatninger og refleksjoner om veien videre

*Viktige dimensjoner som er tatt opp i kapittel 4:*

- Algoritmiske systemer er komplekse, noe som gjør at forbudet mot indirekte diskriminering kan være enklere å anvende enn forbudet mot direkte diskriminering.
- Algoritme- og KI-relatert ulempe utfordrer vilkåret om forskjellsbehandling. Det handler dels om at forskjellsbehandlingsvilkåret har et noe vagt innhold. Dels handler det om bevismessig utfordringer knyttet til å få innsikt og kunnskap om algoritmers og KI-systemers kriterier og virkninger, slik at det blir mulig å underbygge at forskjellsbehandling har funnet sted.
- KI-systemers bruk av stedfortrederfaktorer gjør det mer krevende å etablere en sammenheng mellom forskjellsbehandling og et eller flere diskrimineringsgrunnlag. Dette gjør at forbudet mot indirekte diskriminering, som retter seg mot systemers virkninger, kan være mer aktuelt enn forbudet mot direkte diskriminering i saker som omhandler diskriminering gjennom stedfortrederfaktorer («proxy discrimination»).

Samtidig viser praksis fra EU-domstolen at forbudet mot direkte diskriminering ikke bare gjelder der forskjellsbehandling eksplisitt er begrunnet i et vernet diskrimineringsgrunnlag. Forbudet kommer også til anvendelse dersom det er anvendt kriterier som uløselig er forbundet med et vernet gruppekjennetegn.

I lys av disse trekkene, kan det vurderes å innføre egne definisjoner av direkte og indirekte algoritmisk forskjellsbehandling i likestillings- og diskrimineringsloven.

- Forbudet mot diskriminering omfatter ikke bare eksisterende forhold. Det omfatter også tilknytningsdiskriminering, samt diskriminering på grunn av «antatte, tidligere eller fremtidige forhold» ved en person. Disse sidene ved diskrimineringsforbudet kan bli aktuelle i spørsmål om algoritmisk diskriminering. De kan bidra til at personkretsen som er omfattet av vernet mot diskriminering blir utvidet.
- Sammensatt diskriminering er omfattet av forbudet mot diskriminering, noe som kan gjøre det enklere å anvende forbudet mot diskriminering i tilfeller der mer enn ett diskrimineringsgrunnlag er involvert. De EU/EØS-rettslige diskrimineringsdirektivene inneholder ikke et forbud mot sammensatt diskriminering.

- Adgangen til tillatt forskjellsbehandling i ldl. § 9 er som utgangspunkt vagt utformet, noe som kan være en utfordring i møte med algoritmisk forskjellsbehandling fordi det gir liten grad av forutsigbarhet. To faktorer påvirker unntaksadgangens rekkevidde etter dagens bestemmelse: hvilket samfunnsområde forskjellsbehandlingen finner sted på, og om det dreier seg om direkte eller indirekte forskjellsbehandling. Utredningen viser at algoritmisk forskjellsbehandling utfordrer begge disse grunnleggende utgangspunktene. Algoritmiske systemer brukes på tvers av alle samfunnsområder og bidrar på ulike måter til avgjørelser som kan ha svært stor betydning for individer. Dette gjør at utgangspunktet om at arbeidslivsfeltet er særlig viktig, bør nyanseres. Videre gjør kompleksiteten i algoritmiske systemer at det kan være vanskelig å trekke en klar grense mellom direkte og indirekte diskriminering. I lys av disse forholdene presenteres muligheten for å innføre en egen og mer konkretisert unntaksbestemmelse for algoritmisk forskjellsbehandling i likestillings- og diskrimineringsloven.
- Kapitlet viser at forbudet mot innholdsbasert diskriminering kan være aktuelt i saker om algoritmisk diskriminering.
- Kapitlet viser at den diskrimineringsrettslige bevisbyrderregelen, særlig slik regelen er presisert i det nye lønntransparensdirektivet (direktiv 2023/970) artikkel 18, kan utgjøre en nøkkel i møte med algoritmisk diskriminering. Dette har sammenheng med at saker om likelønn kjennetegnes av manglende transparens om hva som ligger til grunn for lønnsfastsettelse. Her er det dermed en tydelig parallell til spørsmål om algoritmisk diskriminering. Lønntransparensdirektivet artikkel 18. nr. 2 gir uttrykk for en prosessuell dreining idet bestemmelsen pålegger den som er ansvarlig for lønn å gjennomføre ulike vurderinger og å være åpen om resultatene. Dersom dette ikke gjøres, er det denne parten som har ansvaret for å bevise at diskriminering ikke har funnet sted. I denne tilnærmingen ligger det elementer som muligens kan overføres til spørsmål om håndheving av enkelt saker der tema er algoritmisk diskriminering.

### *Refleksjoner om veien videre:*

Kapittelet har synliggjort at det kan vurderes å innføre noen nye presiseringer i likestillings- og diskrimineringsloven. Presiseringene vil kunne bidra til å sikre et mer robust forbud mot diskriminering. Følgende temaer er tatt opp:

- Det kan vurderes å innføre egne definisjoner av direkte og indirekte algoritmisk forskjellsbehandling. Definisjonen av direkte algoritmisk forskjellsbehandling bør få frem at forbudet både omfatter situasjoner der et algoritmisk system direkte legger vekt på et/flere vernet gruppekjennetegn (se ldl. § 7) og situasjoner der systemet legger vekt på faktorer som uløselig er knyttet til et vernet gruppekjennetegn. Definisjonen av indirekte algoritmisk forskjellsbehandling kan bygge på dagens ordlyd i ldl. § 8, men også presisere at forbudet omfatter bruk av stedfortrederfaktorer – i tilfeller der vernet ikke faller inn under (det foreslåtte) forbudet mot direkte algoritmisk forskjellsbehandling.
- Det kan vurderes å innføre en egen bestemmelse om adgang til lovlig algoritmisk forskjellsbehandling. Bestemmelsen bør være mer konkret enn ldl. § 9 (1), den bør ta høyde for kompleksiteten som foreligger i grenseoppgangen mellom direkte og indirekte algoritmisk forskjellsbehandling og den bør ha som utgangspunkt at vernet mot algoritmisk diskriminering bør være like sterkt på tvers av alle samfunnsområder. Slike mer konkrete vilkår kan etter mitt syn anses utledet av felleseuropeiske regler om diskrimineringsvern og menneskerettigheter.

