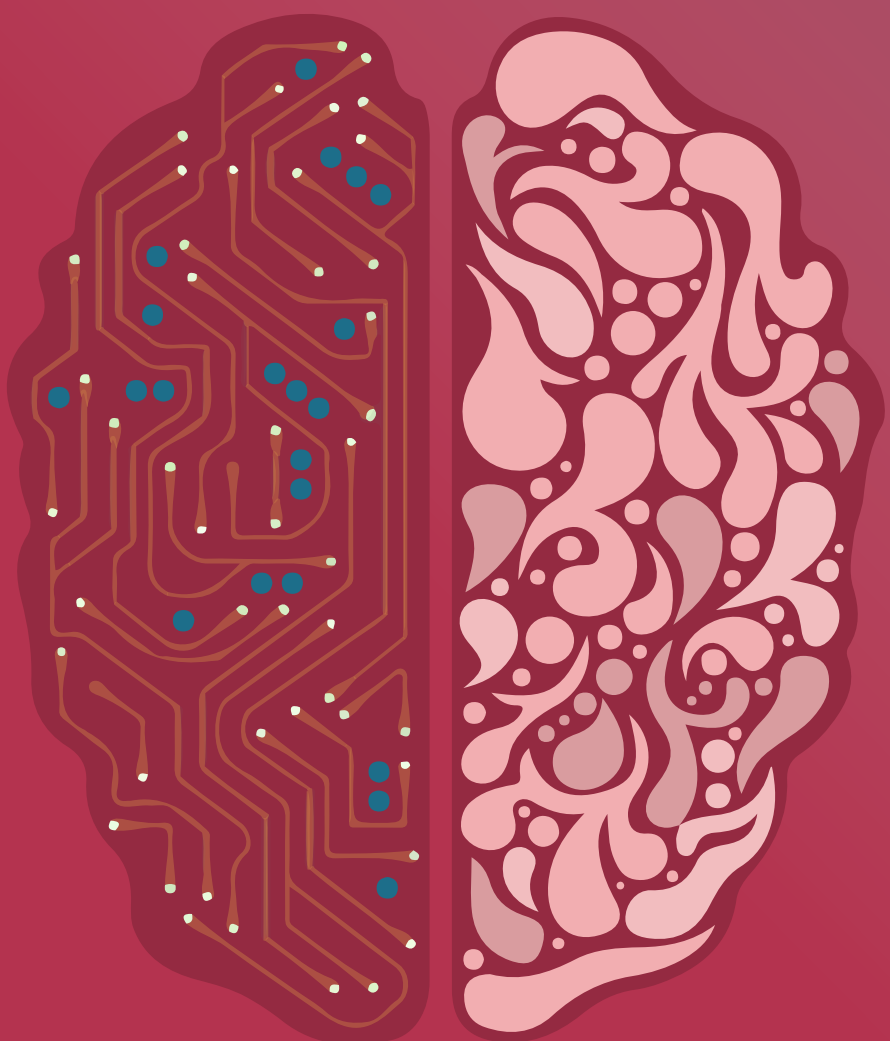
Norges institusjon for menneskerettigheter

Ny teknologi og menneskerettigheter



Innhold

Forord

Den digitale omveltingen tegner til å bli en av de raskeste og mest dramatiske samfunnsendringene i menneskehetens historie.

Ny teknologi åpner blant annet opp for helt nye former for samkvem, styring, formidling og kommunikasjon. Grunnleggende funksjoner i samfunnet vårt er i rask endring på positive måter som følge av digitale systemer, kunstig intelligens og maskinlæring. Men det er også faretegn. I 2021 uttrykte FNs høykommissær for menneskerettigheter dilemmaet slik: «Nye teknologier kan være en enorm positiv kraft, og kan hjelpe samfunn med å løse noen av de største utfordringene i vår samtid. Men ny teknologi kan også få negative, ja sågar katastrofale følger, dersom de rulles ut uten hensyn til hvordan de påvirker menneskerettighetene.»[1]

> «Det neste tiåret vil teknologiutviklingen trolig gå fortere enn det siste hundreåret, og endringene for samfunnet vårt kan bli like omfattende som den industrielle revolusjon.»
>
> *Richard More, sjef MI6, 2021*

Menneskerettighetene er rettigheter som norske myndigheter er forpliktet til å respektere og sikre for alle i Norge. De mest sentrale menneskerettighetskonvensjonene, herunder Den europeiske menneskerettskonvensjonen (EMK) og FNs konvensjoner om sivile og politiske rettigheter (SP) og økonomiske og kulturelle rettigheter (ØSK), har forrang foran norsk lov, og myndighetene har ansvar for å overholde bestemmelsene i disse traktatene når ny teknologi tas i bruk. Teknologiutviklingen gir myndighetene store muligheter til å forbedre måten menneskerettighetene sikres på. Dette gjelder for eksempel retten til liv og helse. Samtidig gir ny teknologi noen betydelige menneskerettslige utfordringer. I noen land er bekymringen rundt ny teknologi at den vil gi statlige myndigheter potente verktøy for gjennomgripende overvåkning og økt undertrykkelse. I Norge er utfordringene særlig de utilsiktede effektene av ny teknologi. Teknologiens konsekvenser for menneskerettighetene er i stor grad upløyd mark for alle stater i verden. Sammenlignet med andre land ligger Norge svært langt fremme med digitalisering av offentlig forvaltning, og er slik sett et slags testlaboratorium. Det fordrer at vi er våkne.

De internasjonale menneskerettighetene gir myndighetene plikt til å være på vakt ovenfor utilsiktede effekter av teknologiske nyvinninger og til å kartlegge hvordan de påvirker menneskerettighetene. Når teknologien endrer flere samfunnsområder samtidig, er det også et myndighetsansvar å fange opp og adressere tverrgående menneskerettslige utfordringer.

Rapporten «Ny teknologi og menneskerettigheter» gir et overblikk over de viktigste menneskerettslige utfordringene som oppstår når teknologiutviklingen gir maskinene

[1] FNs høykommisær for menneskerettigheter, *The right to privacy in the digital age*. FN-rapport, 13. sept. 2021, avsn. 2.

en stadig mer sentral plass i styring og forvaltning av det norske samfunnet. Rapporten tar for seg et stort og teknisk krevende felt. Når ny teknologi kombineres med de internasjonale menneskerettighetene, møtes to verdener hvor den ene krever teknisk ekspertise og den andre krever juridisk ekspertise. Problemstillingene egner seg best for dobbelteksperter, og er gjerne mat for spesialister. Her forsøker vi likevel å gjøre temaet konkret og allment tilgjengelig, og løfte problemstillingene opp på et overordnet plan. Rapporten trekker opp overordnede perspektiver rundt hvordan ny teknologi forsterker velkjente menneskerettighetsproblemer og samtidig gir oss helt nye utfordringer på menneskerettighetsfeltet.

Metoden som er benyttet i rapporten er en kombinasjon av samfunnsvitenskapelig og juridisk metode. Innholdet baserer seg på sosiologiske studier av teknologiens faktiske effekter, prediktive analyser av hvordan teknologien trolig kommer til å virke, samt juridiske analyser av hvilke rettslig bindende regler og prinsipper som aktualiseres av ny teknologi. Rapporten tar sikte på å belyse sider av teknologiutviklingen som gir særlige utfordringer for respekt for og sikring av menneskerettighetene. Avslutningsvis gir den et knippe anbefalinger til norske myndigheter.

De mange temaene i rapporten kunne vært omtalt i mye større dybde og bredde. Men heller enn å zoome inn, ønsket vi med denne rapporten å løfte blikket og tegne et større bilde. Vi håper at det kan inspirere til at mer detaljerte tegninger vil følge.

God lesning.[2]

[2] Denne rapporten er utarbeidet av Norges institusjon for menneskerettigheter (NIM). Hovedforfatter er spesialrådgiver i NIM, Cecilie Hellestveit. Medforfatter er rådgiver Mathilde Wilhelmsen.

1. Hva er ny teknologi?

Med «ny teknologi» mener vi de nye formene for digitale verktøy som utgjør grunnmuren i informasjonssamfunnet.

Teknologiutviklingen gjør at enorme mengder generell og spesiell digital informasjon nå kan samles inn, oppbevares, forvaltes og nyttiggjøres av maskiner. Ved hjelp av kunstig intelligens, kan data analyseres og benyttes av autonome systemer for beslutningsstøtte eller beslutninger. Maskinene lærer å utføre oppgaver som normalt krever menneskelig intelligens eller skjønnsutøvelse. De kan gjenkjenne mønstre, drive med erfaringsbasert læring, trekke logiske slutninger, foreta prediksjoner eller treffe beslutninger om handlinger og deretter utføre dem – digitalt eller som smart software for autonome fysiske systemer. «Ny teknologi» omfatter både teknologiske nyvinninger og velkjent teknologi brukt på en ny måte.

> «Den digitale tidsalderen har åpnet nye grenser for menneskelig velferd, kunnskap og utforsking, [...]digitale teknologier gir nye virkemidler for å promotere, beskytte og utøve menneskerettighetene.»
>
> *FNs Generalsekretær, 2019*

Selv om ny teknologi utvikles og benyttes digitalt, får de i økende grad effekter i den fysiske verden. Etter hvert som «tingenes internett» blir mer vanlig, vil konvertering fra virtuell til reell øke i omfang. Da vil også forskjellen mellom det digitale og det tinglige bli mindre tydelig.

Terminologi

Kunstig intelligens refererer til datateknologi (hardware og software) som kan imitere intelligent (menneskelig) oppførsel. Det er et paraplybegrep som omfatte mange typer maskinlæring. Algoritmer beskriver fremgangsmåten for å løse en oppgave, og kunstig intelligens gjør at software kan lære fra akkumulert data gjennom å benytte intelligente algoritmer.

Maskinlæring er en underkategori av kunstig intelligens som baserer seg på automatisert modellbygging gjennom analyse. Makskinlæring kan beskrives som et sett teknikker og verktøy som lar maskiner «tenke» ved å lage matematiske algoritmer basert på akkumulert data. Maskinlæringskapasitet innebærer at maskiner er i stand til å «lære» ny informasjon og benytte den uten å være pre-programmert til det. Her utstyres datasystemer med evne til å lære fra data uten eksplisitt programmering. Enkelte metoder er inspirert av funksjoner fra den menneskelige hjerne, nettverk, statistikk og fysikk, og maskinlæring kan identifisere

skjulte og nye innsikter som ligger i datamengder av et visst omfang. For enkelte typer funksjoner forutsetter effektiv maskinlæring tilgang til store mengder data. Maskinlæring reiser derfor en rekke problemstillinger knyttet til innsamling, prosessering og vurdering av omfangsrike data.

Dyplæring er en underkategori av maskinlæring, som kan være bygget opp rundt de samme prinsippene som det nevrale nettverket i hjernen. Slike systemer tar ofte utgangspunkt i et kjent treningsdatasett som hjelper de selvlærende algoritmene å få nettverket til å utføre en oppgave. Dette forutsetter at nettverket selv kan avgjøre hva som er riktig respons for å løse oppgaven. De siste årene har gjennombrudd i dyplærling ført til at maskiner kan utføre oppgaver som ikke kan direkte programmeres, som generering eller klassifisering av bilder eller språkoversettelse. Maskinenes kapasitet til analyse og presisjon overgår etter hvert menneskelige evner. Dyplæring gjør at en algoritmemodell blir mer kapabel, men også mer komplisert. I neste omgang kan dyplæring tåkelegge innsikt i modellens indre logikk, og gjøre det umulig å vurdere eller etterprøve modellene eller dataene. Maskinene forstår mer, menneskene forstår mindre.

2. Hva gjør ny teknologi?

Teknologiutviklingen kan bidra til å styrke menneskerettighetene. Men ny teknologi setter også menneskerettighetene på en hard prøve.

