

HVA VET VI OM...?

Kunstig intelligens og likestilling

En kartlegging av norsk forskning



UTGITT AV: Kilden kjønnsforskning.no

Adresse: Postboks 564, 1327 Lysaker

Ansvarlig: Linda Marie Rustad

Forfattere: Trine Rogg Korsvik, Marie Hulthin og Anne Sæbø

ISBN: 978-82-12-03870-7

PUBLISERT: OKTOBER 2020

LAYOUT: KILDEN KJØNNSFORSKNING.NO

ALLE FOTO: ISTOCKPHOTO

TITTEL:

Hva vet vi om kunstig intelligens og likestilling?
En kartlegging av norsk forskning

**Kilden kjønnsforskning.no er et nasjonalt kunnskapssenter
for kjønnsperspektiver og kjønnsbalanse i forskning.**

Forord

Målet med denne kunnskapsoversikten er å kartlegge og belyse hva norsk forskning sier om likestillingskonsekvenser av at kunstig intelligens tas i bruk på stadig flere samfunnsområder. Kartleggingen viser at dette fortsatt er et utforsket tema i Norge.

Kunnskapsoversikten er utarbeidet av Kilden kjønnsforskning.no, og er finansiert av Likestillings- og diskrimineringsombudet (LDO) og Kilden. Gjennomføringen av litteratursøket i kunnskapsoversikten er utført i samarbeid med Universitetsbiblioteket ved Universitetet i Oslo. Anne Sæbø fra Universitetsbiblioteket har hatt ansvar for kapittel 3 om metode. Marie Hulthin og Trine Rogg Korsvik fra Kilden har skrevet de øvrige kapitlene i rapporten.

Kilden ønsker å takke LDO for samarbeidet og Anne Sæbø ved Universitetsbiblioteket for stor hjelp til å systematisere litteratursøkene og samle inn forskning. En stor takk også til prosjektets faglige referansegruppe som har bistått både når det gjelder tips til problemstillinger, søkeord og relevant litteratur, samt kommentarer til rapportens kapitler underveis. Referansegruppas medlemmer har vært:

- Hege Kristin Andreassen, forsker og senterleder ved Senter for kvinne- og kjønnsforskning ved UiT Norges arktiske universitet
- Erling Barth, forsker ved Institutt for samfunnsforskning
- Morten Goodwin, professor ved Institutt for informasjons- og kommunikasjonsteknologi og nestleder ved Centre for Artificial Intelligence Research (CAIR), Universitetet i Agder
- Tereza Østbø Kuldova, forsker II ved Arbeidsforskningsinstituttet ved OsloMet
- Roger A. Søråa, forsker ved Institutt for tverrfaglige kulturstudier og Institutt for nevromedisin og bevegelsesvitenskap ved NTNU

Takk også til Taina Bucher, førsteamanuensis ved Institutt for medier og kommunikasjon, Universitetet i Oslo, og Kristine Ask, førsteamanuensis ved Institutt for tverrfaglige kulturstudier, NTNU, som har kommet med innspill til kapitlet om digitale sosiale plattformer og underholdning.

Lysaker, oktober 2020

Marie Hulthin og Trine Rogg Korsvik

INNHold

Hva vet vi om kunstig intelligens og likestilling? En kartlegging av norsk forskning

Forord	s 3
Sammendrag	s 5
Kapittel 1 Innledning	s 9
Kapittel 2 Hva er kunstig intelligens?	s 14
Kapittel 3 Metode og datagrunnlag	s 21
Kapittel 4 Tjenester	s 25
Kapittel 5 Arbeidsliv	s 31
Kapittel 6 Digitale sosiale plattformer og underholdning	s 40
Kapittel 7 Diskusjon og avslutning	s 47
Referanser	s 55
Vedlegg 1 Søkeord	s 61
Vedlegg 2 Søkelogg	s 64

Sammendrag

I Norge er det et uttalt politisk mål å øke bruken av kunstig intelligente systemer for å effektivisere og levere mer treffsikre og brukertilpassede tjenester. Nye digitale løsninger ses som en forutsetning for at Norge skal kunne opprettholde og utvikle konkurransekraft og velferdsnivå, og som en nøkkel for å kunne oppnå FNs bærekraftsmål. I januar 2020 la regjeringen fram *Nasjonal strategi for kunstig intelligens* (KMD 2020). Et av målene med strategien er å stimulere til offentlig debatt om etisk bruk av kunstig intelligens (KI). Denne rapporten er et bidrag til denne debatten.

Mange av oss er bevisst eller ubevisst daglig i befatning med KI-baserte verktøy. For eksempel når Facebooks algoritmer anbefaler oss nyheter, reklame og nye venner, når skatteetaten regner ut skatten vår, eller når Lånekassen og NAV behandler søknadene våre. I tillegg erfarer arbeidstakere at arbeidsoppgaver helt eller delvis tas over av robotteknologi og automatiserte løsninger.