En slik unntaksbestemmelse kan slå fast at et algoritmisk system som forskjellsbehandler, for å være tillatt, må:

- ▶ søke å realisere et saklig formål, noe som forutsetter at systemet ikke kan ha formål som strider mot nasjonal og internasjonal lovgivning (herunder Norges menneskerettslige forpliktelser);
  - ▶ at forskjellsbehandlingen må være egnet og nødvendig for at det algoritmiske systemets formål skal kunne realiseres. Her kan det også presiseres at dersom det dreier seg om systemer som direkte legger vekt på vernede gruppekjennetegn eller på faktorer som uløselig er knyttet til gruppekjennetegnene, må forskjellsbehandlingen ha avgjørende betydning for realiseringen av systemets formål. (En slik presisering vil innebære at det etableres et særlig sterkt vern mot direkte algoritmisk diskriminering.)
  - ▶ til vilkåret om at forskjellsbehandlingen ikke må være uforholdsmessig inngripende kan det presiseres at det skal legges vekt på om forskjellsbehandlingen omhandler individers tilgang til grunnleggende rettigheter og tjenester, slik som helse, utdanning, arbeid, bank- og forsikringstjenester mv.
- Det kan vurderes å innføre en egen bevisbyrdebestemmelse om algoritmisk diskriminering, med inspirasjon fra lønnstransparensdirektivet artikkel 18.

# KAPITTEL 5

## Aktørene: hvem er vernet mot diskriminering, og hvem kan holdes ansvarlig?

### 5.1 Innledning

I dette kapitlet rettes oppmerksomheten mot diskrimineringsvernets aktører. Hvem er omfattet av vernet mot diskriminering, og hvem kan holdes ansvarlig for å ha diskriminert? Jeg belyser også om forbudene mot instruksjon og medvirkning kan utvide kretsen av hvilke aktører som kan holdes ansvarlig.

### 5.2 Diskrimineringsvernede grupper

Ldl. § 6 (1) forbyr diskriminering som har sammenheng med «kjønn, graviditet, permisjon ved fødsel eller adopsjon, omsorgsoppgaver, etnisitet, religion, livs-syn, funksjonsnedsettelse, seksuell orientering, kjønns-identitet, kjønnsuttrykk, alder eller kombinasjoner av disse grunnlagene.» Bestemmelsen presiserer dessuten at grunnlaget etnisitet blant annet omfatter «nasjonal opprinnelse, avstamning, hudfarge og språk.»<sup>72</sup>

Det er imidlertid også en rekke gruppebaserte kjennetegn som ikke er omfattet av lovgivningens opplisting av diskrimineringsgrunnlag. Det er dermed fare for at algoritmer- og KI-systemer kan bidra til ulemper – og til forsterket ulempe – for grupper som i dag ikke er omfattet av likestillings- og diskrimineringsloven. Dette kan for eksempel skje dersom KI-systemene bidrar til å knytte ulemper opp mot faktorer som sosial

status, helse, bostedsadresse, utdanningsnivå og inntekt.<sup>73</sup> Det er også påvist at KI-systemer kan forskjellsbehandle på grunn av gruppekaraktistikker som man i utgangspunktet tenker at er irrelevante, men som likevel viser seg å føre til ulempe. Et eksempel på dette er dersom KI-systemer legger til grunn at det å være hundeeier gjør at man er mer mottakelig for visse former for reklame.<sup>74</sup>

Algoritmiske systemers kompleksitet kan gjøre det aktuelt å diskutere muligheten for å åpne diskrimineringsvernet opp for nye grupper. Dette reiser imidlertid grunnleggende spørsmål om forholdet mellom gruppekjennetegn som er begrunnet i historisk strukturell ulikhet, slik som kjønnsdiskriminering, og ulemper som oppstår for ulike grupperinger – i ulike kontekster – i dagens samfunn.

At det kan oppstå utfordringer i møtet mellom algoritmisk diskriminering og et regelverk som i utgangspunktet bygger på en uttømmende tilnærming til hvilke grupper som er vernet mot diskriminering, er allerede satt på agendaen i europeisk sammenheng.<sup>75</sup> I akademiske artikler og utredninger tas det blant annet opp om EU Charteret artikkel 21 (som er en ikke-uttømmende bestemmelse om diskrimineringsvern) kan brukes som mekanisme for å fange opp situasjoner med diskriminering mot grupper som i dag ikke er vernet mot diskriminering gjennom diskrimineringsdirektivene.

72 På arbeidslivets område er det i tillegg etablert forbud mot diskriminering på grunn av politisk syn, medlemskap i arbeidstakerorganisasjon, deltidsarbeid og midlertidig ansettelse, se lov 17. juni 2005 nr. 62 om arbeidsmiljø, arbeidstid og stillingsvern mv. (arbeidsmiljøloven) § 13-1.

73 Gerards og Xenidis (2020) s. 64.

74 Wachter (2022).

75 Gerards og Xenidis (2020).

Begrensningene som følger av en additiv tilnærming til hvilke grupper som er vernet mot diskriminering, kan i noen grad løses gjennom å inkludere en åpen sekkekategori i angivelsen av diskrimineringsgrunnlag. Som vi har sett, bygger ldl. § 6 på en uttømmende opplisting av diskrimineringsgrunnlag. Samtidig har formålsbestemmelsen i ldl. § 1 (1) en ikke-uttømmende tilnærming til hvilke grupper som er omfattet av diskrimineringsvernet, jf. formuleringen «og andre vesentlige forhold ved en person». Formålsbestemmelsen har klare linjer til de ikke-uttømmende diskrimineringsforbudene i internasjonale menneskerettskonvensjoner, for eksempel EMK artikkel 14, SP artikkel 26 og ØSK artikkel 2 nr. 2.

Det er en utfordring at lovens formålsbestemmelse er ikke-uttømmende i sin angivelse av lovens diskrimineringsgrunnlag, mens diskrimineringsforbudet i § 6 (1) bygger på en uttømmende opplisting. Dette skaper uklarhet som det er behov for at lovgiver rydder opp i. Spørsmål om algoritmisk diskriminering kan gi foranledning til å se nærmere på dette.

Det kan også vurderes å innføre en egen diskrimineringsforbudsbestemmelse om algoritmisk diskriminering – som bygger på en positiv opplisting av venede diskrimineringsgrunnlag, men som også ivaretar behovet for fleksibilitet gjennom en åpen sekkekategori.

### 5.3 Hvem skal påstanden om diskriminering rettes mot?

Algoritmiske systemer (herunder KI-systemer) utvikles gjennom flere faser. Det er først i implementeringsfasen at systemene mates med nye data og kan generere svar («output») som skal ivareta formålene som systemene er konstruert for å oppfylle.<sup>76</sup> Det er også i denne fasen at individer eventuelt kan utsettes for diskriminering.