Den første store europeiske rapporten om samfunnsmessige følger av kunstig intelligens fra 2017 predikerte at kunstig intelligens vil påvirke vår sikkerhet, etikk, lover og regler, demokrati, relasjoner, innsynsmuligheter, privatliv, arbeid og utdannelse.[3] Digital teknologi kan øke effektiviteten og treffsikkerheten i offentlig styring og forvaltning. Ny teknologi kan sette myndighetene bedre i stand til å innfri sine positive menneskerettighetsforpliktelser, som retten til utdannelse, retten til arbeid og retten til helse. Kunstig intelligens kan benyttes til å finne personer som får utbetalt mer penger enn de har krav på fra Nav, identifisere mennesker som har forhøyet risiko for bivirkninger ved medisinbruk eller de som er i risikogruppen for å bli langtidsledige. Kunstig intelligens og maskinlæring kan også benyttes til å predikere fremtidige handlingsmønstre, og bistå myndighetene med å hindre kriminalitet og sørge for mer treffsikre tiltak for å sikre andre menneskerettigheter.

Den europeiske rapporten fra 2017 forutså imidlertid også noe annet – at utstrakt bruk av kunstig intelligens vil øke økonomisk ulikhet.[4]

Kjernen i våre dagers bekymringer rundt maskinlæring og kunstig intelligens ligger nettopp her: frykt for at våre samfunn vil formes og kontrolleres av digitale virkemidler med ukjente effekter, istedenfor at vi tar kontroll over hvordan teknologien brukes og utnyttes til beste for samfunnet. For ny teknologi kan true menneskerettighetene på en rekke måter og områder.

> «Vi skaper våre teknologier, og deretter omskaper de oss.»
>
> *Marshall McLuhan, mediefilosof*

Ny teknologi kan forsterke diskriminerende praksis. Den kan også skape nye former for diskriminering. Tilgang til digitale flater er gjerne ulikt fordelt når det gjelder språk, innhold og infrastruktur, og eksisterende skjevheter kan forsterkes av ny teknologi.[5] Hvilke data som er tilgjengelige vil også gjerne reflektere ulikheter av sosial, økonomisk, politisk eller annen art.[6] Ettersom maskinene baserer seg på algoritmer som mennesker har laget, kan diskriminerende praksis videreføres digitalt. I Storbritannia har for eksempel bruk av kunstig intelligens for å forutse fremtidige lovbrudd vist seg å diskriminere de som kommer fra lavinntektsområder. Resultatet vil alltid bero på den informasjonen man legger inn, og diskriminerende praksis vil derfor kunne gjentas og forsterkes digitalt. Måten teknologien

[3] Den europeiske økonomiske og sosiale komite, *Opinion on AI & Society* (INT/806), 2017.
[4] Den europeiske økonomiske og sosiale komite, 2017.
[5] FNs generalsekretærs rapport. *Roadmap for Digital Cooperation: Implementation of the Recommendations of the High-Level Panel on Digital Cooperation*, avsn. 6, 2020.
[6] FNs menneskerettighetsråd. *Human Rights of Older Persons: The Data Gap*, rapport, 2020.

fungerer kan også skape nye former for diskriminering. I USA har for eksempel ansiktsgjenkjenning hatt problemer med å identifisere mørkhudede, og diskriminering har blitt en konsekvens.

Med digitale analyseverktøy tiltar også samfunnskontrollen. Ny teknologi gir nye muligheter for å avdekke misbruk av statlige tjenester. Ytelser som statlige myndigheter gir individer for å sikre deres menneskerettigheter kan gjøres mer effektive og rettmessige av ny teknologi. Utfordringen er at når vi som samfunn får tilgang til teknologi som er egnet til å avdekke misbruk, vil bruken av analyseverktøyene gjerne rettes inn mot å «ta dem som lurer systemet». Når styresmaktene bruker teknologien til å identifisere alle som kan tenkes å jukse istedenfor å gjøre nytte av teknologien til å øke individenes velferd, øker risikoen for en digital velferdsdystopi.[7]

Under covid-19-pandemien var kunstig intelligens svært viktig for å bidra til å finne en vaksine raskt, men også for å effektivisere smittesporing. Kunstig intelligens bidro til å styrke retten til liv og helse. Men her kommer kunstig intelligens raskt på kollisjonskurs med andre menneskeretter som gir vern knyttet til informasjon om hver av oss, fra den samles inn frem til den benyttes eller slettes. Informasjonen kan gi grunnlag for å fatte beslutninger som innebærer inngrep i våre materielle menneskerettigheter, for eksempel et bevegelsesforbud knyttet til smittevern. Teknologiutviklingen kan sette retten til liv og helse hos noen, opp mot andres rett til privatliv, ytringsfrihet, forsamlingsfrihet, samt forbudet mot diskriminering.[8] Når vi tar i bruk ny teknologi med stor samfunnsnytte og effekt for enkelte sentrale menneskerettigheter, øker presset på andre av våre menneskerettigheter.

Ny teknologi gjør det mulig å samle inn enorme mengder av data om hver enkelt av oss. Innholdsdata viser hva som blir skrevet, mens kommunikasjonsdata viser hvem som har kommunisert med hvem. Slik data kan for eksempel gi oversikt over hvem som har vært til stede på en demonstrasjon mot myndighetene og hvem som har sosial omgang, og kan brukes til å kartlegge både kommunikasjon og bevegelse. Informasjonen kan avdekke allerede utførte handlinger, sosial aktivitet eller medlemskap i gruppe eller nettverk.[9] Basert på disse dataene, gjør digitale analyseverktøy det mulig å identifisere handlingsmønstre.

DATATYPER

- **Innholdsdata:** informasjon som viser innhold i data som deles.
- **Kommunikasjonsdata:** informasjon som viser hvem man kommuniserer med. Kommunikasjonsdata er *metadata*.
- **Metadata:** informasjon *om* innholdsdata. Metadata beskriver relevant informasjon rundt innholdsdata som gir kontekst.
- **Bulkdata:** samling av store mengder rådata, for eksempel metadata.
- **Stordata:** store, mangeartede, ustrukturerte og dynamiske data som ikke uten videre kan analyseres med tradisjonelle databehandlingsverktøy.
- **Smart data:** informasjon som organiseres og forberedes for analyse allerede på innsamlingstidspunktet.

[7] Philip Alston, *Report from the UN special rapporteur on extreme poverty and human rights to the GA*, 2020.
[8] *Szabo and Vissy v. Ungarn* (37138/14), 12. jan. 2016, avsn. 57.
[9] Government of the United Kingdom, «Operational Case for Bulk Powers», 2016, pkt. 5.6.

Ny teknologi åpner også for nye typer uønsket vilkårlighet. For å virke best mulig er en del digitale løsninger avhengige av store mengder informasjon. Verktøyene må også vedlikeholdes og justeres jevnlig. Jo mer data og oppdaterte verktøy, desto mer treffsikker vil teknologien være, med fordelaktige konsekvenser for oss som samfunn (det blir billigere) og som enkeltindivider (vi får mer skreddersydde løsninger). Datamining og bruk av kunstig intelligens til å sette sammen data og statistikk på nye måter kan gi ny viten om oss som enkeltindivider. Når data om en person settes sammen, kan teknologien skape ny informasjon om personen (profilering). Imidlertid forsvinner da den opprinnelige konteksten. Samles slike data inn uten personens vitende og vilje kan det bryte med retten til selvbestemmelse over informasjon om en selv. Når informasjon om en person samles sammen gjennom et helt liv, vil den også forfølge vedkommende for alltid. I tillegg kan dataene være ufullstendige, upresise eller utdaterte, og kan sette personen i et feil lys.

Pålitelig og sikker bruk av kunstig intelligens er derfor avhengig av data av betydelig kvalitet og i betydelige mengder. Uten dette kan kunstig intelligens i liten grad utnyttes til det beste for samfunnet. For å bøte på den betydelige faren for feil, de-kontekstuell og selektiv informasjon som kan medføre utilsiktet diskriminering, er det tale om å bevege seg fra «stor data» til «smart data». Slik kan vi i noen grad løse utfordringene om «feil» informasjon. Til gjengjeld samles betydelig mer og bedre informasjon om den enkelte, som igjen reiser krevende problemstillinger om retten til privatliv.

Den samfunnsmessige nytten av ny teknologi kan være betydelig. Når digitale analyseverktøy gjør det mulig å identifisere handlingsmønstre, kan man etter hvert også forutsi fremtidige handlinger. Da er vi over i prediksjon om fremtidig adferd. Slike verktøy kan være med på å hindre for eksempel terrorisme eller cyber-kriminalitet. Økt digital overvåkning kan dermed bidra til å tilrettelegge for bedre gjennomføring av statens menneskerettighetsforpliktelser, som statens plikt til å sikre retten til liv og frihet.

Med teknologi som kan utføre prediksjon av fremtidige handlingsmønstre, kan man også i større grad forutsi fremtidig fare. Når slik informasjon finnes eller kan teknologisk fremskaffes, blir *prevensjon* (avverging av fare) mer naturlig som en sentral myndighetsoppgave. Med prevensjon som sentralt mål, blir monitorering og overvåkning et viktigere virkemiddel. Det blir et slags rustningskappløp mellom overvåkning og kontroll på den ene siden, og personvern på den andre. Slik øker presset på hver av oss om å akseptere mer vidtrekkende inngrep, fordi teknologien til å «gjøre noe med et samfunnsproblem» finnes og er noe vi kan nyttiggjøre oss av. Det kan ofte være vanskelig å argumentere mot slike forslag, som gjerne har gode intensjoner og skal ivareta viktige samfunnsinteresser, og som vil tjene oss både som enkeltpersoner og som fellesskap. Den europeiske menneskerettsdomstol (EMD) har ved flere anledninger formulert dette dilemmaet for statlig overvåkning ved at man risikerer å «undergrave demokratiet under dekke av å forsvare det».

Ny teknologi åpner for helt nye former for kartlegging, aldri tidligere muliggjort, verken online eller offline. Profilering kan brukes til å avlede personlige egenskaper eller predikere individers fremtidige oppførsel. Profilering kan også gi innsikt i fremtidig adferd på måter som gir muligheter til målrettede aksjoner mot enkeltmennesker for å påvirke deres adferd, for eksempel stemmegiving i valg. Det er særlig her ny teknologi eksponerer oss for nye typer overgrep. Informasjon om enkeltpersoner som er samlet inn uten samtykke og viten gjennom

en lang rekke prosesser, kan utnyttes til å manipulere enkeltmennesker, for eksempel til å profilere folks politiske preferanser og manipulere demokratiske prosesser. EMD har klargjort at økt overvåkning muliggjort gjennom moderne teknologi har potensialet til både å underminere individuelle rettigheter og et effektivt deltakende demokrati.[10]

Når styresmaktene i økende grad belager seg på data og digitale analyseverktøy, reduserer man samtidig saksbehandleres menneskelige skjønnsutøvelse. Digitalisering og kunstig intelligens gir også anledning til å beregne borgeres fremtidige adferd, med en viss grad av sikkerhet. Dette fører oss mot en verden hvor myndighetene får stadig mer innsyn og kontroll med individene, mens individene får stadig mindre innsyn og kontroll med myndighetene og deres maskiner. Teknologiutviklingen legger derfor press av en helt ny art på menneskerettighetene.

PÅVIRKNING I DEMOKRATISKE VALGPROSESSER

- Under valget i USA i 2016, benyttet Trump-kampanjen nesten seks millioner ulike kampanjeannonser på Facebook. De som ble delt og likt mest ble spredd sammen med donasjonsforespørsler. Analyseselskapet Cambridge Analytica hadde tilgang til data fra 87 millioner Facebook-brukere, og tilbød profilering av disse brukerne til Trump-kampanjen.
- Før Storbritannias folkeavstemning om Brexit i 2016, ble Vote-no-sidens målrettede kampanjemateriell spredd rundt en milliard ganger på Facebook. Britiske myndigheter har gitt en rekke bøter for ulovlig bruk av personlig data for å påvirke stemmegivingen under Brexit.

[10] *Szabo and* Vissy *v. Ungarn*, (37138/14), 12. jan. 2016, avsn. 57.

3. Menneskerettighetsforpliktelsene er teknologinøytrale

Menneskerettighetene gir et internasjonalt, legitimt og nyttig rammeverk for å forutse, forhindre og adressere en del av risikoene og farene ved kunstig intelligens, maskinlæring og digitalisering.