Til tross for de store sosiale og økonomiske mulighetene KI medfører, har de «smarte» løsningene også vist seg å by på utfordringer. Det finnes flere eksempler på hvordan KI-teknologi har fått utilsiktede og diskriminerende utfall, for eksempel når det gjelder ansettelse eller beregninger av forsikringspremier. Nettopp derfor mener Likestillings- og diskrimineringsombudet (LDO) og Kilden kjønnsforskning.no at det er relevant å se hva forskningen sier om likestillingskonsekvenser av at KI-teknologi får en stadig større innvirkning på befolkningens hverdag, både når det gjelder offentlige og private tjenester, på jobben og på fritida.

Metode og litteratursøk

Kunnskapsoversikten kartlegger hva som finnes av forskning utført ved norske forskningsinstitusjoner på kunstig intelligens med kjønns- og likestillingsperspektiv. Forskning med kjønnsperspektiver er kartleggingsoversiktens hovedanliggende. Med kjønnsperspektiver mener vi både forskning der kjønn er en sentral del av forskningsdesign og -analyse, og forskning der kjønn kun er én av flere variabler.

Denne formen for systematisk kartleggingsoversikt (*scoping review*) er en litteraturoversikt som med utgangspunkt i et definert spørsmål kartlegger hva som finnes og hva som kjennetegner forskningslitteraturen om temaet. Prosessen går ut på å identifisere, utvelge, trekke ut data og beskrive forskningslitteraturen. I kartleggingsoversikten ser vi primært på empiriske studier som er publisert som vitenskapelige artikler i fagfelleverderte tidsskrifter, men har i visse tilfeller også beskrevet funn fra forskningsrapporter, såkalt grå litteratur.

Litteratursøket med relevante søkeord ble gjort i sentrale norske og internasjonale databaser. I tillegg har vi gjennom håndsøk funnet og supplert med relevante artikler og rapporter fra forskningsinstitusjoner. Det totale antall norske studier som presenteres som egne caser i denne rapporten er 24 (21 vitenskapelige artikler og 3 forskningsrapporter). Ikke alle disse har et tydelig kjønns- eller likestillingsperspektiv, men er likevel inkludert fordi de på mer implisitte måter belyser likestillingskonsekvenser av KI-teknologi. Det kan for eksempel være fordi studiene belyser sosial ulikhet i møte med myndighetene, eller fordi de undersøker endringer som følge av automatisering i arbeidslivet – et arbeidsliv som i Norge fortsatt er relativt kjønnsdelt.

Tematiske områder

Rapporten er strukturert etter tre tematiske områder der KI-teknologi er relevant i et likestillings- og diskrimineringsperspektiv: Tjenester, arbeidsliv og digitale sosiale plattformer og underholdning.

Tjenester

Foreløpig er det gjort relativt få empiriske studier i Norge av samfunnsmessige konsekvenser av at offentlige og private tjenester blir drevet av automatiske løsninger med elementer av KI. Enda færre empiriske studier anlegger et tydelig kjønns- og likestillingsperspektiv. Vi har likevel identifisert forskningsfunn som er relevante når en skal vurdere likestillingskonsekvenser av KI-teknologi i offentlige og private tjenester.

Ekskluderende roboter?

Samfunnsfaglig robotforskning viser hvordan roboter kan påvirke vårt syn på kjønn, og hvordan vårt syn på kjønn påvirker hvordan roboter blir utformet. Kulturen «kjønner» roboter, i den forstand at jo mer menneskelignende en robot er, jo mer tillegges den «maskuline» eller «feminine» egenskaper. Slik kan roboter bidra til å videreføre eller forsterke kjønnsstereotyper. Forskningen viser også hvordan robotteknologi som blir brukt i helsevesenet kan ekskludere grupper som faller utenfor normen for «standardmennesket».

Automatisert saksbehandling – mer likebehandling eller forskjellsbehandling?

Studier som undersøker sosiale effekter av automatisert saksbehandling, er relevante i en likestillings- og diskrimineringsammenheng. Ifølge forskning er det i norske kommuner lav kunnskap om at algoritmer kan videreføre skjevheter fra dataene de bygger på, for eksempel knyttet til kjønn, og som dermed kan forsterke eksisterende forskjellsbehandling. For mange NAV-brukere er digitale automatiske løsninger effektive og tidsbesparende. Andre brukere mestrer dem ikke, og risikerer å ikke få de ytelsene de etter loven har krav på. Forskningen påpeker at det er oppstått et nytt såkalt digitalt skille mellom NAV-brukere som mestrer og ikke mestrer de digitale løsningene. Det digitale skillet dreier seg i liten grad om kjønn, men om utdanningsnivå og alder. Eldre med lite utdanning er mest sårbare for rettighetstap som følge av digitale løsninger. Men også personer med høy digital kompetanse kan falle ut av systemet på grunn av psykiske problemer eller annen sykdom. NAVs mål om effektiv saksbehandling kan, ifølge forskningen, stå i fare for å gå på bekostning av grunnleggende rettigheter.