I og med at ulike aktører er involvert før et algoritmisk system tas i bruk, kan det stilles spørsmål ved hvem som egentlig kan og skal holdes ansvarlig for å ha diskriminert.

Det klare utgangspunktet må være at en person som mener seg diskriminert skal rette en påstand om dette mot den private eller offentlige aktøren som er avsender av en avgjørelse, en beslutning, et vedtak mv. – og som bruker et algoritmisk system som ledd i sin virksomhet. I KI-forordningen omtales dette som «the deployer», jf. artikkel 3 nr. 4. Dette utgangspunktet må gjelde uavhengig av om det dreier seg om en

helautomatisk algoritmisk avgjørelse eller en situasjon der et KI-system brukes som kunnskapsstøtte for en avgjørelse som til syvende og sist tas av et menneske. En offentlig eller privat arbeidsgiver, utdanningsinstitusjon, helseinstitusjon mv. kan ikke navigere seg bort fra sine forpliktelser etter likestillings- og diskrimineringsloven gjennom å bruke systemer som baserer seg på algoritmer og KI.<sup>77</sup> Denne forståelsen er også i tråd med plikten til å sikre effektiv gjennomføring av diskrimineringsvernet.

EMDs behandling av saken *Glukhin mot Russland* bygger på tilsvarende forståelse.<sup>78</sup> Saken omhandlet blant annet konvensjonsmessigheten av et automatisk ansiktsgjenkjenningssystem som var i bruk på t-banen i Moskva. Det var russiske myndigheters bruk av systemet overfor en enkeltperson som var det sentrale vurderingstemaet i saken, ikke eventuelle aktører som hadde bidratt til å utvikle datasystemet. EMD konkluderte med at bruken utgjorde en krenkelse av individets rett til privatliv etter EMK artikkel 8.

At en påstand om algoritmisk diskriminering skal rettes mot aktøren som bruker slike systemer som del av sin virksomhet, synliggjør et det påhviler disse aktørene et ansvar både ved anskaffelse og ved oppfølging av systemene. Gitt KI-systemers kompleksitet, er det for disse aktører et stort behov for tiltak som kan forhindre at diskriminering oppstår i enkelttilfeller. Dette handler både om systemer for å vurdere diskrimineringsrisiko før systemet tas i bruk, samt systemer for løpende kontroll etter at systemet er i bruk. Slike strukturelle mekanismer vil kunne redusere risikoen for diskriminering, men et individvern vil fortsatt eksistere gjennom likestillings- og diskrimineringslovgivningen.

I situasjoner der en aktør som bruker et algoritmisk system blir funnet å ha utsatt enkeltpersoner for diskriminering og må betale oppreisning og/eller erstatning til disse, kan det oppstå spørsmål om mulighet for å kreve regress fra aktørene som har vært involvert i å systemutviklingen. Denne problemstillingen går jeg ikke nærmere inn på her.

<sup>77</sup> Dette har en liten parallell til menneskerettsfeltet, der det er klart etablert at en stat ikke kan unndra seg menneskerettslig forpliktelser ved å overlate bestemte oppgaver til private aktører.

<sup>78</sup> EMD, *Glukhin mot Russland* (sak nr.11519/20) (2023).

#### 5.4 Kan forbudene mot instruksjon og medvirkning utvide kretsen av hvilke aktører som kan holdes ansvarlig?

I ldl. § 15 slås det fast at det er forbudt å instruere noen om å diskriminere. Instruksjonsforbudet gjennomfører de EU/EØS-rettslige diskrimineringsdirektivene i norsk rett, men har et mer vidtrekkende virkeområde enn etter direktivene idet de gjelder generelt innenfor hele likestillings- og diskrimineringslovens virkeområde. Forbudet skal ramme «ordre, retningslinjer, formaninger, oppfordringer etc. overfor en eller flere personer om å diskriminere».<sup>79</sup> Hensikten er at ikke bare den som direkte eller indirekte diskriminerer noen kan holdes ansvarlig, men også den som initierer eller beordrer diskriminering.

Likestillings- og diskrimineringsloven oppstiller videre et forbud mot å medvirke til diskriminering. Forbudet kommer til uttrykk i ldl. § 16 og lyder:

*Det er forbudt å medvirke til å diskriminere i strid med § 6, trakassere i strid med § 13, gjengjelde i strid med § 14, og instruere i strid med § 15.*

I forbindelse med at forbudet mot medvirkning i 2005 ble innført i likestillingsloven 1978, uttalte departementet at bestemmelsen skulle «sikre at handlinger eller unnlatelser som ellers vil kunne falle utenfor» lovens diskrimineringsforbud ble omfattet.<sup>80</sup> Departementet fremhevet videre at

*rammene for når det foreligger medvirkning ikke kan fastlegges generelt, men vil måtte bero på en konkret tolkning sett i forhold til ulovlig forskjellsbehandling, instruks, gjengjeldelse og trakassering. Både fysisk og psykisk medvirkning vil kunne rammes av forbudet.<sup>81</sup>*

De EU/EØS-rettslige diskrimineringsdirektivene inneholder ikke forbud mot medvirkning. Medvirkningsforbudet i likestillings- og diskrimineringsloven innebærer dermed at norsk lovgivning etablerer et strengere vern mot diskriminering enn minimumskravene som følger av direktivene.<sup>82</sup>

Et aktuelt problemkompleks handler om hvilken potensiell betydning forbudene mot instruksjon og mot medvirkning har i møte med algoritmisk diskriminering. Kan for eksempel aktører som har vært med på å utvikle et diskriminerende KI-system anses for å

ha medvirket til diskrimineringen? Eller kan et kunnskapsstøttende KI-system som gir anbefalinger og prediksjoner om hvilken avgjørelse som bør tas, omfattes av instruksjonsforbudet i tilfeller der resultatet er at en person utsettes for diskriminering?