Norge har en av de mest digitaliserte offentlige sektorene i verden. Vi har en offensiv strategi om digitalisering og bruk av kunstig intelligens i offentlig forvaltning, og stadig mer av myndighetenes samhandling med borgerne preges av ny teknologi.[11]

> «En stat som inntar en pionerrolle i utviklingen av ny teknologi bærer et spesielt ansvar for å finne en god balanse mellom retten til privatliv og andre menneskerettigheter.»
>
> *Den europeiske menneskerettsdomstolen, 2008*

Menneskerettighetene regulerer forholdet mellom hver stat og de individene som befinner seg under statens jurisdiksjon. I likhet med andre konvensjoner og traktater er menneskerettighetskonvensjoner rettslig bindende avtaler mellom de statene som velger å slutte seg til dem. Myndighetene er forpliktet til å respektere og sikre menneskerettighetene, uavhengig av hvilken teknologi som benyttes. Menneskerettighetsforpliktelsene er teknologinøytrale.

Grunnlovens § 92 sier at «Statens myndigheter skal respektere og sikre menneskerettighetene slik de er nedfelt i denne grunnlov og i for Norge bindende traktater om menneskerettigheter». Det er altså offentlige myndigheter på alle nivåer, både statlig og kommunalt, som er forpliktet av menneskerettighetene. EMK, SP og ØSK beskytter de rettighetene som er særlig aktuelle i møte med ny teknologi, som retten til privatliv, ytringsfrihet og ikke-diskriminering. Alle tre konvensjonene er inkorporert i menneskerettsloven.[12] Lovens § 3 gir menneskerettighetene forrang foran annen norsk lovgivning hvis det skulle oppstå regelkonflikt.

Når oppgaver som tidligere ble forvaltet av mennesker, delegeres eller settes bort til maskiner, oppstår nye utfordringer rundt oversikt, innsyn og etterprøvbarhet. Mange systemer, for eksempel innenfor helsevesenet eller Nav, er per i dag utviklet av algoritmer som lages av mennesker. Man taler gjerne om *algoritmebaserte* beslutninger, hvor beslutninger foretas av mennesker, men de er

[11] Regjeringens digitaliseringsstrategi (2019), Regjeringens nasjonale strategi for kunstig intelligens (2020) og Hurdalsplattformen (2021). Se også SILO, *The Nordic State of AI - The 2021 Report.*

[12] Lov om styrking av menneskerettighetenes stilling i norsk rett (menneskerettsloven) av 21. mai 1999. Andre konvensjoner som er inkorporert i denne loven er FNs barnekonvensjon og FNs kvinnekonvensjon.

basert på informasjon som er generert av eller støttet av algoritmeberegninger. En annen type beslutninger er *algoritmefremskaffede* beslutninger, hvor beslutningene hviler på menneskelige vedtak basert på valgmuligheter som er fremlagt av maskingenerert informasjon, men hvor det er lagt inn et visst handlingsrom for menneskelig skjønn. Endelig har man *algoritmedeterminerte* beslutninger, hvor det er maskinbaserte systemer som «fatter» beslutningen. Etter hvert som maskinlæring og selvlærende, autonome systemer blir mer utbredt, vil dette gjøre oversikt, innsyn og etterprøvbarhet stadig mer krevende. Jo mer komplekse selvlærende maskiner blir, desto vanskeligere (eller umulig) blir det for mennesker å etterprøve hvordan et system har kommet til et resultat.

TEKNOLOGINØYTRAL

- Norske myndigheters menneskerettighetsforpliktelser er teknologinøytrale.
- EMK, SP, Grunnloven og menneskerettsloven pålegger norske myndigheter å respektere og sikre menneskerettighetene uavhengig av hvilken teknologi et inngrep hviler på.
- Inngrep som baserer seg på ny teknologi må tilfredsstille alminnelige krav under menneskerettighetene.

Tiltak som griper inn i menneskerettighetene (inngrep) kan være lovlige dersom de er tillatt etter de internasjonale menneskerettighetsbestemmelsene. Det innebærer at et inngrep må ivareta et legitimt formål, være forholdsmessig og ha hjemmel i lov. Dersom disse vilkårene er innfridd, har stater anledning til å begrense enkelte menneskerettigheter med formål om for eksempel å beskytte andre rettigheter, ivareta nasjonal sikkerhet eller opprettholde offentlig ro og orden. Dersom en statlig myndighet får tilgang på de elektroniske helsedataene til innbyggerne, må denne tilgangen (inngrepet) være basert på en begrensning i retten til privatliv som tilfredsstiller vilkårene om legitimt formål, forholdsmessighet og lovhjemmel. Masseinnsamling og lagring av elektroniske data anses å utgjøre et inngrep i retten til privatliv selv om innsamlingen ikke er knyttet til enkeltindivider.[13] Slik masseinnsamling er kun tillatt dersom overvåkningen er i tråd med de begrensningene i retten til privatliv som menneskerettighetene tillater.[14]

Når ny teknologi tas i bruk som påvirker forholdet mellom staten og individene, kan dette føre til tilsiktede eller utilsiktede menneskerettighetsbrudd fra myndighetene. Rollen til de internasjonale menneskerettighetene i møte med ny teknologi, er slik sett å være rettesnor for myndighetenes handlinger, og bidra til at ny teknologi blir mer til nytte enn til skade for samfunnet og den enkelte.

[13] Se eksempelvis *Uzun v. Tyskland*, (35623/05), 2. sept. 2010, avsn. 61.

[14] *Big Brother Watch and Others v. Storbritannia*, (51170/13, 62322/14, 24960/15), 25. mai 2021.

4. Nye utfordringer for menneskerettighetene

Når ny teknologi blir utbredt som verktøy for statlige myndigheter, private aktører og borgerne, oppstår nye menneskerettslige problemstillinger.

Digitalisering og nye teknologiske verktøy forsterker noen velkjente utfordringer for menneskerettighetene. Enkelte av menneskerettsutfordringene er også helt nye – nærmest like banebrytende som teknologien selv.

> «Den raske teknologiutviklingen innenfor kunstig intelligens og måten teknologien brukes av næringsliv og statlige myndigheter, har helt overkjørt innsatsen for å sikre menneskerettighetene i møte med kunstig intelligens.»
>
> *FNs høykommissær for menneskerettigheter, 2021*

Maskinenes veier blir uransakelige

Systemer som er basert på maskinlæring, kunstig intelligens eller autonomi er ofte ugjennomtrengelige for menneskelige beslutningstakere, et problem kjent som «sort boks-problemet». I motsetning til tradisjonelle algoritmer, vil beslutninger basert på dyplæring kunne bli nær sagt umulige å ettergå for mennesker. Dermed vil de heller ikke kunne overprøves eller på andre måter forklares for publikum eller for dem som har ansvar for kontrollmekanismene. Det tales i økende grad om et «forklaringsprinsipp».

«Alle» er sårbare i møte med ny teknologi

De internasjonale menneskerettighetene skal beskytte enkeltmennesket i møte med statsmaktene. Menneskerettighetene er gjerne spesielt viktige for personer som er sårbare eller som tilhører utsatte grupper.

Problemstillingene knyttet til menneskerettigheter og ny teknologi omfatter alle medlemmer av samfunnet. Ny teknologi og kommunikasjonsdata gjør at informasjon ikke lenger trenger å være «personliggjort» for å identifisere et individ.[15] Teknologien kan også avsløre svært private og intime detaljer.[16] Ny teknologi representerer slik sett menneskerettighetsutfordringer som rammer en mye større del av samfunnet enn bare de som er i kontakt med for eksempel rettsvesenet, barnevernet eller eldreomsorgen. I møte med ny teknologi er alle mennesker sårbare.

Ny teknologi øker sårbarheten til utsatte grupper

Fra et menneskerettighetsperspektiv vil sårbare grupper gjerne være særlig utsatte i møte med ny teknologi. Menneskerettigheter handler i stor grad om minoritetsvern. Vernet av utsatte

[15] FNs høykommissær for menneskerettigheter, *Report on the right to privacy in the digital era*, 2018, avsn. 54.
[16] FNs høykommissær for menneskerettigheter, *Right to privacy*, 2018 avsn. 25.

grupper ivaretas gjerne ved å sikre en god saksbehandling som er egnet til å identifisere hvilke særlige behov en gruppe har. For eksempel gir FNs barnekonvensjon rettigheter til enhver person under 18 år. Myndighetene må ta hensyn til «barnets beste» under saksbehandling og i konkrete avgjørelser som treffes.[17] Barnekonvensjonen stiller krav til barnets muligheter for å bli hørt og bidra under saksutredning, og det påligger statlige myndigheter et særlig ansvar for å etterleve konvensjonsforpliktelsene gjennom saksbehandlingen. Ved bruk av ny teknologi blir gjerne store deler av saksbehandlingen automatisert og underlagt algoritmebaserte beslutninger. Det vil kunne oppstå risiko for at myndighetene ikke ivaretar barns rett til å bli hørt i alle saker som berører dem, dersom myndighetene ikke har et særlig blikk på denne problemstillingen. Når ny teknologi gjør saksbehandlingen mer utilgjengelig, blir sårbare grupper mer utsatte.

Digitalt utenforskap

Ny teknologi skaper også helt nye sårbarheter. Av ulike grunner vil mange falle utenfor i den digitale utviklingen. I takt med at digitaliserte løsninger overtar stadig større deler av hverdagen vår, også oppgaver som er helt dagligdagse, vokser bekymringene rundt digitalt utenforskap. Den digitale utviklingen kan innebære at personer med nedsatt funksjonsevne ikke får tilgang på tjenester, noe som vil kunne utgjøre brudd på FNs konvensjon om rettighetene til mennesker med nedsatt funksjonsevne (CRPD). For eksempel har en stor andel eldre lav digital deltakelse eller svak digital kompetanse, og risikerer å falle utenfor.

Ny teknologi kan skape nye former for utenforskap.

Teknologiselskaper som inngrepsaktører

Staten har ansvar for menneskerettighetsbrudd som statlige myndigheter selv forårsaker. Staten er i tillegg forpliktet til å *sikre* menneskerettighetene i den forstand at det må iverksettes tiltak for å hindre at tredjeparter bryter individenes menneskerettigheter. For eksempel har statene omfattende plikter til å forebygge og forhindre vold mellom enkeltpersoner, når myndighetene visste eller burde visst om denne faren. De har også plikt til å etterforske og straffeforfølge slik vold.[18] Dette kalles gjerne positive forpliktelser.

Ny teknologi kjennetegnes imidlertid av at det er store selskaper, gjerne store multinasjonale selskaper som GAFAM og BATX, som utvikler, forvalter og styrer teknologiutviklingen. I tillegg er det gjerne teknologiselskapene som setter rammene for den digitale aktiviteten – de står selv for reguleringen. Teknologiselskapene representerer dermed en form for regulerende myndighet som på mange måter både er sterkere og mer direkte enn myndighetene. Ny teknologi gir derfor selskaper usedvanlig stor makt over individene.

GAFAM og BATX

GAFAM er forkortelse for de amerikanske selskapene Google, Amazon, Facebook, Apple, og Microsoft, mens BATX står for de kinesiske motsatsene Baidu, Alibaba, Tencent og Xiaomi.

[17] FNs barnekonvensjon (BK), vedtatt 20. nov. 1989, trådte i kraft 2. sept. 1990, art. 3 nr. 1. BK artikkel 12 stiller mer presise krav til saksbehandlingen.

[18] Se f.eks. *Opuz v. Tyrkia* (33401/02), 9. juni 2009, avsn. 150, og *Mustava Tunç og Fecire Tunç v. Tyrkia* (24014/05), 14. apr. 2015, avsn. 171.

Teknologiselskapene blir stående mellom staten og innbyggerne når det gjelder kontroll og myndighet. Under menneskerettighetene er selskaper derimot ikke direkte pliktsubjekter. Det er staten som har plikt til å sikre menneskerettighetene. Dette kan etter omstendighetene medføre krav om at statene tar aktive grep for sikre menneskerettighetene i forholdet mellom teknologiselskaper og enkeltmennesker.

Utfordringen med at teknologiselskaper har mye makt, men fravær av direkte menneskerettighetsforpliktelser, kan løses på to måter. En tilnærming er å vurdere hvor langt myndighetene har plikt til å sikre rettigheter som ytringsfrihet og personvern dersom teknologiselskapene ikke selv gjør nok for å ivareta disse rettighetene. Det er fortsatt mye upløyd mark når det gjelder hvordan enkelte rettigheter skal ivaretas i møtet med ny teknologi.