Forutseende analyser som verktøy i flyktningebosetting og kriminalitetsforebygging

I Norge pågår det en diskusjon om å ta i bruk algoritmiske verktøy i offentlig sektor. Prediktive (forutseende) analyser basert på historiske data kan brukes som redskap for å bosette flyktninger, men også til å forutsi om noen kan tenkes å utføre uønskede og kriminelle handlinger. På grunn av risikoen for urettferdig behandling, anbefaler forskere at myndighetene venter med å ta i bruk KI-systemer for å bosette flyktninger i norske kommuner. I stedet anbefales det fleksible modeller for bosetting av flyktninger som tar hensyn til alle familiemedlemmers behov, og som ivaretar både barneperspektivet og kjønnsperspektivet. Forskere advarer også mot å ta i bruk algoritmebaserte prediktive analyser for å forebygge kriminalitet fordi de kan true enkeltmenneskers rettssikkerhet.

Arbeidsliv

Kartleggingen viser at sosiale konsekvenser av KI i arbeidslivet, som også inkluderer likestillings- og kjønnsperspektiver, fortsatt er relativt utforsket i Norge. Imidlertid er det gjort relevante forskningsfunn når det gjelder omstilling i arbeidslivet som følge av automatisering, både når det gjelder hvilke yrkesgrupper som er mest og minst utsatt, og for medbestemmelse på arbeidsplassen. Selv om ikke kjønns- eller likestillingsperspektivet er sentralt i denne forskningen, er den relevant for å analysere hvordan KI og automatisering slår ut i typiske kvinne- eller mannsdominerte yrker.

Algoritmestyrte rekruttering

Internasjonal forskning har påvist diskriminerende effekter av KI-systemer som peker ut de beste jobbsøkerkandidatene basert på historisk skjeve data. Algoritmene har vist seg å videreføre fordommer når det gjelder kjønn og hudfarge. Vår kartlegging tyder på at det ennå ikke er publisert vitenskapelige artikler som utforsker mulige diskriminerende effekter av rekrutteringsalgoritmer når det gjelder norske forhold.

Polarisering av arbeidslivet

Forskning har vist at automatisering kan føre til en polarisering i arbeidslivet ved at det blir større økonomiske forskjeller. Allerede nå er det blitt færre mellomlønte ansatte og flere høytlønte og lavtlønte i Norge. Den såkalte plattformøkonomien der privatpersoner kan tilby tjenester på digitale plattformer gjør at nye serviceyrker oppstår. Plattformøkonomien har innvirkning på ansettelsesforhold ved at det blir færre faste ansettelser, flere frilansere, selvstendig næringsdrivende og løsarbeidere (også kalt prekariatet), og at fagorganiseringen dermed svekkes. Usikre arbeidsforhold og ekskludering av lavt utdannede og personer med dårlig helse fra arbeidsmarkedet, kan skape større sosial ulikhet, som igjen kan ha negativ innvirkning for folkehelsen og trolig også for likestillingen.

Yrkesgruppers utsatthet for automatisering

Når det gjelder spørsmålet om hvilke yrkesgrupper som er mest eller minst utsatt for automatisering og hva dette har å si for likestillingen, må kjønn ses i sammenheng med variabler som utdanning, kompetanse og alder. Typisk kvinnedominerte yrker i helse-, omsorgs- og utdanningssektoren er blant de minst automatiserbare, og for dem kan robotteknologi og KI fungere som nyttige hjelpemidler i jobben. Samtidig har mange av arbeidsoppgavene til kontormedarbeidere, et annet typisk kvinnedominert yrke, allerede blitt automatisert bort.

Høyere utdanning, der kvinner er i flertall, framheves i forskningen som et vern mot å miste jobben som følge av automatisering. Men forskningen tyder på at uansett kjønn vil yrker som fordrer kompleks problemløsning og sosial samhandling, ikke kunne erstattes av KI og robotteknologi med det første.

I landbruket har innføringen av KI-teknologi ført til større fleksibilitet i bondens arbeidshverdag. Men melkeroboter, som blir et stadig vanligere redskap blant melkebønder, har så langt ikke ført til mer likestilling på gårdene. Uansett om bøndene har melkerobot eller ikke, gjør kvinnelige bønder fortsatt mer husarbeid enn mannlige bønder.

Gir automatisering mer makt til sjefene?

Hvilke konsekvenser automatisering med KI-baserte systemer har for arbeidsmiljøet er fortsatt relativt utforsket. Forskningen vi har kartlagt som studerer dette, peker imidlertid i retning av at automatisering gir ledelsen mer enn de ansatte. I den høyteknologiske industrien ser det ut til at omstilling med digitale verktøy basert på Industri 4.0-konseptet foreløpig er mer tilpasset ledelsen enn arbeidernes behov. I mediebransjen har digitaliseringen ført til nedbemanning som kvinnelige journalister mer enn mannlige rapporterer at de er misfornøyde med. I NAV opplever veilederne at «skjermbyråkratiet» gir ledelsen økte muligheter for rigid detaljstyring. Den digitale detaljstyringen gjør at de ansatte mister kontroll over eget arbeid og muligheten til å utøve skjønn i saksbehandlingen. Forskere argumenterer for at digitale omstillinger må forankres blant de ansatte slik at digitale verktøy støtter dem og gir dem mer medbestemmelse og autonomi i arbeidet.