Gerards og Xenidis påpeker at forbudet mot instruks ikke er klart definert verken i EU-direktivene eller gjennom EU-domstolens tolkning. De mener imidlertid at «an innovative interpretation of the notion of ‘instruction to discriminate’ could effectively mitigate the substantive and procedural hurdles that arise in the context of algorithmic discrimination.»<sup>83</sup> De utelukker ikke at instruksjonsforbudet rettslig sett kan tolkes slik at det rammer situasjoner der kunnskapsstøttende KI-systemer er en del av årsaken til at diskriminerende avgjørelser tas. Videre påpeker de at instruksjonsforbudet kan være positivt ved at det kan utgjøre et tillegg/supplement til forbudene mot direkte og indirekte diskriminering (og den vanskelige grenseoppgangen mellom disse). Til sist peker de på at instruksjonsforbudet kan fungere som en nøkkel til å ansvarliggjøre ulike aktører i verdikjeden, for eksempel ved at de som designer algoritmer som brukes i KI-systemer vil søke å få sine produkter sertifisert slik at de unngår anklager om å ha instruert noen om å diskriminere.<sup>84</sup>

Etter mitt syn vil medvirkningsforbudet også kunne være aktuelt i saker om algoritmisk diskriminering. Medvirkningsbegrepet er vagt, og det er derfor noe uklart hva som skal til for å bli ansett for å ha medvirket til diskriminering. Det er imidlertid potensiale for at forbudet mot medvirkning innehar en fleksibilitet som gjør det egnet til å fange opp diskriminering i komplekse situasjoner. Sammen med forbudet mot direkte og indirekte diskriminering, gjør forbudet mot medvirkning det mulig å holde mer enn én aktør ansvarlig for diskriminering.

En sak behandlet av Diskrimineringsnemnda synliggjør muligheten til å holde flere aktører ansvarlig for det samme bruddet på likestillings- og diskrimineringsloven. Saken omhandlet en gravid kvinne (A) som var ansatt i 50 % stilling hos bemanningsbyrået C.<sup>85</sup> Hun var innleid på et midlertidig oppdrag hos B. Etter at det ble kjent at kvinnen var gravid, ble oppdraget ikke forlenget av B. Dette var begrunnet i at kvinnen som følge av kommende fødselspermisjon ikke kunne være vikar i alle sommermånedene.

<sup>79</sup> Ot.prp. nr. 35 (2004–2005) punkt 5.2.

<sup>80</sup> Ot.prp. nr. 35 (2004–2005) punkt 6.5.

<sup>81</sup> Ot.prp. nr. 35 (2004–2005) punkt 6.5.

<sup>82</sup> Ot.prp. nr. 35 (2004–2005) punkt 6.1.

<sup>83</sup> Gerards og Xenidis (2020) s. 143.

<sup>84</sup> Gerards og Xenidis (2020) s. 143–144.

<sup>85</sup> Diskrimineringsnemnda, *DIN-2019-115* og *DIN-2019-196* (felles behandling).

Bemanningsbyrået C tok ikke grep for å bidra til at kvinnen likevel fikk forlenget oppdraget sitt hos B. Nemnda uttalte følgende om dette trepartsforholdet:

*Nemnda tar som utgangspunkt at et slikt trepartsforhold ikke innebærer at utleierens eller innleierens plikt til opptre i samsvar med diskrimineringslovgivningen, bortfaller eller reduseres. Utleieren kan altså ikke uten videre vise til at han ikke har ansvar for innleierens handlemåte. Innleieren kan heller ikke uten videre vise til at han bare har et kontraktsforhold med utleieren og henvide den ansatte til utleieren som arbeidsgiver. Diskrimineringsvernet står på egne ben, og det må vurderes konkret om innleieren eller utleieren – eller begge – har handlet på en måte som innebærer at den aktuelle ansatte rent faktisk er kommet i en dårligere situasjon enn ansatte som ikke er omfattet av det aktuelle diskrimineringsgrunnlaget.*

Nemnda konkluderte med at innleiers beslutning om ikke å forlenge oppdragsavtalen med kvinnen utgjorde diskriminering på grunn av graviditet og permisjon i strid med ldl § 6. I vurderingen av om bemanningsbyrået hadde medvirket til diskrimineringen uttalte nemnda at «både bemanningsbyrået og innleieren må ivareta sitt ansvar for å hindre at arbeidstakeren blir satt i en dårligere stilling på grunn av graviditet og permisjon ved fødsel.» Bemanningsbyråets passivitet og unnlattelse av å handle overfor innleier innebar etter nemndas syn at byrået hadde medvirket til diskrimineringen i strid med medvirkningsforbudet i § 16 jf. § 6. Kvinnen ble tilkjent oppreisning fra både bemanningsbyrået og innleier.<sup>86</sup>

I den nevnte saken opptrådte både bemanningsbyrået og innleier på en klanderverdig måte. Diskriminerings- og medvirkningsforbudet forutsetter imidlertid ikke at det foreligger en intensjon om å diskriminere, noe som stort sett vil være situasjonen i møte med algoritmisk diskriminering. Eksempelet fra nemnda synliggjør at medvirkningsforbudet vil kunne være aktuelt for aktører som har handlet på en måte som muliggjør diskriminering, der selve diskrimineringen skjer i en annen relasjon.

Etter mitt syn kan medvirkningsforbudet ha potensiale til å kunne brukes i situasjoner med algoritmisk diskriminering og det vil – i likhet med instruksjonsforbudet – også kunne ha en disiplinerende effekt på aktører som arbeider med å utvikle og teste algoritmiske systemer. Forbudet mot instruksjon og forbudet

mot medvirkning overlapper nok i en del tilfeller, men muligens kan medvirkningsforbudet være egnet til å fange opp et større mangfold av aktører enn det som er aktuelt for instruksjonsforbudet. Det dominerende er imidlertid likhetstrekkene som foreligger mellom instruksjonsforbudet og medvirkningsforbudet, idet begge forbud er egnet til å slå ned på andre aktørers bidrag til at en bruker av et algoritmisk system ender opp med å diskriminere.

## 5.5 Sammenfatninger og refleksjoner om veien videre

Viktige dimensjoner som er tatt opp i kapittel 5:

- Likestillings- og diskrimineringsloven angir uttømmende hvilke diskrimineringsgrunnlag som er omfattet av forbudet mot diskriminering, og bygger med dette på samme tilnærming som EU/EØS-rettens diskrimineringsdirektiver. Det er i økende grad rettet oppmerksomhet mot at algoritmisk diskriminering også kan ramme en rekke andre grupper enn de som er opplistet. For eksempel faller faktorer som sosial status, helse og utdanningsnivå utenfor opplistingen av diskrimineringsgrunnlag. Algoritmisk diskriminering synliggjør at det kan være behov for å diskutere om det bør innføres en fleksibilitet i lovens angivelse av diskrimineringsgrunnlag, slik at også andre grupper kan omfattes av forbudet mot diskriminering.

Likestillings- og diskrimineringslovens formålsbestemmelse inkluderer formuleringen «og andre vesentlige forhold ved en person». Det er en spenning mellom den uttømmende opplistingen av diskrimineringsgrunnlag i diskrimineringsforbudet i ldl. § 6 og den ikke-uttømmende tilnærmingen i formålsbestemmelsen i § 1. Algoritmisk diskriminering kan gi en foranledning til å se nærmere på dette.