En annen måte å gripe problemet an på er å se det gjennom «linsen» selskapers menneskerettighetsansvar. De siste par tiårene har oppmerksomheten økt rundt selskapers ansvar for menneskerettigheter i tilknytning til den virksomheten de driver. Både FN og OECD har forsøkt å «bygge broer» over det gapet som finnes mellom statenes rettslige plikter og selskapers ansvar. FNs veiledende prinsipper for næringsliv og menneskerettigheter (UNGP) utgjør de ledende internasjonale standardene for næringslivets menneskerettighetsansvar, og henviser blant annet til selskapers ansvar for retten til privatliv og ytringsfrihet.[19]

Teknologiselskapenes store makt ovenfor enkeltindividene gjør at lovgivere har en særlig interesse av og behov for å underlegge teknologiselskaper direkte menneskerettighetsforpliktelser.[20] FNs spesialrapportør for ytringsfrihet, David Kaye, har for eksempel fokusert på teknologiselskapers plikt til å fremme digital ytringsfrihet og hindre begrenset tilgang til nettet.[21] I norsk sammenheng har vi nylig fått en åpenhetslov som pålegger større virksomheter å gjøre aktsomhetsvurderinger for menneskerettigheter i verdikjeder.[22] Selskapenes plikter knytter seg til grunnleggende menneskerettigheter, dermed også retten til privatliv, ytringsfrihet og diskrimineringsvernet, noe som etter omstendighetene vil kunne få følger for teknologiselskapene.

Fremmede stater som inngrepsaktører

Moderne teknologi opphever i noen grad geografiens begrensninger. Cyberspace er ikke knyttet til fysiske territorier, men er desentralisert, og dominert av ikke-statlige aktører. Teknologiutviklingen eksponerer derfor mennesker i Norge i økende grad for inngrep fra utenlandske staters myndigheter som kan bryte med menneskerettighetene.

En utfordring er den digitale infrastrukturen. Norske myndigheters forvaltning av Norge er avhengig av digital infrastruktur i utlandet, som dermed ligger under andre staters jurisdiksjon. Rundt 75 prosent av skytjenestene som benyttes av det offentlige tilbys av kommersielle aktører, og store mengder data lagres i skytjenester i utlandet. Andre stater kan

[19] FNs veiledende prinsipper for næringsliv og menneskerettigheter (UNGP).

[20] FNs spesialrapportør for ytringsfrihet, David Kaye, *Report on the freedom of expression in the digital era*, 2017, avsn. 82–83, FNs høykommissær for menneskerettigheter, *Right to privacy*, 2018, avsn. 43–49.

[21] FNs spesialrapportør for ytringsfrihet, David Kaye, *Report on the freedom of expression*, 2017.

[22] Lov om virksomheters åpenhet og arbeid med grunnleggende menneskerettigheter og anstendige arbeidsforhold (åpenhetsloven), som trer i kraft 1. juli 2022.

dermed få myndighet over data samlet inn av norske myndigheter om norske borgere i Norge. For å sikre «digital autonomi» har for eksempel Tyskland etablert *Bundescloud*, en offentlig skytjeneste som gir tyske myndigheter eksklusiv kontroll og jurisdiksjon over tyske data. Spørsmålet er om og i så fall hvordan myndighetene plikter å beskytte lagrede data om egne borgere mot andre stater. I Shrems-II dommen gikk EU-domstolen langt i å forby lagring i USA fordi det kan gi utenforstående (selskaper eller stater) tilgang til data på måter som bryter med EU-rettens regler om blant annet personvern.[23]

En annen utfordring handler om fremmede staters aktiviteter i Norge og hvilket menneskerettighetsansvar en stat har for digitale operasjoner utenfor egne landegrenser. EMD har ikke tatt stilling til såkalte ekstraterritorielle menneskerettighetsforpliktelser for digital virksomhet.[24] I 2020 konkluderte imidlertid Den tyske konstitusjonsdomstolen med at tysk utenlandsetterretning er underlagt grunnlovfestede forpliktelser til å respektere menneskerettigheter også når de overvåker utlendinger i utlandet. Dette er en rettsforståelse som ikke deles av de statene i verden med mest offensive digitale kapasiteter, som Kina, Russland og USA. Det er altså uenighet om hvordan menneskerettighetene gjelder når utenlandske statsmakter opererer digitalt i Norge på måter som kan innebære inngrep i menneskerettighetene her i landet.

Digital teknologi kan manipuleres

En offentlig forvaltning som i økende grad belager seg på digitale løsninger i sitt forhold til borgerne, må være oppmerksom på at data, prosesser og systemer kan manipuleres av inntrengere. Det kan dreie seg om hackere som driver med rampestreker, inntrengere som saboterer av personlige, politiske eller økonomiske årsaker, eller manipuleringen kan foretas av en terror-gruppe eller en fremmed statsmakt.

Manglende oversikt over teknologiens effekter for menneskerettighetene

Det er en økende bekymring for at de omfattende effektene som digitalisering og ny teknologi har for menneskerettighetene ikke blir analysert og debattert i tilstrekkelig grad. FNs høykommissær for menneskerettigheter anbefalte i 2021 at stater ikke bør innføre kunstig intelligens-teknologi på områder hvor det er «uklart hvilke effekter dette har for menneskerettighetene».[25] En dansk rapport fra 2021 som kartlegger rettighetsutfordringer knyttet til bruk av profileringsteknologi i offentlig forvaltning, konkluderte med at en rekke tiltak bør settes i verk for å sikre at kunstig intelligens og profilering som benyttes av offentlige myndigheter foregår på en betryggende måte fra et menneskerettighetsperspektiv.[26]

EMD har klargjort at stater som ligger langt fremme i utvikling og bruk av ny teknologi bærer et spesielt ansvar for å finne en god balanse mellom retten til privatliv og andre

[23] Shrems-II, EU-domstolen, storkammeravgjørelse 16. juli 2020.

[24] EMDs storkammerdommer fra 2021, *Big Brother Watch and Others v. Storbritannia* (58170/13, 62322/14 og 24960/15) og *Centrum för Rättvisa v. Sverige* (35252/08), 25. mai 2021, valgte å ikke berøre det kontroversielle spørsmålet.

[25] FNs høykommissær for menneskerettigheter, *Right to privacy in the digital era, 2021*.

[26] Dansk institut for menneskerettigheder. Når algoritmer sagsbehandler – rettighteder og retssikkerhed i offentlige myndigheders brug af profileringsmodeller, 11. okt. 2021.

menneskerettigheter.[27] Dersom styresmaktene skal kunne ivareta sine teknologinøytrale menneskerettighetsforpliktelser, er det en forutsetning at myndighetene i tilstrekkelig grad skaffer til veie kunnskap om hvilke mulige negative konsekvenser nye teknologiløsninger kan få for menneskerettighetene.

[27] *S. og Marper v. Storbritannia* (30562/04 og 30566/04), 4. des. 2008, avsn. 112.

5. Prosessuelle prinsipper for ny teknologi

Ny teknologi gjør at spørsmålet om hvorvidt et inngrep er i tråd med menneskerettighetene både vil dreie seg om *hva* inngrepet går ut på og *hvordan* beslutningen om inngrepet har blitt fattet.

Når styresmaktene utøver myndighet, setter menneskerettighetene grenser for hvilke vedtak offentlige organer kan treffe, med andre ord grenser for innholdet i beslutningene. Ved utstrakt bruk av ny teknologi flyttes fokus for menneskerettighetene i større grad over på *måten* styresmaktene kommet frem til beslutningene, altså saksbehandlingen eller det prosessuelle. Ny teknologi skyver i større grad menneskerettighetsfokuset over på prosessene: Hvordan kommer myndighetene frem til innholdet i et vedtak, og på hvilke måter kan dette etterprøves? Dersom det foreligger et statlig inngrep fra myndighetene som bryter med menneskerettighetene, vil spørsmålet både bli *hvorfor* (materielt spørsmål), men også *hvordan* maskinene har kommet frem til at dette inngrepet skulle foretas (prosessuelt spørsmål).

> «Statlig bruk av kunstig intelligens må sikre rettferdighet, ansvarlighet, etterprøvbarhet, samt innsyn og kontroll.»
>
> *FNs generalsekretær, 2020*

Menneskerettighetskonvensjoner har få regler som omhandler prosessuelle garantier for offentlige myndigheters saksbehandling. Menneskerettighetene forutsetter imidlertid at inngrep i menneskerettighetene har et rettslig grunnlag, som hjemmel i lov, og baserer seg dermed på et legalitetsprinsipp. Under EMK har krav til saksbehandling blitt utviklet gjennom domstolpraksis.[28] EMD legger vekt på om myndighetene har tilbudt individer tilstrekkelig deltakelse (medvirkning) under saksbehandlingen og gitt dem rimelig mulighet til å ivareta sine interesser. Domstolen legger særlig vekt på hvilke prosedyrer myndighetene har for opplysning av saken, kontradiksjon og mulighet for å få overprøvd beslutninger.

Når myndighetene innfører og benytter ny teknologi, oppstår spørsmålet om hvilke krav som må stilles for å sikre de menneskerettslige kravene til saksbehandling. Det er særlig fire prinsipper som er sentrale for å verne om og sikre menneskerettighetene ved bruk av ny teknologi.

Ansvarlighet

Digitale løsninger må anvendes på ansvarlig måte.[29] Man snakker gjerne om algoritmisk

[28] F.eks. *T.P. og K.M. v. Storbritannia* (28945/95), 10. mai 2001, avsn. 72.

[29] GDPR artikkel 5 (2), EUs *erklæring* om *digitale rettigheter og prinsipper for det digitale tiår*, 2022. FNs høynivåpanel for digitalt samarbeid, UNESCOs erklæring 2019/2021, OSCE-prinsippene fra 2019.

ansvarlighet eller algoritmeintegritet. Noen styresmakter har lagt seg på en linje hvor teknologi som kunstig intelligens kun skal brukes *hvor det gir verdi for borgerne og alltid på en etisk ansvarlig måte*.[30] EU-erklæringen om digitale rettigheter og prinsipper fra 2022 sier at «teknologi bør tjene og være til nytte for alle europeere [...] med full sikkerhet og respekt for deres grunnleggende rettigheter».[31]

Staten er forpliktet til å unngå eller avslutte bruk av teknologiske løsninger som bryter med menneskerettighetene. Fra et menneskerettslig ståsted innebærer ansvarlighet at ved innføring av ny teknologi eller bruk av teknologi på nye måter, må staten sørge for å kartlegge de menneskerettslige konsekvensene av slik teknologibruk. Når statlige myndigheter beslutter å innføre ny teknologi eller å endre bruken til eksisterende teknologi som kan innebære inngrep i menneskerettighetene, innebærer legalitetsprinsippet at det må foreligge hjemmel for beslutningen. Lovprosesser er gjerne grundig opplyst, og gir statlige myndigheter god tid til å kartlegge hvilke effekter en planlagt teknologi har for menneskerettighetene. Likevel fremgår ikke plikten til å utrede forholdet til menneskerettighetene eksplisitt av dagens utredningsinstruks.

Plikten til å kartlegge og å skaffe kunnskap omfatter både tilsiktede og utilsiktede effekter for menneskerettighetene. Kartleggingsplikten vil omfatte hele tidsperioden fra teknologien besluttes benyttet til bruken har opphørt. Myndighetene plikter også å iverksette tiltak for å avbøte konsekvenser som kan være i strid med menneskerettighetene, eller å unnlate å innføre teknologiske løsninger dersom effektene for menneskerettighetene er store eller ikke kan bøtes på. Alt dette forutsetter imidlertid at det foreligger tydelige krav til utredninger som er egnet til å sikre at menneskerettighetene ivaretas når myndighetene innfører tiltak eller systemer som bygger på ny teknologi.

Ansvarlighet innebærer krav til gode konsekvensutredninger om menneskerettslige aspekter fra myndighetenes side. I tillegg fordrer det at det brede sivilsamfunn, domstoler, ombud og tilsynsmyndigheter får tilstrekkelig kunnskap og kompetanse til å fungere som den motmakten de er forutsatt å være. Ansvarlighet medfører derfor også at dimensjoneringen av «motmakten» må reflektere den makten og det misbrukspotensialet som ny teknologi innebærer.