Manglende likestilling i IKT-sektoren

Ettersom KI-systemer blir stadig viktigere, øker behovet for flere fagfolk i IKT-sektoren, et fagfelt der kvinner utgjør et klart mindretall. For å øke IKT-kompetansen i Norge, også når det gjelder IKT-sikkerhet, hevder forskere at det er nødvendig å styrke IKT-utdanningen og å rekruttere flere kvinner til fagfeltet. Foreløpig ser det imidlertid ut til å være få konkrete likestillingstiltak i IKT-sektoren.

Digitale sosiale plattformer og underholdning

Også når det gjelder forskningen på bruken av kunstig intelligens i uformelle og hverdagslige sammenhenger, på digitale sosiale plattformer og i underholdningstjenester som for eksempel strømmetjenester, viser kartleggingen at kjønnspektiver og likestilling er relativt lite tematisert i norsk forskning. Samtidig viser kartleggingsoversikten at enkelte forskningsbidrag likevel kan belyse tematikken.

Lite kjennskap til hva algoritmer er

Forskning viser at store deler av Norges befolkning har svært liten kjennskap til algoritmer, og at dette har sammenheng med kjønn, alder, utdanningsnivå og hvor man bor. Kjønnforskjellene er store, idet menn langt oftere enn kvinner oppgir at de har kjennskap til og synspunkter på algoritmer. En stor andel kvinner, spesielt eldre og med lav utdanning, oppgir at de ikke har noen kjennskap til algoritmer overhodet. Menn med høy utdanning som er bosatt i urbane strøk oppgir oftest at de har god kjennskap til algoritmer. Ifølge forskerne kan manglende kunnskap om algoritmer være et demokratisk problem ettersom algoritmer stadig får større betydning for hva internettbrukere finner på nett, tilgang på informasjon og deltagelse på digitale plattformer.

Algoritmestyrt kuratering

KI og algoritmer har vekket forskningsmessig interesse også i medievitenskapen. Relevant i et likestillingsperspektiv er diskusjonen om hvordan algoritmiske anbefalingssystemer i strømmetjenester kan stå i fare for å standardisere og reprodusere konvensjonelle kjønnsrollemønstre. Forskning viser at KI-baserte anbefalingsfunksjoner i strømmetjenester kan føre til såkalte «kjønnsbobler» hvor systemet forsøker å predikere hva som er typisk jente- og guttemusikk.

Seksualitet, trakassering og kunstig intelligens

Den raske teknologiske utviklingen aktualiserer nye spørsmål om forholdet mellom menneske og maskin, inkludert forholdet mellom menneskelig seksualitet og ny teknologi. I forskningslitteraturen diskuteres etiske og psykologiske problemstillinger knyttet til kommersielle roboter som er designet for å tilfredsstille mennesker seksuelt og emosjonelt. Forskerne finner tydelige kjønnsforskjeller, hvor kvinner er mer skeptiske enn menn til roboter generelt, og til sexroboter spesielt.

En annen tilnærming til seksualitet og teknologi er å se på hvordan mennesker uttrykker sin seksualitet, sitt kjønn eller opplever uønsket seksuell oppmerksomhet på digitale sosiale plattformer. Forskning viser at å være en såkalt «jentegamer» blir negativt vurdert, og at kjønn i dataspill har en klar over- og underordning. Kjønsdynamikken som utspiller seg i dataspillverdenen, kan ifølge forskere trolig påvirke spillere også i det virkelige liv.

Forskningen viser at kjønnsbasert seksuell trakassering forekommer i digitale spill og at kvinner tar ulike forholdsregler for å unngå uønsket seksuell oppmerksomhet, blant annet ved å skjule at de er kvinner.

KAPITTEL 1

INNLEDNING



Kapittel 1

Innledning

Digitalisering og ny teknologi preger samfunnet i stadig større grad, og utviklingen går raskt. Nye digitale løsninger ses som en forutsetning for å kunne møte ulike samfunnsutfordringer, for å opprettholde dagens velferdsnivå og som en nøkkel for å kunne oppnå FNs bærekraftsmål. Blant løsningene som forventes å ha et stort potensial, er teknologi som baserer seg på kunstig intelligens (KI).

Lovnadene om hvilke muligheter som ligger i KI har gjort at kunstig intelligente løsninger høster allmenn, forskningsmessig, og politisk interesse. Å nyttiggjøre ulike teknologiske løsninger ses nå som en forutsetning for at Norge skal kunne opprettholde og utvikle konkurransekraft og velferd. I januar 2020 la regjeringen fram *Nasjonal strategi for kunstig intelligens* som skisserer hvordan Norge skal gå foran i utvikling og bruk av kunstig intelligens, med respekt for den enkeltes rettigheter og friheter (Kommunal- og moderniseringsdepartementet, KMD 2020).

Selv om KI innebærer store muligheter både for samfunnet og for enkeltpersoner, fører teknologien også med seg en rekke pro-

blemstillinger, blant annet i et kjønns- og likestillingsperspektiv. Formålet med denne kunnskapsoversikten er å kartlegge hva norsk forskning sier om likestillingskonsekvenser av at kunstig intelligente digitale systemer tas bruk på stadig flere områder i samfunnet.