- En påstand om diskriminering fra et individ skal rettes mot den private eller offentlige aktøren som har fattet en beslutning, utformet et vedtak mv. overfor individet, og som – som ledd i sin virksomhet – har brukt et algoritmisk system.
- Likestillings- og diskrimineringslovens forbud mot instruks og mot medvirkning kan bidra til at forbudet mot diskriminering kommer til anvendelse på flere aktører, slik at for eksempel også utviklere i noen tilfeller kan omfattes av forbudet mot diskriminering. Innholdet i de to forbudene er lite utviklet. Dette er et område der likestillings- og diskrimineringsloven har et sterkere vern enn det som følger av EU/EØS-rettens diskrimineringsdirektiver.

<sup>86</sup> Dette beløp seg til kr 25 000 fra bemanningsbyrået og kr 15 000 fra innleier.



Lovens forbud mot instruks gjelder innenfor hele lovens virkeområde («på alle samfunnsområder», jf. ldl. § 2 (1)), som er videre enn direktivenes virkeområde. EU/EØS-direktivene inneholder ikke noe forbud mot medvirkning.

*Refleksjoner om veien videre:*

Kapittelet har synliggjort behov for å vurdere veien videre for noen av likestillings- og diskrimineringslovens bestemmelser.

- Det er behov for å vurdere innføring av en åpen sekkekategori i diskrimineringsforbudets angivelse av diskrimineringsgrunnlag. Dersom det ikke er ønskelig å innføre dette generelt, kan det vurderes å innføre en egen diskrimineringsforbudsbestemmelse om algoritmisk diskriminering – som bygger på en positiv opplisting av vernede diskrimineringsgrunnlag, men som også ivaretar behovet for fleksibilitet gjennom en åpen sekkekategori.
- Det er behov for belysning av hvilken betydning forbudene mot instruks og mot medvirkning kan ha i møte med algoritmisk diskriminering.

# KAPITTEL 6

## Håndheving og sanksjoner

### 2.1 Innledning

Enkelt personer som mener seg utsatt for algoritmisk diskriminering, har mulighet til enten å klage en sak inn for Diskrimineringsnemnda eller å bringe en sak inn for ordinære domstoler.<sup>87</sup>

Videre i dette punktet er det to hovedtemaer som er særlig aktuelle i møte med algoritmisk diskriminering. Det første handler om organisasjoners adgang til å bringe en sak om slik diskriminering inn for nemnda eller for ordinære domstoler. Det andre hovedtemaet handler om disse organenes sanksjonskompetanse.

### 6.2 Organisasjoners mulighet til å få håndhevet spørsmål om algoritmisk diskriminering

Det er ikke alltid at enkeltindivider ønsker å ta belastningen med å klage på diskriminering – enten ved å bringe en sak inn for Diskrimineringsnemnda eller for ordinære domstoler. Dette utløser spørsmål om andre aktørers mulighet til å bringe saker om algoritmisk diskriminering inn for nemnda eller domstolene. Særlig aktuelle aktører er Likestillings- og diskrimineringsombudet, samt ulike NGO'er og interesseorganisasjoner som har som formål å arbeide mot diskriminering. Situasjoner der organisasjoner mv. kan ønske å få håndhevet spørsmål om algoritmisk diskriminering kan oppstå der det foreligger kunnskap om at personer har blitt diskriminert gjennom hvordan et system er brukt, men der personene som har opplevd dette ikke selv ønsker å bringe en sak inn for nemnda eller for ordinære domstoler. En annen aktuell

situasjon er dersom det foreligger såkalt innholdsbasert diskriminering.<sup>88</sup>

Det følger av diskrimineringsombudsloven (diskrol.) § 8 at: «Nemnda behandler de sakene som bringes inn for den. En part, ombudet eller andre med rettslig klageinteresse kan bringe en sak inn for nemnda.» Bestemmelsen gir uttrykkelig ombudet en mulighet til å bringe spørsmål om algoritmisk diskriminering inn for nemnda. Bestemmelsen åpner også for at interesseorganisasjoner og NGO'er kan klage en sak inn for nemnda, forutsatt at de oppfyller kravet til «rettslig klageinteresse.»<sup>89</sup>

Twisteloven<sup>90</sup> § 1-4 (1) slår fast at foreninger og stiftelser «kan reise søksmål i eget navn om forhold som det ligger innenfor organisasjonens formål og naturlige virkeområde å ivareta», forutsatt at de har «et reelt behov for å få kravet avgjort overfor saksøkte», jf. lovens § 1-3 (2). Twisteloven § 1-4 (2) slår fast at: «Offentlige organer med oppgave å fremme særskilte interesser kan på tilsvarende måte reise søksmål for å ivareta disse.» Bestemmelsen åpner dermed for at både LDO og organisasjoner/NGO'er/stiftelser som arbeider for å forhindre diskriminering kan reise søksmål i saker som omhandler algoritmisk diskriminering.<sup>91</sup> I stedet for å reise eget søksmål etter § 1-4, kan en organisasjon også ha adgang til å opptre i et søksmål som en annen

<sup>87</sup> Vilkårene for å få saken behandlet følger av diskrimineringsombudsloven § 8 og av tvisteloven § 1-3.

<sup>88</sup> Se utredningens punkt 4.7.

<sup>89</sup> Hva som nærmere ligger i dette kravet er berørt i Hellum og Strand (2022) s. 686 flg.

<sup>90</sup> Lov 17. juni 2005 nr. 90 om mekling og rettergang i sivile tvister (tvisteloven).