Beskyttelse mot vilkårlighet

En forutsetning for ansvarlig bruk av ny teknologi er at det lar seg gjøre å unngå vilkårlige beslutninger. Digital teknologi gir økt treffsikkerhet i offentlig styring og forvaltning, og er derfor et godt redskap for å øke forutsigbarhet og redusere uønsket vilkårlighet. Riktige beslutninger avhenger imidlertid både av kvaliteten på de digitale analyseverktøyene og på informasjonen som verktøyene benytter. Systemer som gir beslutningsstøtte eller som selv treffer beslutninger kan ha svakheter som innebærer at offentlig myndighetsutøvelse får uønskede eller utilsiktede følger, og resulterer i vilkårlig forskjellsbehandling eller usaklig diskriminering som er vanskelig å oppdage fra utsiden av de digitale systemene.

Datakvaliteten må være inkluderende og dekkende. Dersom informasjonen som benyttes

[30] Digitalstyrelsen, Danmark. Kommuner og regioner skal afprøve kunstigintelligens for at løfte kvaliteten i den offentlige service. 2019.

[31] EUs erklæring om digitale rettigheter, 2022, kap. 1.

av kunstig intelligens til å analysere eller lære, er ufullstendig, ubalansert eller på andre måter utilstrekkelig, kan dette produsere beslutninger som er urettferdige eller vilkårlige.[32]

Datakvaliteten kan svikte i mange ulike ledd. For eksempel ved:

- innledende beskrivelse av et problem (f.eks. sosioøkonomiske eller etniske karakteristikker)
- innsamling av data (f.eks. utsatt gruppe underrepresentert i treningsdataen)
- behandling av dataen
- type algoritmer
- skjevheter og bias
- mangel på vedlikehold og jevnlig justering av algoritmene

Algoritmestyrt forvaltning kan ha uønskede effekter som innebærer at ulikhet mellom folk og grupper vokser, og at utsatte grupper rammes særlig. Grunnen til dette er at algoritmer i større grad enn andre typer data kan gi skadelige «feedback loops» som øker forskjellene fordi de skjer innenfor algoritmens automatiseringssyklus, og ikke blir gjenstand for oversyn. Jo mer avanserte analyseverktøyene blir, desto vanskeligere blir det å få innsyn og kontroll som trengs for å sikre seg mot slik vilkårlighet.

Innsyn og kontroll

En forutsetning for ansvarlig bruk av ny teknologi er at systemene som benyttes er åpne og kan forstås av mennesker. De digitale verktøyene må tilfredsstille krav til innsyn og kontroll (*transparency*).[33] Det innebærer at digitale løsninger må sikre muligheter for innsikt i måten et system fungerer på, «resonnerer» og treffer beslutninger. Her kan vi skille mellom to ulike typer utfordringer. På den ene siden er algoritmer kompliserte og vanskelige å forklare i menneskelig forståelige termer. Fullt innsyn er teknologisk umulig i en del systemer. På den andre siden kan innsyn i algoritmene være begrenset på grunn av kommersielle hensyn, nasjonal sikkerhet eller personvernhensyn. For eksempel vil statlige myndigheter gjerne ikke ha full innsynsrett i beslutningsprosesser og softwaredesign som utvikles på kommersiell basis. Det er et myndighetsansvar å sørge for at teknologien som benyttes tilfredstiller krav til innsyn og kontroll. Ny personvernforordning i 2022 har gjort at Datatilsynet etablerer en oppdatert metodikk for algoritmetilsyn for å tilfredstille de nye kravene i GDPR.

Overprøvbarhet

En annen forutsetning for ansvarlig bruk av ny teknologi er at systemene som benyttes må være etterprøvbare.[34] Etterprøvbarhet og kontradiksjonsmulighet (*explainability* og *contestability*) innebærer at teknologien må kunne ettergås for å se hvordan en beslutning har blitt fattet. Dette er nødvendig for å kunne forstå vedtak, begrunne beslutningen, og gjøre det mulig å overprøve den. Den europeiske erklæringen om digitale rettigheter fra 2022 omtaler det som å «sørge for at algoritme-systemer er basert på passende datasett som [...] muliggjør menneskelig overvåkning av effektene som rammer mennesker». [35]

[32] Europaparlamentets personvernforordning (GDPR) 2016, art. 5.

[33] EUs erklæring om digitale rettigheter, 2022, kap. III; EUs ekspertpanel om KI, 2019; FNs høynivåpanel for digitalt samarbeid, UNESCO-erklæring 2019/2021, OSCE-prinsippene 2019. GDPR art. 5.

[34] EUs ekspertpanel om KI, 2019 og EUs erklæring om digitale rettigheter, 2022, FNs høynivåpanel for digitalt samarbeid, UNESCO-erklæring 2019/2021, OSCE-prinsippene 2019.

[35] EUs erklæring om digitale rettigheter, 2022, kap. III.

Retten til få overprøvd vedtak er en grunnforutsetning for utøvelse av menneskerettighetene. Det kan gjelde alt fra enkeltvedtak i forvaltningen til dommer i rettssystemet. Retten til en forklaring av automatiserte beslutninger er berørt i fortalen til EUs personvernforordning (GDPR). Overprøvbarhet innebærer at teknologien må være av en karakter som gjør at vedtak lar seg etterprøve. Det innebærer at det må være en reell mulighet til å begrunne vedtak og for å utøve retten til kontradiksjon

OVERSIKT OVER IKKE BINDENDE MENNESKERETTSLIGE INSTRUMENTER FOR KUNSTIG INTELLIGENS

EUs ekspertpanel for kunstig intelligens (KI) kom i 2019 med en rapport som anbefalte at syv elementer må besørges for at utvikling, oppsett og bruk av KI skal skje på en tillitsvekkende måte:

(1) menneskelig inngripen og oversikt
(2) teknisk robusthet og sikkerhet
(3) personvern og forsvarlig data forvaltning
(4) innsyn og kontroll
(5) mangfold, ikke-diskriminering og rimelighet
(6) miljø og sosial velferd
(7) ansvarlighet

I februar 2022 vedtok **EU verdens første erklæring om digitale rettigheter og prinsipper**.[36] Den gir anvisning på prinsipper som vil være sentrale når myndighetene tar i bruk ny teknologi, og har som formål å «styrke den menneskelige dimensjonen av det digitale økosystemet». Erklæringen er ikke rettslig bindende.

Det finnes etter hvert mange **etiske kodekser for kunstig intelligens**.[37] Felles er at de identifiserer prinsipper som en eller flere institusjoner forplikter seg til å ivareta ved utvikling eller bruk av kunstig intelligens.[38] Disse kodeksene forholder seg gjerne til vage begreper som *rettferdighet*, *menneskeorienterte verdier*, *innsyn og kontroll*, *overprøvbarhet*, *robusthet* og *ansvarliggjøring*.[39] Garantier for konkrete menneskerettigheter er tvert imot gjerne fraværende i slike dokumenter. Av G20-landenes prinsipper for kunstig intelligens fremgår det at «KI-aktører bør respektere rettsstatsprinsipper, menneskerettigheter og demokratiske verdier gjennom hele KI-systemets livssyklus».[40]

[36] Erklæringen er ikke direkte rettslig bindende for Norge, men vil påvirke tolkningen av regler som er rettslig bindende for oss gjennom EØS-avtalen.
[37] En oversikt over etiske regelverk for KI ved viktige institusjoner finnes i Fjeld, J. m.fl, *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*, Berkman Klein Center Research Publication nr. 1, 14. feb. 2020.
[38] En gjennomgang av fordeler og ulemper med KI i etiske koder og menneskerettslige rammeverk, se Business for Social Responsibility and World Economic Forum, *Responsible Use of* Technology, aug. 2019, s. 7 (med synspunktet at etikk og menneskerettigheter bør vær «synergetiske»).
[39] G20 Ministerial Statement on Trade and Digital Economy, (G20 prinsippene om KI), 2019.
[40] G20-prinsippene 12 (a).

6. Materielle menneskerettigheter i den digitale æraen

I møte med ny teknologi er enkelte menneskerettigheter særlig viktige.

Ny teknologi gir helt nye muligheter for hvilke mengder av data som kan samles inn, hvor lenge dataene om oss kan lagres, og hva de kan brukes til. Datasyklusen består av en rekke ledd, fra innhenting, lagring, systematisering, bearbeidelse, analyse, til langtidslagring og sletting. Menneskerettighetene som da blir særlig viktige er de materielle rettighetene som samtidig gir prosessuell beskyttelse: de som verner informasjon om hver av oss *gjennom hele syklusen* fra dataene samles inn frem til de slettes. Retten til privatliv gir slikt særlig vern ved innsamling, lagring og bruk av private data. Retten til ytringsfrihet gir beskyttelse av private uttrykk fra det øyeblikk de er ytret, også i den forstand at ytringer ikke skal påvirke ens øvrige rettigheter. Retten til ikke-diskriminering gir beskyttelse i den forstand at data ikke må samles inn, behandles eller benyttes på måter som forsterker eksisterende diskriminerende praksis eller foranlediger nye typer diskriminerende praksis. Etter hvert som digitale løsninger integreres dypere i hverdagen vår, blir retten til privatliv, ytringsfrihet og ikke-diskriminering stadig viktigere som forutsetninger for oppfyllelse av andre menneskerettigheter.

> «Retten til privatliv er en inngangsport til å fremme og verne andre menneskerettigheter som påvirkes av kunstig intelligens.»
>
> *FNs menneskerettighetsråd, 2019*

Retten til privatliv

Teknologiutviklingen medfører at retten til privatliv blir stadig viktigere fordi dataene om hver borger i større grad blir grunnlag for både rettigheter og for inngrep.[41]

I en informasjonsbasert verden blir retten til privatliv og til personlige data en grunnrettighet som tilrettelegger for andre menneskerettigheter, både online og offline.[42] Fra å være en alminnelig menneskerettighet som ga beskyttelse mot enkelte typer inngrep, er retten til privatliv blitt en grunnleggende rettighet som er en forutsetning for oppfyllelsen av andre rettigheter, på linje med retten til ytringsfrihet eller retten til ikke-diskriminering. Retten til privatliv spiller i økende grad en avgjørende rolle i maktbalansen mellom myndighetene og individet, og er en grunnleggende menneskerettighet i et demokratisk samfunn.[43]

EMK artikkel 8 beskytter privat informasjon i den forstand at staten som utgangspunkt

[41] FNs barnekomité, *General Comment No. 25*, 2021, avsn. 67–68; og A/HRC/39/29, avsn. 11; samtA/HRC/48/31 avsn. 6.
[42] Se FNs barnekomité, *General Comment No. 25*, 2021, avsn. 67–68, og A/HRC/39/29, avsn. 11.
[43] FNs høykommisær for menneskerettigheter, *Right to privacy*, 2021, avsn. 6.

verken skal ha eller gripe inn i privat informasjon. Privat informasjon inkluderer informasjon som finnes eller kan bli avledet om en person og vedkommendes liv, samt beslutninger fattet basert på slik informasjon. Digitaliseringen av livene våre gjør imidlertid at grensen mellom hva som er privat informasjon og hva som er offentlig er i endring, noe som gjerne kalles det private paradoks. Når skillet viskes ut mellom det personlige og det upersonlige, privat og offentlig data, anonym informasjon og informasjon hvor opphavskilde kan identifiseres, og denne sammenblandingen samtidig tjener interessene til mektige kommersielle aktører, blir det mer krevende å avgjøre hvilken type informasjon som er beskyttet av retten til privatliv.

RETTEN TIL PRIVATLIV

- Retten til privatliv tar utgangspunkt i menneskets verdighet og er knyttet til beskyttelse av individets autonomi og identitet.
- Retten til privatliv er vernet av Grl. § 102, EMK art. 8 og SP art. 17.
- Høyesterett har slått fast at Grl. § 102 og EMK art. 8 har sammenfallende innhold.
- Under SP og EMK omfatter retten til privatliv også retten til privat informasjon.