Vi har valgt ut tre samfunnsområder der KI-teknologi blir stadig viktigere, og som utgjør hvert sitt kapittel i rapporten: offentlige og private tjenester, arbeidslivet og digitale sosiale plattformer og underholdning. Disse samfunnsområdene er særlig viktige i et likestillingsperspektiv. Hva har det for eksempel å si for likebehandlingsprinsippet at maskiner med innslag av KI tar beslutninger i saksbehandlingen uten at menneskelig skjønn er involvert? Hva har det å si for likestillingen på arbeidsmarkedet at mange arbeidsoppgaver blir overtatt av roboter? Og hva har det å si at kvinner og menn kan få anbefalt ulikt innhold på digitale plattformer som de bruker i hverdagen?

Kunstig intelligens, kjønnslikestilling og diskriminering

Til tross for mulighetene som ligger i KI-teknologi, har de smarte løsningene vist seg å innebære betydelige utfordringer, også i et kjønns- og likestillingsperspektiv. Det finnes flere eksempler på hvordan systemer basert på kunstig intelligens har fått både utilsiktede uheldige utfall, og også direkte diskriminerende konsekvenser. Et kjent eksempel er hvordan et rekrutteringsvektøy basert på KI-teknologi utviklet av det amerikanske selskapet Amazon, viste seg å diskriminere kvinnelige jobbsøkere (Dastin 2018). I et annet tilfelle førte KI-systemet til Apples kredittkort at kvinner fikk tilbud om lavere kreditttramme enn menn (Vigdor 2019). Flere internasjonale studier har dokumentert liknende tendenser. Studien «Gender Shades» fra MIT (Massachusetts Institute of Technology) har vist at ansiktsgjenkjenningssystemer fra noen av de største teknologiselskapene fungerte dårligere på å korrekt identifisere kvin-

ner enn menn, og aller dårligst på svarte kvinner (Buolamwini og Gebru 2018). En annen studie har vist at Googles algoritmer kun viste reklame for høyt betalte prestisjejobber til menn og ikke til kvinner (Datta, Tschantz og Datta 2015).

Forskere har påpekt at KI-teknologi kan bidra til å reproducere skadelige kjønnsstereotyper. Adams og Loideain (2019) hevder for eksempel at det at virtuelle assistenter som Apples Siri og Amazons Alexa er utformet med kvinnenavn og -stemmer bidrar til å reproducere forestillingen om at kvinner er underordnet menn, og satt til å være menns assistenter og hjelpere. Forskerne mener disse eksemplene på kulturell reproduksjon av maktubalanse er en form for indirekte diskriminering og et brudd på FNs kvinnekongresskonvensjon (CEDAW).

Også Europarådet har påpekt at KI og algoritmer kan reproducere og forsterke eksisterende stereotyper knyttet til kjønn og dermed bidra til å opprettholde sexistiske holdninger (Europarådet 2019: 18). Kjønnsdiskriminerende holdninger kan også gjøre seg utslag i ulike former for vold mot kvinner. Sexistiske kommentarer, kjønnsbaserte hatytringer og såkalte «deepfake» – falske videoer – forekommer stadig oftere på internett, og rammer kvinner og jenter hardest. Særlig «deepfake»-bilder og -video, basert på avanserte KI-programmer der ansikt og stemme til en person kan manipuleres inn i for eksempel pornografisk materiale, brukes stadig oftere som

en form for seksuell trakassering og overgrep mot kvinner (EU-kommisjonen 2020: 6; Harwell 2018).

Kjønns- og likestillingsperspektiv er ikke sentralt i regjeringens strategi for kunstig intelligens. I strategiens kapittel om ansvarlig og pålitelig kunstig intelligens er imidlertid ett av de sju etiske prinsippene for kunstig intelligens at «KI-systemer skal legge til rette for inkludering, mangfold og likebehandling». Her heter det «ved utvikling og bruk av kunstig intelligens, er det særlig viktig å være oppmerksom på at KI bidrar til inkludering og likestilling, og at diskriminering unngås» (KMD 2020: 59).

Problemstillinger, begrepsavklaringer og utvalg av litteratur

Eksemplene ovenfor aktualiserer flere spørsmål knyttet til likestilling og kunstig intelligens, og illustrerer behovet for en kunnskapsbasert tilnærming til hvilke positive og negative konsekvenser kunstig intelligens har i et kjønn- og likestillingsperspektiv. Målet med denne kunnskapsover-

sikten er å kartlegge hva vi vet om kjønnslikestilling og kunstig intelligens i Norge i dag. Står kunstig intelligente systemer i fare for å reproducere eksisterende kjønnsforskjeller? Eller gir KI-teknologien mulige verktøy for å oppnå større grad av likestilling og forhindre diskriminering?

Likestilling

Begrepet «likestilling» kan forstås som å innebære frihet fra diskriminering. Tradisjonelt har begrepet vært forbundet med like rettigheter og muligheter uavhengig av kjønn, men i dag handler det også om likestilling på bakgrunn av andre forhold, som for eksempel alder, etnisitet, religion, funksjonsevne, seksuell orientering, utdanningsnivå, sosioøkonomisk status eller klasse. Av hensyn til rapportens omfang har vi her tatt utgangspunkt i kjønnslikestilling. Der forsknin-

gen inkluderer andre forhold som er relevante, belyser vi også dette. I den grad forskningen også vektlegger et interseksjonelt perspektiv, hvordan ulike diskrimineringsgrunnlag som for eksempel kjønn og etnisitet kan virke sammen og forsterke hverandre, løfter vi også fram dette. Bruk av likestilling og likestillingskonsekvenser i denne kunnskapsoversikten henviser til at KI-systemer kan ha implikasjoner for rettigheter og muligheter for ulike samfunnsgrupper.