<sup>91</sup> For en nærmere beskrivelse av organisasjoners søksmålsadgang, se Ot.prp. nr.51 (2004–2005) kapittel 29, spesialmerknader til § 1-4.

har reist, jf tvl. § 15-7 om partshjelp<sup>92</sup> og § 15-8 om innlegg til belysning av allmenne interesser.<sup>93</sup>

LDOs og organisasjoners adgang til å gå til søksmål i saker om diskriminering, er i tråd med det systemet som de EU/EØS-rettslige diskrimineringsdirektivene legger opp til. Alle diskrimineringsdirektivene inneholder bestemmelser om foreningers og organisasjoners adgang til å gå til sak om diskriminering på vegne av eller sammen med enkeltpersoner. Rammedirektivet (2000/78) artikkel 9 nr. 2 lyder:

*Member States shall ensure that associations, organisations or other legal entities which have, in accordance with the criteria laid down by their national law, a legitimate interest in ensuring that the provisions of this Directive are complied with, may engage, either on behalf or in support of the complainant, with his or her approval, in any judicial and/or administrative procedure provided for the enforcement of obligations under this Directive.*<sup>94</sup>

I *Accept-saken*, som omhandlet vernet mot direkte diskriminering av homofile, uttalte EU-domstolen at forbudet mot direkte diskriminering i direktiv 2000/78 kan håndheves selv om det ikke er en «identifiable complainant who claims to have been the victim of such discrimination».<sup>95</sup> Videre presiserte domstolen at direktivet «in no way precludes a Member State from laying down, in its national law, the right of associations with a legitimate interest in ensuring compliance with that directive to bring legal or administrative proceedings to enforce the obligations resulting therefrom without acting in the name of a specific complainant or in the absence of an identifiable complainant.»<sup>96</sup>

<sup>92</sup> Det følger av tvl. § 15-7 (1) bokstav b at partshjelp skal tillates for «foreninger og stiftelser, samt offentlige organer med oppgave å fremme særskilte interesser, i sak som ligger innenfor deres formål og naturlige virkeområde etter § 1-4.»

<sup>93</sup> Tvl. § 15-8 (1) lyder:  
«Skriftlig innlegg til å belyse allmenne interesser som en sak reiser, kan gis av  
a. foreninger og stiftelser innenfor rammen av deres formål og naturlige virkeområde, eller  
b. et offentlig organ innenfor sitt ansvarsområde.»

<sup>94</sup> Slik også rasediskrimineringsdirektivet (2000/43) artikkel 7 nr. 2; kjønnslikestillingsdirektivet (2006/54) artikkel 17 nr. 2; og kjønnslikestillingsdirektivet om varer og tjenester artikkel 8 nr. 3.

<sup>95</sup> C-81/12 (*Accept*) (2013), premiss 36.

<sup>96</sup> C-81/12 (*Accept*), premiss 37.

Tvl. § 1-4 oppstiller ikke noe krav om at det må foreligge konkrete individer som er diskriminert og/eller involvert for at LDO eller en interesseorganisasjon skal kunne reise søksmål om diskriminering. Tvistelovens tilnærming ligger innenfor EU/EØS-rettens rammer, slik EU-domstolen tydeliggjør i *Accept-saken*.

Dette er en praktisk problemstilling, fordi det vil kunne tenkes situasjoner der det finnes kunnskap om at et algoritmisk system har diskriminerende virkninger selv om det ikke foreligger kunnskap om enkeltpersoner som har blitt diskriminert av det aktuelle systemet. For eksempel kan personvernregler utgjøre en hindring for å få tak i informasjon på enkeltpersonnivå. I slike tilfeller vil det være viktig at det likevel foreligger en mulighet til rettslig håndheving av spørsmålet om diskriminering.

For at likestillings- og diskrimineringsloven skal kunne fungere på en robust måte i møte med KI-relatert diskriminering, vil organisasjoners søksmålsadgang kunne stå sentralt. Etter mitt syn gir tvl. § 1-4 en god ramme, fordi bestemmelsen gir organisasjoner søksmålsadgang som gjelder uavhengig av om diskriminerte enkeltindivider er identifisert eller involvert i søksmålet. I praksis handler imidlertid dette i stor grad om hvilke ressurser organisasjonene, for eksempel LDO, har til å prioritere slike saker. Dette er det god grunn til å rette oppmerksomhet mot i tiden fremover.

### 6.3 Diskrimineringsnemndas og ordinære domstolars sanksjonskompetanse

Saker om algoritmisk diskriminering må kunne håndheves gjennom de samme organer som ellers på diskrimineringsfeltet.

I vurderingen av om en sak om algoritmisk diskriminering bør bringes inn for Diskrimineringsnemnda eller for ordinære domstoler, vil det være relevant å kjenne til disse organenes sanksjonskompetanse.

Det følger av de EU/EØS-rettslige diskrimineringsdirektivene at medlemslandene skal sikre at brudd på diskrimineringsforbudet sanksjoneres ved bruk av sanksjoner som er effektive, står i et rimelig forhold til overtredelsen og som har en avskrekkende virkning.<sup>97</sup>

Dette kravet er søkt gjennomført gjennom ldl. § 38, som slår fast at «[d]en som er behandlet i strid med» forbudet mot diskriminering, kan kreve oppreisning og erstatning. I *Feryn-saken* har EU-domstolen

<sup>97</sup> Se rasediskrimineringsdirektivet (2000/43) artikkel 15; rammedirektivet (2000/78) artikkel 17; kjønnslikestillingsdirektivet (2006/54) artikkel 25; og kjønnslikestillingsdirektivet om varer og tjenester (2004/113) artikkel 14.

tydeliggjort at kravet om effektiv sanksjonering også gjelder der en organisasjon går til sak om diskriminering og det ikke er et krenket individ involvert:

*In a case such as that referred by the national court, where there is no direct victim of discrimination but a body empowered to do so by law seeks a finding of discrimination and the imposition of a penalty, the sanctions which Article 15 of Directive 2000/43 requires to be laid down in national law must also be effective, proportionate and dissuasive*<sup>98</sup>

Ldl. § 38 om oppreisning og erstatning vil for å være i tråd med EU/EØS-direktivene måtte tolkes slik at den rommer situasjoner der det er en organisasjon som bringer en sak inn for nemnda eller anlegger søksmål for ordinære domstoler.

Det er ingen materielle begrensninger i ordinære domstolars kompetanse til å behandle saker om diskriminering, og domstolene kan også tilkjenne oppreisning og/eller erstatning i saker på alle samfunnsområder. Dette vil si at domstolene er en aktuell adressat for alle saker om algoritmisk diskriminering.

Når det gjelder nemndas kompetanse, er bildet mer komplisert.

Nemndas kompetanse til å tilkjenne oppreisning og erstatning er regulert i diskrol. § 12 der det fremkommer at den kan treffe vedtak om oppreisning kun i «arbeidsforhold». Algoritmisk diskriminering kan oppstå i tilknytning til arbeidsforhold, og i så fall vil nemnda som utgangspunkt kunne tilkjenne oppreisning dersom en krenkelser av diskrimineringsforbudet har funnet sted. På alle andre områder må oppreisning eventuelt tilkjennes av ordinære domstoler.

Når det gjelder kompetansen til å tilkjenne erstatning, er nemndas kompetanse begrenset til såkalte enkle saksforhold, det vil si situasjoner der «innklagede ikke gjør gjeldende annet enn manglende betalingsevne eller andre åpenbart uholdbare innsigelser», jf. diskrol. § 8 (2).