Lagring av informasjon kan utgjøre et brudd med retten til privatliv under EMK.[44] Når det gjelder innsamling og prosessering av kommunikasjonsdata, har EMD klargjort at slik data er beskyttet av retten til privatliv også når det ikke anses som personlig informasjon.[45] Retten til privatliv omfatter også rett til konfidensialitet i kommunikasjon.[46] For eksempel retten til å sende en e-post til en annen uten at kommunikasjonen fanges opp av statlige myndigheter. Om e-posten inneholder sensitiv informasjon er underordnet.[47] Den skal ikke leses av andre mennesker. Teknologiutviklingen gjør at spørsmålet endrer karakter. Dersom det er maskiner som leser e-posten, oppstår spørsmålet om retten til privatliv brytes i samme grad. Det er uenighet om retten til privatliv slår inn dersom den private informasjonen behandles av maskiner gjennom algoritmeanalyse. Svaret vil trolig avhenge av formålet med maskinens undersøkelse. Dersom det er for å sikre nettverkets operasjoner, vil det enten ikke implisere retten til privatliv eller det vil være et inngrep som vil være rettferdiggjort. Der det benyttes filtre som vil innebære at utvalgte e-poster til slutt vil bli undersøkt av mennesker, kan saken likevel stille seg annerledes.[48]

Dersom staten foretar inngrep i retten til privatliv, krever EMK at tre vilkår er oppfylt: Inngrepet må ha et rettslig grunnlag, det må ivareta et legitimt formål, og det må være nødvendig i et demokratisk samfunn.

Et inngrep vil foreligge dersom innhenting av informasjon (data eller metadata) er systematisk, informasjonen lagres over tid og utleveres til andre. Behandling av opplysningene vil kunne utgjøre et inngrep i privatlivet etter EMK artikkel 8, selv om informasjonen er offentlig tilgjengelig. EMD legger til grunn at masseinnsamling er en

[44] Se f.eks. *Leander v. Sweden*, (9248/81), 26. mars 1987, avsn. 48. Se også FNs høykommisær for menneskerettigheter, *Right to privacy, 2021, avsn. 20.*

[45] *Big Brother Watch and Others v. Storbritannia avsn*. 330 og *Centrum för Rättvisa v. Sverige* avsn. 244.

[46] FNs høykommisær for menneskerettigheter, *Right to privacy*, 2021, avsn. 17; Tallin Manualen 2017, s. 189.

[47] Den europeiske domstolen. *Digital Rights Ireland and Seitlinger and Others*, forente saker C-293/12 og C-594/12, 2014, 8. apr. 2014, avsn. 33.

[48] *Big Brother Watch and Others v. Storbritannia*, avsn. 330.

gradvis prosess, der inngrepet i individets rettigheter etter artikkel 8 tiltar underveis i prosessen.[49] Selv på det første stadiet, der innhentingen og lagringen ikke er rettet mot konkrete individer, vil dette kunne utgjøre et inngrep i artikkel 8. Kravet til sikkerhetsmekanismer vil imidlertid være størst på slutten av prosessen, når innholdet i kommunikasjonen kan bli nærmere undersøkt.[50] Under EMK vil altså det å samle inn data i store mengder (bulk) i seg selv være en innblanding i retten til privatliv.[51] Slik innsamling er likevel ikke et brudd med retten til privatliv dersom den tredelte testen er innfridd.[52]

Til illustrasjon vil et overvåkningsregime som samler inn digital informasjon i bulk (kommunikasjonsdata) være et inngrep i konvensjonenes forstand og må underkastes den tredelte testen.[53] Det første spørsmålet blir om det finnes en rettslig hjemmel under nasjonal rett som er tilstrekkelig til å beskytte mot vilkårlige inngrep i individers menneskerettigheter. Det andre spørsmålet blir om overvåkningen har et legitimt formål. Det tredje spørsmålet blir om overvåkningen er nødvendig i et demokratisk samfunn, i betydningen å svare på et presserende sosialt behov, og om inngrepet er proporsjonalt med det legitime formålet.[54] Overvåkning vil gjerne anses å være legitimt i den grad formålet er å beskytte nasjonal sikkerhet eller offentlig ro og orden.[55] Vurderingen av hva som er nødvendig i et demokratisk samfunn krever da at man vurderer den potensielle nytteverdien av teknologien sett opp mot alvorligheten i inngrepet i en gitt menneskerettighet. Når nytteverdien handler om nasjonale sikkerhetsbehov, er det gjerne svært krevende å ha en åpen og innsiktsfull vurdering av nytten.

EMDS KRAV TIL HJEMMEL FOR BULKOVERVÅKNING

EMDs storkammerdommer *Big Brother Watch and Others v. Storbritannia* og *Centrum för Rättvisa v. Sverige* har knesatt åtte krav for at bulkovervåkning skal være i tråd med EMK. Hjemmelen for bulkovervåkningen må inneholde

1. klar formulering av grunnlagene for bulkovervåkning
2. presisering av forhold som kan lede til at en persons kommunikasjonsdata overvåkes
3. fremgangsmåten for å gi tillatelse til å hente ut en persons kommunikasjonsdata
4. fremgangsmåten for å velge ut, undersøke og bruke slik kommunikasjonsdata
5. forholdsregler som skal tas dersom materialet kan deles med andre
6. klare begrensninger for varighet av overvåkningen og lagring av materialet, samt presisering av forhold som gjør at materialet må slettes og ødelegges
7. regler om uavhengig instans med oversyn og om dens kompetanse ved brudd
8. regler om uavhengig etterfølgende overprøvelse og rettsfølger ved brudd

[49] *Big Brother Watch and Others v. Storbritannia,* avsn. 325–331.
[50] *Big Brother Watch and Others v. Storbritannia*, avsn. 330.
[51] *Barbulescu v. Romania*, (61496/08), 5. sept. 2017.
[52] *Big Brother Watch and Others v. Storbritannia* og *Centrum för Rättvisa v. Sverige*. Se også *Telegraaf Media Nederland Landelijke Media BV and Others v. Nederland,* (39315/06), 22. nov. 2012, avsn. 88, og *Tele2 Sverige AB v Post-och telestyrelsen Secretary of State for the Home Department v Watson and Others,* 2016.
[53] *S. and Marper v. Storbritannia,* avsn. 101.
[54] Testen under FNs konvensjon om politiske og sivile rettigheter referer til nødvendighet og proposjonalitet. Vurderingen av nødvendighet forholder seg ikke alltid til «demokratisk samfunn», se f.eks. *General Comment* No. 34, art. 19.
[55] *Weber and Saravia v. Tyskland,* (54934/00), 29. juni 2006, avsn. 103–104.

I tillegg til EMK, gjelder også EUs personvernregler og SPs regler om privatliv for norske myndigheter. EU-pakten for grunnleggende rettigheter opererer med en rettighet til beskyttelse av persondata som er skilt ut fra retten til privatliv.[56] Pakten gjelder ikke direkte for Norge, men vil påvirke tolkningen av reglene for persondatabeskyttelse i EU, som Norge er bundet av gjennom EUs personvernforordning (GDPR). Lagring av data utgjør «prosessering av personlig data» og faller under GDPR. Forordningen gjelder som norsk lov med enkelte tilpasninger, jf. personopplysningsloven § 1. I den grad en stat lagrer persondata, vil dette implisere retten til privatliv.[57]

Etter GDPR artikkel 5 (2) må teknologi ha innebygget personvern. GDPR artikkel 35 pålegger personvernkonsekvensvurdering i forhold til de registrertes rettigheter og friheter ved ny teknologi eller teknologi som brukes i nye sammenhenger. De fleste land, inkludert Norge, har en MÅ-gjøre-en-personvernkonsekvensvurderings-liste etter artikkel 35 (4). Dertil sier GDPR artikkel 25 at den som tar i bruk et KI-verktøy der personopplysninger behandles, plikter å se til at de velger verktøy som har innebygd personvern i seg. Det innebærer at programmet eller algoritmen skal ivareta personvernprinsippene (artikkel 5), de registrertes rettigheter (artiklene 12–22) og friheter (fortalen (4) og EMK 8).

Sensitive personopplysninger er gitt et særlig vern i GDPR artikkel 9. Hovedregelen er at behandling av slike data er forbudt, men det er unntak ved samtykke, dersom den registrerte selv har offentliggjort opplysningene, eller dersom det er nødvendig av hensyn til allmenne interesser. I sistnevnte tilfelle må behandlingen av personopplysningene stå i et rimelig forhold til det mål som søkes oppnådd, være forenlig med det grunnleggende innholdet i retten til vern av personopplysninger, samt at egnede og særlige tiltak for å verne den registrertes grunnleggende rettigheter må sikres.[58]

Under SP kreves det at inngrep i retten til privatliv ikke må være vilkårlig eller ulovlig.[59] Ulovlig betyr at stater kun kan gripe inn i retten til privatliv i den grad det foreligger hjemmel. Hjemmelen må være i overenstemmelse med bestemmelsene og formålene med SP, den må være klar og den må spesifisere vilkårene for inngrep i detalj. Inngrep i retten til privatliv kan ikke være vilkårlig, noe som innebærer at selv inngrep som er foreskrevet i lov, må være i tråd med SP, tjene et legitimt formål, være nødvendig for å oppnå det legitime formålet,[60] og være proporsjonal.[61] Enhver begrensning av retten til privatliv må være det minst inngripende alternativ, og må ikke røre ved kjernen i retten til privatliv. [62]

---

[56] Se Pakten for grunnleggende rettigheter i Europa Unionen, som gir rett til respekt for kommunikasjon i art. 7 og rett til beskyttelse av personlig data i art. 8.

[57] Europarådet, direktiv 95/46/EC (1995), art. 2(b).

[58] GDPR art. 9 nr. 2 bokstav g.

[59] For forskjell i terminologien «vilkårlig» og «ulovlig» FNs høykommissær for menneskerettigheter, *Report on the right to privacy in the digital era,* 2014, avsn. 21–27.

[60] FNs høykommissær for menneskerettigheter, *Right to privacy,* 2014, avsn. 21 flg, og 2018, avsn.10.

[61] *Toonen v. Australia* (CCPR/C/50/D/488/1992), avsn. 8.3, *Van Hulst v. Netherlands* (CCPR/C/82/D/903/1999), avsn. 7.3 og 7.6, *Madhewoo v. Mauritius* (CCPR/C/131/D/3163/2018), avsn. 7.5, and CCPR/C/USA/CO/4, avsn. 22.

[62] FNs menneskerettighetskomite, *General Ccmment* No. 31 (2004), avsn. 6; FNs høykommissær for menneskerettigheter, *Right to privacy,* 2014, avsn. 22, and FNs høykommissær for menneskerettigheter, *Right to privacy,* 2018, avsn. 10.

Ytringsfrihet

Stadig flere av våre ytringer foregår digitalt. Nye digitale verktøy gjør at retten til ytringsfrihet må sikres på nye måter.

Ytringsfriheten er beskyttet av EMK artikkel 10, SP artikkel 19 og Grunnloven § 100. Ytringsfriheten skal blant annet verne om sannhetssøken, demokrati og individets frie meningsdannelse.[63] Ytringsfriheten omfatter både retten til å søke og motta informasjon, og til å dele ideer og meninger, og kan derfor sees på som en forutsetning for realiseringen av andre menneskerettigheter. Ytringsfrihet i Grunnloven og internasjonale traktater gjelder for ethvert medium, og omfatter også moderne teknologiske løsninger.[64] Beslektet med ytringsfriheten er retten til menings- og tankefrihet.[65] I det øyeblikket en mening er uttrykt, vil utrykket være beskyttet av og begrenset av ytringsfriheten.[66] Ytringsfriheten er en forutsetning for forsamlings- og demonstrasjonsfriheten, som vernes av EMK artikkel 11, SP artikkel 21 og Grunnloven § 101.

Ny teknologi øker flatene hvor man kan utøve sin ytringsfrihet. Digitale flater gir et globalt publikum mulighet til å søke, motta og dele informasjon, ideer og andre typer innhold for å oppnå kunnskap, bidra i debatter og delta i demokratiske prosesser. Det er også enklere å nå et større publikum med ytringer gjennom digitale flater, hvor enhver kan være sin egen redaktør. Ved at flere kan få tilgang til og dele større mengder informasjon raskere og enklere enn før, kan ny teknologi spille en positiv rolle for ytringsfriheten.

Samtidig representerer ny teknologi en trussel mot retten til å motta informasjon og retten til å ytre seg. Denne trusselen retter seg mot individer, grupper og hele samfunnssegmenter. Algoritmer dikterer hva brukere ser og hvilken prioritet dette gis, og de kan programmeres til å begrense eller prioritere innhold som igjen påvirker og manipulerer mottaker. De mest sentrale digitale plattformene eies av noen få, internasjonale selskaper, som har stor påvirkningskraft, er lite regulert, og som handler som portvoktere for den digitale verden.