Kunstig intelligens

Kunstig intelligens (KI), på engelsk *artificial intelligence* (AI), refererer til informasjonsteknologi som kan justere sin egen aktivitet og derfor kan framstå som intelligent. Det inkluderer både *regelbaserte systemer* som er programmert på forhånd etter prinsippet «hvis x, så y», som brukes i en rekke digitale skjemaer, og maskinlæring. *Maskinlæring* er et sentralt område innenfor KI, og går ut på å gjøre

dataprogrammer i stand til å «lære» ut fra mønstre i store datamengder. Eksempler på bruk av maskinlæring er selvkjørende biler, bilde- og språkgjenkjenning, anbefalingsalgoritmer på nettet, og prediktive analysemodeller for å forutsi sannsynlighet for sykdommer eller uønsket atferd. I kapittel 2 forklares kunstig intelligens mer inngående.

Norsk forskning

Denne kunnskapsoversikten er en systematisk kartlegging av norsk forskning på kjønn, likestilling og kunstig intelligens. Norsk forskning definerer vi som vitenskapelige publikasjoner som er utgitt av forskere hvor minst én har tilknytning til en norsk institusjon. Bakgrunnen for dette er ønsket om å kartlegge hva som er status for forskningen på og kunnskapsgrunnlaget for mulige likestillingskonsekvenser av økt bruk av KI-systemer i Norge. Det

Utvalg av forskningslitteratur

I kartleggingen ser vi primært på empiriske studier som er publisert som vitenskapelige artikler i fagfelleverderte tidsskrifter. I visse tilfeller har vi også beskrevet funn fra relevante forskningsrapporter, såkalt grå litteratur som ikke er utgitt på vitenskapelige forlag.

Fordi KI er et område hvor den teknologiske utviklingen går svært raskt, kan forskningsresultatene fort bli utdatert. Aktører fra spill- og hackerkulturer, dataingeniører, programmerere og journalister er oftere enn forskere i front når det gjelder undersøkelser av hvordan KI-systemer og algoritmer kan virke (Kiberg 2019: 1). At det er få vitenskapelige artikler som omhandler kjønns- og likestillingsaspekter er en av grunnene til at vi har valgt å inkludere særlig relevant og forskningsbasert grå litteratur i kartleggingen. Når det gjelder tidsaspektet, har vi valgt å inkludere forskning som er publisert fra 2010 og fram til i dag. Vi anser dette som tilstrekkelig tidsperiode for å fange opp publikasjoner som kan anses for å være relevante i dag. De fleste forskningsartiklene vi har identifisert er publisert fra og med 2018 og framover.

Kunstig intelligens er primært et forskningstema innenfor informatikk og datateknologi, men fordi disse fagområdene sjeldent tar for seg kjønns- og likestillingsperspektiver, legger denne kartleggingsoversikten vekt på forskning innenfor samfunnsvitenskap, og til dels humaniora. Innenfor andre fagområder foreligger det trolig forskning på kunstig intelligente systemer som kan være relevant i et likestillingsperspektiv, men siden verken dette eller andre sosiale aspekter er tematisert eksplisitt, fanges ikke denne litteraturen opp av våre litteratursøk.

Flere av forskningsbidragene i kartleggingen har en kritisk inngang til studiene av samfunnsmessige konsekvenser av KI-teknologi, inkludert maskinlæringsalgoritmer og automatiserte beslutningssystemer. Internasjonalt har det

er svært få norske studier som omhandler kunstig intelligens i et kjønnslikestillingsperspektiv, og ettersom forskningsfronten er internasjonal, refererer vi også til sentrale forskningsbidrag fra andre land, uten at disse er del av den systematiske kartleggingen av forskningslitteraturen. Gjennomføringen av litteratursøket beskrives nærmere i kapittel 3 om metode.

utviklet seg forskningsmiljøer innenfor såkalte *digital divide studies* og *critical algorithm studies*.

Digital divide studies, eller studier av digitale (klasse-)skiller eller digital ulikhet, problematiserer hvordan tilgangen til nye digitale verktøy og evnen til å bruke dem varierer mellom ulike grupper i befolkningen (Gran et al. 2020; Lutz 2019; van Dijk 2006). Flere studier har funnet at digital ulikhet i stor grad speiler eksisterende sosial ulikhet som følge av sosioøkonomiske forskjeller i utdanning, kjønn, alder, bosted, tilknytning til arbeidslivet og etnisitet (Robinson et al. 2015).