Nemnda har kompetanse til å treffe vedtak om brudd på forbudet mot diskriminering i saker på de aller fleste samfunnsområder, jf. diskrol § 11 (1). Det er noen begrensninger i nemndas kompetanse, men disse har trolig ikke noen særlig betydning i møte med algoritmisk diskriminering.<sup>99</sup> Dette vil si at dersom det viktigste er å få et vedtak om at diskriminering har

funnet sted, vil nemnda kunne være en viktig aktør. Dette må ses i sammenheng med nemndas rolle som et gratis lavterskel håndhevingsorgan. Samtidig utløser algoritmisk diskriminering kompliserte faktiske og rettslige problemstillinger, som gjør at det kan stilles spørsmål ved om nemnda i dag er rigget på en måte som gjør den egnet til å behandle slike saker.

#### **6.4 Sammenfatninger og refleksjoner om veien videre**

*Viktige dimensjoner som er tatt opp i kapittel 6:*

- Organisasjoners mulighet til å få håndhevet spørsmål om algoritmisk diskriminering kan vise seg å bli viktig i tiden fremover. Både LDO og organisasjoner/NGO'er/stiftelser som arbeider for å forhindre diskriminering kan etter tvisteloven § 1-4 gå til søksmål i saker om algoritmisk diskriminering. Bestemmelsen oppstiller ikke noe krav om at konkrete individer må være diskriminert og involvert for at organisasjonene skal kunne reise søksmål.
- De EU/EØS-rettslige diskrimineringsdirektivene stiller krav om effektiv sanksjonering av forbudet mot diskriminering. Dette kravet gjelder også i tilfeller der det er en organisasjon som er klager eller saksøker.
- Nemndas begrensede sanksjonskompetanse i saker utenfor arbeidslivet gjør at det kan bli særlig aktuelt å bringe saker om algoritmisk diskriminering inn for ordinære domstoler. Dette synliggjør at det kan være behov for en diskusjon om hva som skal være nemndas rolle på dette feltet.

*Refleksjoner om veien videre:*

- Det vil bli viktig med bevissthet om LDOs og ulike interesseorganisasjoners rolle i å sikre at spørsmål om algoritmisk diskriminering håndheves og kontrolleres i praksis.

Dette kan handle om en bevisst strategi når det gjelder muligheten til å bringe saker inn for domstolene eller for Diskrimineringsnemnda. Det kan også handle om arbeid med å sikre at det i praksis finner sted tilsyn og kontroll med algoritmiske systemer og slike systemers virkninger – sett opp mot diskrimineringsforbudet.

Arbeidet videre bør ta utgangspunkt i det materielle forbudet mot diskriminering, slik det i dag er regulert i ldl. §§ 6-9 – og slik det eventuelt kan bli justert i lys av de nye utfordringene som bruken av algoritmiske systemer fører med seg.

<sup>98</sup> C-54/07 (*Feryn*) (2008), premiss 38. Se også C-81/12 (*Accept*), premiss 62.

<sup>99</sup> Begrensningene i nemndas kompetanse er behandlet i Hellum og Strand (2022) s. 665 flg.





# LITTERATUR

**Anne Jorun Bolken Ballangrud** og **Margrethe Søbstad**, *Likestillings- og diskrimineringsloven. Lovkommentar*. Universitetsforlaget (2021).

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

**Anupam Datta**, **Matt Fredrikson**, **Gihyuk Ko**, **Piotr Mardziel** og **Shayak Sen**, *Proxy Discrimination in Data-Driven Systems*, arXiv:1707.08120 (2017).

**European Parliament. Directorate General for Parliamentary Research Services**, rapport av Giovanni Sartor og Francesca Lagioia. *The impact of the general data protection regulation on artificial intelligence* (2020), <https://data.europa.eu/doi/10.2861/293>.

**Janneke Gerards** og **Raphaële Xenidis**, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*. European network of legal experts in gender equality and non-discrimination (2020), <https://www.equalitylaw.eu/downloads/5361-algorithmic-discrimination-in-europe-pdf-1-975>.

**Åse Haram**, *Forklarbarhet i algoritmisk velferdsforvaltning. Demokrati og rettssikkerhet ved bruk av maskinlæringsmodeller i saksbehandlingen*. Masteroppgave ved Det juridiske fakultet, UiO. Senter for rettsinformatikk, Complex 6/2021. <https://www.jus.uio.no/ifp/forskning/om/publikasjoner/complex/2021/2021-06.html>.

**Mathias Karlsen Hauglid**, *Bias and Discrimination in Clinical Decision Support Systems Based on Artificial Intelligence*. Avhandling innlevert til bedømmelse for graden Ph.D. ved Universitetet i Tromsø (2023).

**Anne Hellum** og **Vibeke Blaker Strand**, *Likestillings- og diskrimineringsrett*. Gyldendal (2022).

**Laurens Naudts**, *How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them*, s. 71–92 i Ronald Leenes, Rosamunde van Brakel, Serge Gutwirth og Paul De Hert (red.) «Data Protection and Privacy: The Internet of Bodies», Hart Publishing (2019).

**Dag Wiese Schartum**, *Digitalisering av offentlig forvaltning: fra lovtekst til programkode*. Fagbokforlaget (2019).

**Inga Strümke**, *Maskiner som tenker: algoritmenes hemmelighet og veien til kunstig intelligens*. Kagge forlag (2023).

**Inga Strümke, Marija Slavkovik og Clemens Stachl**, *Against Algorithmic Exploitation of Human Vulnerabilities*, arXiv:2301.04993 (2023).

**Marvin van Bekkum og Frederik Zuiderveen Borgesius**, *Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?* i *Computer Law & Security Review* (2023). <https://doi.org/10.1016/j.clsr.2022.105770>.

**Sandra Wachter, Brent Mittelstadt og Chris Russell**, *Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI*, i *Computer Law & Security Review* (2021). [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3547922](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922).

**Sandra Wachter**, «The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law» *Tulane Law Review* (2022) vol. 97 no. 2. <http://dx.doi.org/10.2139/ssrn.4099100>.

**Raphaële Xenidis**, «Tuning EU equality law to algorithmic discrimination: Three pathways to resilience» (2020) 27 *Maastricht Journal of European and Comparative Law* 736–58.