Noen snakker om et tiltagende digitalt forurensningsmiljø i kommunikasjonsmiljøet, med falske profiler, «deep fakes», manipulasjon og ugjennomtrengelige avsendere.[67] Roboter sprer informasjon og desinformasjon raskt og i store mengder. Kunstig intelligens kan spisse politiske budskap rettet mot enkelte samfunnssegment med bestemte holdninger eller bekymringer eller i gitte livssituasjoner. Vår digitale offentlighet styres i økende grad av algoritmer på måter som forsterker motsetninger og polarisering, som regel uten at vi er klar over det selv. Den digitale informasjonsverdenen er desentralisert og fragmentert, og store internasjonale selskaper forvalter stadig mer av de sentrale ytringsflatene. Utviklingen går så raskt at rettslig rammeverk og reguleringstiltak for å beskytte brukere mot nye verktøy stadig havner på etterskudd.

Ny teknologi kan dermed utfordre selve formålet med ytringsfriheten – å skape et klima for en mangfoldig offentlig debatt som er tilgjengelig og inkluderende for alle. Brudd på retten til privatliv og selvbestemmelsesrett over

[63] Grunnloven § 100 annet ledd.

[64] Tallinn Manual 2017 s. 189.

[65] SP art. 19 (2), General Comment No. 34.

[66] EMK art. 10 (1), FN-erklæringen art. 19, SP art. 19 (2).

[67] Dansk institut for menneskerettigheder, Når algoritmer sagsbehandler, (2021).

informasjon om en selv kan hindre mennesker fra å uttrykke seg fritt. Algoritmer kan fragmentere det offentlige rom og skape ekkokamre, hvor enkelte typer nyheter blir forsterket, og fører til økt polarisering i samfunnet, som igjen kan true sosiale bånd. Individets frie meningsdannelse kan bli truet dersom vi blir presentert kun med et utsnitt av virkeligheten, samtidig som vi selv tror at vi blir presentert et helhetlig bilde. Nye teknologiske løsninger kan bryte med retten til ytringsfrihet i en slik grad at den underminerer individuelle rettigheter og den måten deltakende demokrati er ment å fungere.[68]

Overvåkningsmulighetene som ny teknologi bringer med seg, utfordrer særlig ytringsfriheten gjennom nedkjølingseffekten. Teknologien gir mulighet til å fange opp og kartlegge hvem man kommuniserer med (kommunikasjonsdata) og hva som blir delt (innholdsdata). Vi legger igjen stadig flere digitale spor om oss selv, som kan lagres, systematiseres og sammenstilles. Ytringsfriheten og personvernet vil ofte kunne stå i et spenningsforhold til hverandre, for eksempel der en persons ytring griper inn i privatlivet til en annen. Ytringsfriheten og personvernet er gjerne begge involvert i de menneskerettslige avveiningene som må foretas i tilknytning til ny teknologi. Frykt for negative virkninger på privatlivet og personvernet kan medføre at man lar være å ytre seg, eller ytrer seg på en annen måte enn man ellers ville ha gjort. Dette gjelder særlig ved frykt for myndighetenes overvåking, men også private selskapers lagring og håndtering av data kan sette ytringsfriheten under press når vi ikke vet hvilken informasjon som samles inn, hvordan den kan systematiseres eller hvem den kan gis ut til. Den nedkjølende effekten på ytringsfriheten er spesielt alvorlig for journalister, som nyter et sterkt vern av sine kilder og sin korrespondanse. EMD har konstatert krenkelse av ytringsfriheten i EMK artikkel 10 der bulkovervåkingssystemer ikke inneholder gode nok sikkerhetsmekanismer for å beskytte konfidensiell journalistisk informasjon.[69] I andre tilfeller vil påvirkningen på ytringsfriheten kunne være mindre konkret og vanskelig å påvise, slik at det blir snakk om et strukturelle ytringsfrihetsspørsmål snarere enn krenkelser begått overfor enkeltindivider.

Måten ny teknologi kan spre informasjon på, gjør at uønskede ytringer slik som desinformasjon, hatefulle ytringer eller annet ulovlig innhold, kan få større spredning og større negativ effekt enn tidligere. Dette eksponerer særlig minoriteter og marginaliserte grupper. Private plattformer har gjerne egne regler for moderering av innhold, og baserer seg ofte på hel- eller delautomatiserte systemer for å avdekke og blokkere uønsket innhold. Slike verktøy kan være med på å beskytte mot ulovlig og skadelig innhold. En for restriktiv modereringspraksis kan derimot også være problematisk i lys av ytringsfriheten. Det kan å ha motsatt effekt, ved at innhold publisert av personer som tilhører gruppen man i utgangspunktet ønsker å beskytte mot hatefulle ytringer, blir moderert bort.

I møte med ulovlig eller skadelig innhold kan også stater ty til verktøy som å blokkere enkelte individer fra tilgang til utvalgte IP-adresser, ta ned nettsider, benytte filterteknologi for å nekte tilgang til sider som inneholder bestemt innhold, blokkere kommunikasjon av e-poster og andre typer inngrep i ytringsfriheten. Slike virkemidler kan lett misbrukes. Inngrep av denne typen vil være i strid med ytringsfriheten etter EMK dersom de ikke har hjemmel i lov og

[68] *Szabo and Vissy v. Ungarn*, (37138/14), 12 jan. 2016, avsn. 57.
[69] *Big Brother Watch and Others v. Storbritannia*.

er nødvendige og proporsjonale.[70] Enkelte typer ytringer har dessuten et særlig sterkt menneskerettighetsvern. Det gjelder ytringer som angår myndighetenes maktbruk, politikk, demokratiske valg, samt rapportering av menneskerettigheter, myndighetsaktiviteter og korrupsjon.

Ikke-diskriminering

Digitale verktøy og kunstig intelligens kan både forsterke eksisterende diskriminerende tendenser og kan skape helt nye former for diskriminering som bryter med menneskerettighetsvernet mot diskriminering.

Diskrimineringsvernet står sterkt i Norge. Grunnloven § 98 og EMK artikkel 14 slår fast forbudet mot diskriminering. EMK oppstiller et aksessorisk diskrimineringsvern som innebærer at inngrep i rettigheter under EMK ikke kan praktiseres på en måte som gjør forskjell i konvensjonens rettigheter basert på rase, hudfarge, kjønn, språk, trosoppfatning, politiske eller andre meninger, nasjonalt eller sosialt opphav, eiendomsforhold, fødsel eller annen status. SP oppstiller et generelt diskrimineringsvern basert på de samme grunnlagene.[71]

Ny teknologi eksponerer oss for at allerede eksisterende diskriminerende praksiser overføres til det digitale domenet. Et eksempel er ansiktsgjenkjenningsverktøy som i liten grad klarer å gjenkjenne mørkere hudfarge. Algoritmene til maskinene lages per i dag av mennesker, og maskinene kan arve våre stereotypier og vår forutinntatthet. Diskrimineringspraksis smitter over til de digitale flatene, og kan forsterkes som følge av menneskelig programmering.

Et annet diskrimineringsaspekt ved ny teknologi er *tilgang* på nye digitale flater. Digitalt utenforskap betegner personer som av ulike grunner ikke har like god tilgang på ny teknologi som andre. Dette kan for eksempel gjelde personer med funksjonsnedsettelser, personer uten fast bopel eller personer med livsførsel som avviker fra normen. Overføring av stadig flere av samfunnets offentlige funksjoner fra analoge til digitale flater kan gjøre at informasjon og funksjoner som blir mer tilgjengelig for de fleste, blir mindre tilgjengelige for et mindretall. Digitalt utenforskap kan reflektere og forsterke allerede eksisterende sosiale, kulturelle og økonomiske ulikheter. Manglende tilgang på teknologi kan også skape nye former for utenforskap som er problematisk fra et diskrimineringsperspektiv.

I tillegg kan maskinlæring og kunstig intelligens medføre helt nye former for diskriminering. Maskinene kombinerer etter hvert informasjon på nye måter. Livssyklusen til dataene består av generering, lagring, kuratering, prosessering og gjenbruk. Enkelte skillelinjer som kjønn, legning, etnisitet og religion, er karakteristika som for en rekke beslutninger ikke skal tillegges vekt etter menneskerettighetene, for eksempel på arbeids- eller boligmarkedet. Selv om vi bevisst skjuler slike opplysninger for maskinen, gjør summen av andre opplysninger at selvlærende maskiner likevel kan slutte seg frem til disse kjennetegnene. Algoritmer har *annen* informasjon som vil bli benyttet som for eksempel kjønnspredikator. Dersom dataene tilsier at det er en viktig faktor, vil autonome systemer dermed kunne ende med å vektlegge

[70] EMD har eksempelvis konstatert krenkelse av art. 10 i saker om blokkering av nettsteder uten tilstrekkelig begrunnelse og overprøvingsmulighet, se *Ahmet Yildirim v. Tyrkia* (3111/10), 8. des. 2012. Heller ikke blokkering av YouTube var i tråd med EMK art. 10, se *Cengiz and Others v. Tyrkia* (48226/10 og 14027/11), 1. des. 2005.

[71] SP art. 2(1), 3 og 26.

faktorer på måter som vil bryte med menneskerettighetene som beskytter mot nettopp slik form for diskriminering. Et analyseverktøy benyttet av Amazon til å plukke ut relevante CVer har for eksempel gjort at kvinner i stor grad har blitt avvist. Maskinenes veier til beslutninger blir utilgjengelige, samtidig som resultatet blir diskriminerende.

7. Menneskerettslig SciFi: Nye menneskerettigheter for ny teknologi?

Ny teknologi gjør mennesket om fra subjekt til objekt i samfunnet. Utviklingen gir grunn til å tenke nytt rundt menneskerettighetene.

Utfordringene for menneskerettighetene i møte med ny teknologi som er skissert opp i denne rapporten gjør at stadig flere uttrykker bekymring for at den tradisjonelle tolkningen av menneskerettighetene er «håpløst utilstrekkelig» eller metodisk svak. Både internasjonalt og nasjonalt blir det etterspurt klarere rammer og retningslinjer.[72] Det tales derfor stadig oftere om nye typer rettigheter for å fange opp nye former for overgrep, uintenderte effekter og negative konsekvenser av ny teknologi. Man ser for seg både utvidende eller tilpassende tolkninger av allerede eksisterende menneskerettigheter og helt nye regler. FNs spesialrapportør for ytringsfrihet har for eksempel oppfordret til at det utvikles et rettslig rammeverk som kan gi bedre tilpassede normer for beskyttelse av privatlivet i den digitale tidsalderen.

> «Det er simplistisk og naivt å slå seg til ro med at 'det som er beskyttet offline er beskyttet online'.»
>
> *FNs spesialrapportør for ytringsfrihet, 2018*

Nye rettigheter som kommer med denne utviklingen omfatter retten til å bli glemt,[73] retten til å bli informert og retten til å ikke følge ordre eller være ulydig. Dette er alle rettigheter som anses nødvendige i ethvert solid rettssystem, altså et rettssystem som vil være effektivt i møte med algoritmisk myndighetsutøvelse og styring.

En rett til ikke å bli underlagt automatiserte beslutninger er også i støpeskjeen.[74] GDPR gir en registrert person rett til ikke å være gjenstand for en avgjørelse utelukkende basert på automatisert behandling når avgjørelsen i betydelig grad påvirker vedkommende.

En beslektet rettighet er retten til menneskelig behandling, altså retten til kontakt med mennesker i møte med offentlige tjenester. I Frankrike arbeider man for eksempel med utkast til ny lov som skal garantere menneskelig kontakt med helsepersonell selv om selve behandlingen er basert på teknologiske løsninger, altså en slags rett til menneskelig kontakt i beslutningsprosessen. EU-parlamentet la til grunn i 2017 at menneskelig kontakt utgjør en grunnkomponent for personlig helse.[75]

---

[72] Dansk institut for menneskerettigheder (2021), Når algoritmer sagsbehandler, s. 4.

[73] Europeisk erklæring om digitale rettigheter og prinsipper, kap. V: «Everyone should be able to determine their digital legacy, and decide what happens with the publicly available information that concerns them, after their death», 2022.

[74] Den er allerede delvis reflektert i GDPR art. 22.

[75] Résolution du Parlement européen contenant des recommandations à la commission concernant des règles de droit civil sur la robotique. 16. feb. 2017; No. 32.