På et annet og beslektet felt har forskere innen såkalte *kritiske algoritmestudier* påpekt et behov for mer kunnskap om algoritmer, fordi de påvirker hva vi som brukere møter på nett (Beer 2017; Diakopoulos 2015; Gran et al. 2020). Slike studier anvender et kritisk perspektiv på særlig maskinlæringsalgoritmenes «sorte boks», en metafor som viser til manglende innsyn i hvordan algoritmene fungerer og tar beslutninger. For eksempel hvordan algoritmene bidrar til å utvelge informasjon som anses for å være relevant for enkeltpersoner, og dermed setter rammene for informasjon, forbruk og deltagelse i det offentlige liv (Bucher 2018; Gillespie 2013).

Et annet perspektiv innenfor kritiske algoritmestudier studerer hvordan KI-systemer påvirker samfunnet på et mer overordnet nivå. For eksempel brukes begrepet *algorithmic governance* for å karakterisere en styringsmodell som er basert på automatisering, digitalisering og kvantifisering (Katzenbach og Ulbricht 2019). Norske arbeidslivsforskere studerer hvordan algoritmer i økende grad styrer beslutninger i arbeidslivet, og argumenterer for at standardiserte og automatiserte styringsstrategier svekker ansattes og mellomlederes medbestemmelse og dermed muligheten for å utøve faglig skjønn (Kuldova et al. 2020).

Internasjonalt er studier av kunstig intelligens i et kjønnsperspektiv ikke noe nytt fenomen. Flere forskere har kritisert KI-teknologien for å ha iboende skjevheter med hensyn til kjønn, og har introdusert «teknofeministiske» (*techno feminist*) tilnærminger til forholdet mellom kjønn

og teknologi (Adam 1998; 1995; Wajcman 2007). Forskere har i senere tid også argumentert for at man gjennom såkalt datafeminisme (*data feminism*) kan studere innebygde maktrelasjoner i ny KI-teknologi (D'Ignazio og Klein 2020).

Rapportens oppbygging

Rapporten består av i alt sju kapitler, inkludert dette innledningskapitlet. I neste kapittel forklarer vi hva kunstig intelligens er og presenterer sentrale begreper innenfor KI-teknologi, slik som «sterk» og «svak» kunstig intelligens, maskinlæring, regelbaserte systemer for automatisering, algoritmer og algoritmenes «sorte boks». I kapittel 2 diskuterer vi også hvorfor likestillingsperspektiver er relevante i studier av KI-teknologi. Kapittel 3 redegjør for den metoden vi har brukt i kartleggingen av forskningslitteraturen. Det gjelder både litteratursøk i norske og internasjonale databaser og håndسøk. Kapittel 4, 5 og 6 presenterer de norske empiriske studiene vi har kartlagt.

Kapittel 4 presenterer forskning som er relevant for å forstå hvordan innføringen av KI-teknologi for å effektivisere og levere mer treffsikre og brukertilpassende tjenester, kan ha konsekvenser for likestillingen. Temaer som tas opp er KI-teknologi i offentlig saksbehandling og automatiserte beslutningsprosesser i velferdstjenestene (NAV), fordomsfulle roboter, og forutseende (prediktive) analyser brukt som verktøy for bosetting av flyktninger og i kriminalitetsforebygging.

Kapittel 5 tar for seg studier av hvilke effekter KI-teknologi og automatisering av arbeidsoppga-