# KILDER

## Nasjonale lover

- Lov 17. juni 2005 nr. 62 om arbeidsmiljø, arbeidstid og stillingsvern mv. (arbeidsmiljøloven)
- Lov 17. juni 2005 nr. 90 om mekling og rettergang i sivile tvister (tvisteloven)
- Lov 16. juni 2017 nr. 50 om Likestillings- og diskrimineringsombudet og Diskrimineringsnemnda (diskrimineringsombudsloven)
- Lov 16. juni 2017 nr. 51 om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven)
- Lov 15. juni 2018 nr. 38 om behandling av personopplysninger (personopplysningsloven)

## Konvensjoner

- Konvensjon om beskyttelse av menneskerettighetene og de grunnleggende friheter, Roma 4. november 1950 (EMK)
- De forente nasjoners internasjonale konvensjon 16. desember 1966 om økonomiske, sosiale og kulturelle rettigheter (ØSK)
- De forente nasjoners internasjonale konvensjon 16. desember 1966 om sivile og politiske rettigheter (SP)

## EU/EØS-rettslig regulering

- Traktaten om Den europeiske unions virkemåte (TEUV), vedtatt 13. desember 2007, ikrafttredelse 1. desember 2009
- EUs charter om grunnleggende rettigheter (EU Charteret), vedtatt 7. desember 2000, ikrafttredelse 1. desember 2009
- Rådets direktiv 2000/43/EF af 29. juni 2000 om gjennomførelse af prinsippet om likebehandling af alle uanset race eller etnisk oprindelse (EUs rasediskrimineringsdirektiv)
- Rådets direktiv 2000/78/EF af 27. november 2000 om generelle rammebestemmelser om likebehandling med hensyn til beskæftigelse og erhverv (EUs rammedirektiv)
- Rådets direktiv 2004/113/EF af 13. december 2004 om gennemførelse af prinsippet om likebehandling af mænd og kvinder i forbindelse med adgang til og levering af varer og tjenesteydelser. (Kjønnslikestillingsdirektivet om varer og tjenester)
- Europa-Parlamentets og Rådets direktiv 2006/54/EF af 5. juli 2006 om gennemførelse af prinsippet om lige muligheder for og likebehandling af mænd og kvinder i forbindelse med beskæftigelse og erhverv (omarbejdning). (Kjønnslikestillingsdirektivet)
- Europaparlaments- og rådsforordning (EU) 2016/679 av 27. april 2016 om vern av fysiske personer i forbindelse med behandling av personopplysninger og om fri utveksling av slike opplysninger samt om oppheving av direktiv 95/46/EF (generell personvernforordning, GDPR)

Europa-Parlamentets og Rådets direktiv (EU) 2023/970 af 10. maj 2023 om styrkelse af anvendelsen af princippet om lige løn til mænd og kvinder for samme arbejde eller arbejde af samme værdi ved hjælp af løngennemsigtighed og håndhævelsesmekanismer (Lønnstransparensdirektivet)

European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act). Ordinary legislative procedure: first reading. P9\_TA(2024)0138 [Ennå ikke i kraft]

### **Forslag til ny EU-rettslig regulering**

Proposal for a Council directive on standards for equality bodies in the field of equal treatment between persons irrespective of their racial or ethnic origin, equal treatment in the field of employment and occupation between persons irrespective of their religion or belief, disability, age or sexual orientation, equal treatment between women and men in matters of social security and in the access to and supply of goods and services, and deleting Article 13 of Directive 2000/43/EC and Article 12 of Directive 2004/113/EC, (2022). COM(2022) 689 final, 7. desember 2022.

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0689>

Proposal for a directive of the European Parliament and of the Council on standards for equality bodies in the field of equal treatment and equal opportunities between women and men in matters of employment and occupation, and deleting Article 20 of Directive 2006/54/EC and Article 11 of Directive 2010/41/EU. COM(2022) 688 final, 7. desember 2022.

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0688>

### **Nasjonale forarbeider**

Prop. 81 L (2016-2017) Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven)  
Ot.prp. nr. 35 (2004–2005) Om lov om endringer i likestillingsloven mv. (Gjennomføring av Europaparlaments- og rådsdirektiv 2002/73/EF og innarbeiding av FN-konvensjonen om avskaffelse av alle former for diskriminering av kvinner med tilleggsprotokoll i norsk lov)

Ot.prp. nr. 51 (2004–2005) Om lov om mekling og rettergang i sivile tvister (tvisteloven)

### **Den europeiske menneskerettsdomstol (EMD)**

EMD, *Glukhin mot Russland (sak nr.11519/20)* (2023)

### **EU-domstolen**

*C-109/88 (Danfoss)* (1989)

*C-127/92 (Enderby)* (1993)

*C-400/93 (Royal Copenhagen)* (1995)

*C-54/07 (Feryn)* (2008)

*C-81/12 (Accept)* (2013)

*C-804/18 og C-341/19 (sammenslåtte saker, WABE og Müller)* (2021)

*C-148/22 (OP)* (2023)

### **Diskrimineringsnemnda**

Diskrimineringsnemnda, *DIN-2019-115 og DIN-2019-196* (felles behandling)

## Nettressurser

Advokatfirmaet Thommessen ved Espen A. Werring, Christopher Sparre-Enger Clausen og Ane Lerheim Mossige, «AI Act: Fremtiden for kunstig intelligens i EU», 20. desember 2023.

<https://www.thommessen.no/aktuelt/ai-act>, lesedato: 29. februar 2024.

Digdir, «Hva er kunstig intelligens?» <https://www.digdir.no/kunstig-intelligens/hva-er-kunstig-intelligens/4133>, lesedato: 1. mars 2024.

European Parliament, News, «Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI», Pressemelding 9. desember 2023. <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>, lesedato 29. februar 2024.

Likestillings- og diskrimineringsombudet, *Innebygd diskrimineringsvern*. Veileder (2023).

[https://www.ldo.no/globalassets/\\_ldo\\_2019/\\_bilder-til-nye-nettsider/ki/ldo.-innebygd-diskrimineringsvern.pdf](https://www.ldo.no/globalassets/_ldo_2019/_bilder-til-nye-nettsider/ki/ldo.-innebygd-diskrimineringsvern.pdf).

Regjeringen, EØS-notatbasen, «Likelønnsdirektivet» (2023). <https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2021/mai/eus-likelonnndirektiv/id2856216/>, lesedato 12. februar 2024.







**Likestillings- og  
diskrimineringsombudet**

## Vi hjelper deg

Gratis juridisk veiledning  
Finn oss på nett: [www.ldo.no](http://www.ldo.no)  
Ring oss: 23 15 73 00



[facebook.com/mittOmbud](https://facebook.com/mittOmbud)



[twitter.com/mittOmbud](https://twitter.com/mittOmbud)



[instagram.com/mittOmbud](https://instagram.com/mittOmbud)