En annen rett i støpeskjeen handler om rett til å beskyttes mot at algoritmer og kunstig intelligens pre-determinerer folks valg i privatlivet, som valg av helse, utdannelse eller arbeid.[76]

En mer revolusjonerende tilnærming til digitale rettigheter argumenterer for at teknologien i økende grad vil overta den rolle som rettsregimer har hatt, fordi digitale strukturer oppfyller den samme sosiale regulerende funksjonen som jussen har, bare raskere og mer effektivt. Fordi dataprosessering former samfunnet vil datasamling bli et nytt instrument for sosiale betingelser som igjen vil påvirke hvordan enkeltmennesker opptrer og handler. Bekymringen handler om «dyp teknologi», altså ubevisste påvirkningsmuligheter.

Noen foreslår å løse dette gjennom en egen type digitale menneskerettigheter hvor man skiller mellom menneskerettigheter offline og online. For sine online rettigheter, trenger man en online *personae,* altså en digital eller virtuell representasjon som kan eksistere og ha rettigheter uavhengig av rettssubjektet som skapte dem. Dette kalles en virtuell tvilling. Begrepet tar utgangspunkt i at alle har en digital identitet, en personlig digital tvilling. Så mye data samles nå om hver enkelt av oss at vi kan etter hvert kan tale om vår egen personlige virtuelle tvilling. Konseptet befinner seg mellom menneske og ting, og diskusjonen om hvordan menneskerettighetene skal tilpasses vår virtuelle tvilling tiltar raskt. For eksempel har noen foreslått å inkludere en ny rettighet i FN-erklæringen av 1948 for å gi hvert individs digitale tvilling tilsvarende rettigheter som individet har i den fysiske verden.[77]

Ny teknologi vil hjelpe samfunnet vårt med store fremskritt og forbedringer de neste årene. Myndighetene vil bli bedre i stand til å sikre menneskerettighetene gjennom disse teknologiske nyvinningene. Teknologien har imidlertid noen mørke baksider. I påvente av klarere internasjonale rettsregler for digital styring, må myndighetene ta konsekvensene av at ny teknologi og kunstig intelligens også kan utgjøre en risiko for samfunnet vårt og ha negative effekter for menneskerettighetene. Noen av disse negative effektene er det vanskelig å forutse, identifisere eller måle, som for eksempel konsekvenser for demokratiet, rettsstaten, rettferdig omfordeling eller virkninger på det menneskelige sinnet. Likefullt er det myndighetenes ansvar å sørge for at ny teknologi blir mer til nytte enn til skade for det menneskerettslige vern. Myndighetene må derfor kartlegge mulige negative effekter og iverksette nødvendige og proporsjonale tiltak for å redusere denne risikoen.

Når offentlige myndigheter anvender ny teknologi til beslutningsstøtte eller til å fatte beslutninger, er det myndighetenes ansvar at menneskerettighetene respekteres og vernes. Ansvaret gjelder både ved beslutninger om å innføre og modifisere digitale løsninger, og ved bruk av slik teknologi i statlig forvaltning.

[76] Europeisk erklæring 2022, kap. II (Valgfrihet).

[77] Change.org. A new UN Human Rights article that protects your Personal Digital Twin. Dette vil neppe føre frem.

8. NIMs anbefalinger

Anbefaling 1:

Stortinget bør be regjeringen om at lovforslag som innebærer økt bruk av informasjonsteknologi, følges av grundige konsekvensanalyser som vektlegger menneskerettighetene, særlig retten til privatliv, ikke-diskriminering og ytringsfrihet.

- **Tilsiktede effekter:** Konsekvensanalysene bør kartlegge tilsiktede effekter av teknologien som kan berøre menneskerettighetene.
- **Utilsiktede effekter:** Det må også undersøkes om utilsiktede, men beregnelige effekter av teknologien kan berøre menneskerettighetene.
- **Tverrgående effekter:** Kartlegging bør søke å fange opp og adressere tverrgående menneskerettslige utfordringer.
- **Diskriminering:** Analysen må identifisere klare rettssikkerhetsgarantier, og sørge for at rase, farge, kjønn, religion, politisk overbevisning, nasjonal eller sosial opprinnelse, eiendomsforhold, fødsel eller annen status ikke gir grunnlag for diskriminerende praksiser.
- **Inkludering:** Mangfold, inkludering og vern av utsatte grupper må tas i betraktning for hele datasyklusen. Rettighetene til mennesker med funksjonsnedsettelser i forhold til digitale løsninger må ivaretas.
- **Periodevis vurdering:** Analyser bør gjøres tidlig i utviklingsfasen og periodevis etter at teknologien tas i bruk.

Anbefaling 2:

Ved innføring av helautonome beslutningssystemer må myndighetene foreta en *særlig* grundig analyse som kartlegger konsekvensene for menneskerettighetene. I slike tilfeller bør konsekvensanalysen som et minimum inneholde følgende elementer:

- **Begrunnelse:** Når myndighetene beslutter å benytte digitale verktøy til beslutningsstøtte eller til autonome beslutningssystemer oppstår en begrunnelsesplikt. Begrunnelsen må inneholde hjemmel og formål, samt egnethet til å oppnå formålet. Den bør også beskrive alternative løsninger.

- **Algoritmisk innsyn og kontroll:** Systemet må tilfredsstille krav om innsyn og kontroll. Det må være klart hvordan man sikrer datakvalitet, hvordan oversyn og kontroll kan gjøres på en tilfredsstillende måte, samt hvordan data som går ut på dato ikke legges til grunn.

- **Overprøving:** Systemet må være konstruert på en måte som gjør det mulig å etterprøve beslutningen, samt å utøve kontradiksjon. Når profileringsmodeller benyttes til beslutninger, må det være vilkår om at modellene er i stand til å produsere konkrete begrunnelser. Det samme vil gjelde profileringsmodeller til beslutningsstøtte.

- **Kartlegging av risiko for menneskerettighetene:** Myndighetene må kartlegge risikoen for menneskerettighetene, herunder risiko for vilkårlige beslutninger, diskriminerende beslutninger, om teknologien tilfredstiller GDPRs vilkår om personvern, dataminimering eller om den baserer seg på korrekte data. Kartleggingen bør beskrive hvilke tiltak som kan iverksettes for å begrense slik risiko. Jo mer inngripende tiltak autonome beslutningssystemer muliggjør, desto strengere vilkår stilles til kartlegging av de menneskerettslige konsekvensene og tiltak for å bøte på dette.

- **Deltakelse for utsatte grupper:** Statlige myndigheter har et særlig ansvar for å sørge for etterlevelse av konvensjonsforpliktelsene *under* saksbehandlingen. Ved bruk av profileringsmodeller for saksbehandling, må man sikre reelle muligheter for at de berørte kan bli hørt. Ved helautomatiserte prosesser, må det klargjøres hvordan disse rammer sårbare grupper og hvilke tiltak som kan virke avbøtende. Ved bruk av beslutningsstøtte, må man sikre at modellen ikke får en uforholdsmessig stor innflytelse og blir beslutningsstyrende istedenfor beslutningstøttende.

Litteraturliste

Acharya, S.,«Tackling Bias in Machine Learning», *Medium: Insight*, 2019.

Buiten, M. C., «Towards Intelligent Regulation of Artificial Intelligence», *European Journal of Risk Regulation*, vol. 10, nr. 1, 2019.

Burgess, M. «UK police are using AI to inform custodial deci- sions—but it could be discriminating against the poor», *Wired Magaazine*, 2. feb. 2018, hentet fra http://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit

Crawford, K. «The hidden biases in big data», *Harvard Business Review*, 2013.

Den internasjonale telekommunikasjonsenheten. *Digital Inclusion for All*, nov. 2019.

Europarådet, *Recommendation of the Committee of Ministers to member States on the protection of human rights with regard to search engines*, CM/Rec(2012)3, 4. apr. 2012, avsn. 1, hentet fra https://wcd.coe.int/ViewDoc.jsp?id=1929429

Europarådet menneskerettighetskommisjonær, «Democratic and Effective Oversight of National Security Services», mai 2015, s. 19–27.

Davidson, T. m. fl., «Racial Bias in Hate Speech and Abusive Language Detection Datasets», 2019, hentet fra https://aclanthology.org/W19-3504.pdf

Fabbrini, F., Celeste, E., «The right to be forgotten in the digital age: the challenges of data protection beyond borders», *German Law Journal*, avsn. 21 og 55-65.

Hao, K., «This Is How AI Bias Really Happens – and Why It's So Hard to Fix», *MIT Technology Review*, 2019.

Jimenez, J. I., Jahankhani, H., Kendzierskyj, S. «Health care in the cyberspace: Medical cyber-physical system and digital twin challenges», *Digital twin technologies and smart cities. Internet of things (technology, communications and computing)*, red. Farsi, M. m. fl., 2020, Cham, Switzerland: Springer.

Joyce, D. «Privacy in the Digital Era: Human Rights Online?», *Melbourne Journal of International Law,* 2015, s. 270 og 273.

Jobin, A., Ienca, M., Vayena, E., «The Global Landscape of AI Ethics Guidelines», *Nature Machine Intelligence*, vol. 1, nr. 9, 2019.

Kluttz D., Kohli N., Mulligan D. «Shaping our tools: contestability as a means to promote responsible algorithmic decision making in the professions», *After the digital tornado: networks, algorithms, humanity*, red. Werbach, K., s. 137-152, Cambridge, Storbritannia: Cambridge University Press.

Kochi, E. «How to Prevent Discriminatory Outcomes in Machine Learning», World Economic Forum 2018.

Latonero, M. «Governing Artificial Intelligence: Upholding Human Rights and Dignity», Report, Data & Society 2018.

O'Reilly, T., «The great question of the 21st century: Whose black box do you trust?», 13. sept. 2016.

Penney, J. m. fl., «Advancing Human Rights-by-Design in the Dual-Use Technology Industry», *Journal of International Affairs*, 2018, s. 103.

SAS Institute. «Artificial Intelligence – What It Is and Why It Matters,» *SAS: The Power to Know*, hentet 10. apr. 2019 fra https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html

Samuel, Arthur, «Some Studies in Machine Learning Using the Game of Checkers,» *IBM Journal of Research and Development 3:3*, juli 1959.

Shany, Y. Contribution to Open Consultation on UN GGE 2015 Norm Proposals, 2018.

Saxena, R. «The social media "echo chamber" is real», *Arstechnica*, hentet 25. sept. 2017 fra https://arstechnica.com/science/2017/03/the-social-media-echo-chamber-is-real/

Sharkey, N. «The Impact of Gender and Race Bias in AI», *ICRC Humanitarian Law and Policy Blog*, 28. aug. 2018, hentet fra https://blogs.icrc.org/law-and-policy/2018/08/28/impact-gender-race-bias-ai/

Stanford University. «A Peek Inside the 'Black Box' of Machine Learning Systems», *Artificial Intelligence Research*, 2017.

Sullivan C. «Digital identity – a new legal concept», *Digital identity: an emergent legal concept*, s. 19-40. Adelaide, Australia: The University of Adelaide Press.

Ribeiro, T. «OSCE Representative on Freedom of the Media», hentet fra https://www.osce.org/files/f/documents/8/f/510332_0.pdf

Ronen, Y., «Big Brothers Little Helpers; The Right to Privacy and the Responsibility of the Internet Service Providers», *Utrecht Journal of International and European Law*, 2015 s. 72.

Rudin, C., Radin, J. «Why Are We Using Black Box Models in AI When We Don't Need To?», *Harvard Data Science Review*, vol. 1, nr. 2, 2019.

Teknologirådet. «Elections, Technology and Political Influencing», 2019. https://teknologiradet.no/wp-content/uploads/sites/105/2019/06/Elections-technology-and-political-influencing.pdf

Teller, M. «Legal Aspects related to digital twin», *Philosophical Transactions of the Royal Society A*, 16. aug. 2021, hentet fra https://royalsocietypublishing.org/doi/10.1098/rsta.2021.0023

USAs forsvarsdepartements strategi om kunstig intelligens, *Harnessing AI to Advance Our Security and Prosperity*, 2018.

Wright, J., Verity, A., «Artificial Intelligence Principles for Vulnerable Populations in Humanitarian Contexts», *Digital Humanitarian Network*, jan. 2020, s. 15.

Norges institusjon for menneskerettigheter

Publisert 28. mars 2022