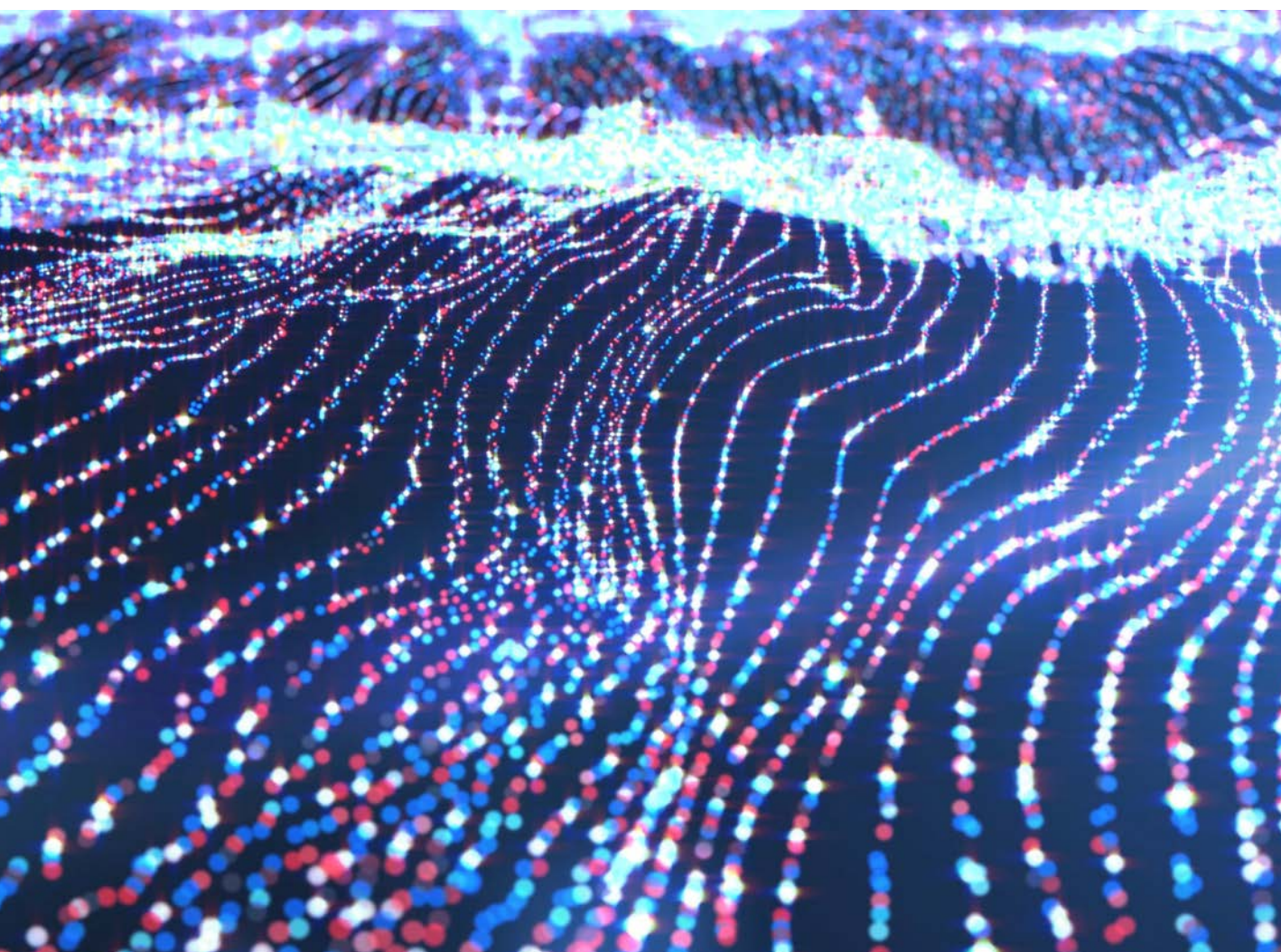
ver har for likestillingen i arbeidslivet. Et kjennetegn ved det norske arbeidslivet er at det fremdeles er relativt kjønnssegregert. I kapitlet presenteres forskning om polarisering og endringer i yrkessammensetning av arbeidslivet som følge av automatisering, og sosiale effekter av den såkalte plattformøkonomien (også kalt gig-økonomi eller delingsøkonomi). Manglende likestilling i IKT-sektoren og arbeidstakeres erfaringer med digital omstilling er andre temaer i dette kapitlet.

Kapittel 6 omhandler forskning om KI-teknologi i digitale sosiale plattformer og i underholdningstjenester mange benytter i hverdagen, og hvilken relevans det har for likestilling. Her presenteres studier av befolkningens kunnskaper og holdninger til algoritmer i hverdagen, hvordan algoritmer påvirker hva vi som brukere blir disponert for på internett, og hvordan det kan oppstå «kjønnsbobler» i en digital musikkultur. Andre tema som tas opp er sexroboter og seksuell trakassering i dataspill.

Kapittel 7 diskuterer hovedfunnene fra de norske empiriske studiene i kapittel 4, 5 og 6 i lys av sentrale politiske styringsdokumenter og internasjonal forskning, og peker på områder der det er behov for mer forskning.

KAPITTEL 2

HVA ER KUNSTIG
INTELLIGENS?



Kapittel 2

Hva er kunstig intelligens?

For å forstå sammenhengen mellom kjønnslikestilling og kunstig intelligens (KI), er det hensiktsmessig å se nærmere på hva kunstig intelligens egentlig er.

Det er ulike måter å definere kunstig intelligens på, og i takt med en rask teknologisk utvikling er definisjonene også i endring. Kort sagt er KI en samlebetegnelse på metoder for utførelse av oppgaver som krever intelligens, der intelligens avgrenses til kompleks problemløsning. Det finnes mange tilnærminger til KI. Det kan dreie seg om simulering av menneskelig intelligens innenfor et IT-system, eller som en del av en fysisk robot. Fagområdet KI bygger på studier fra psykologi, filosofi, nevrovitenskap, matematikk og datateknologi.

Likevel dreier KI-teknologi seg i hovedsak om å gjøre datamaskiner i stand til å løse oppgaver som tidligere har vært forbeholdt mennesker. Det handler om ulike verktøysett bestående av dataprogrammer som er tilsynelatende intelligente. Ved å analysere store mengder data kan systemene utføre oppgaver ved en viss grad av autonomi (Andreasson og Stende 2019: 11; Tørresen 2013: 14). Dette kalles gjerne kunstig intelligente systemer. Regjeringen tar i sin nasjonale strategi utgangspunkt i en definisjon utviklet av EUs ekspertgruppe for kunstig intelligens:

«Kunstig intelligente systemer utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Enkelte KI-systemer kan også tilpasse seg gjennom å analysere og ta hensyn til hvordan tidligere handlinger har påvirket omgivelsene» (Kommunal- og moderniseringsdepartementet, KMD 2020: 9).

I forbindelse med kunstig intelligens hører vi ofte om algoritmer. Algoritmer er matematiske form-ler som kan ses som en oppskrift eller en instruksjon for hva programmer skal gjøre – for eksempel å sortere store mengder data ut fra bestemte kriterier. Et sentralt element er hvordan bestemte algoritmer i KI-systemer er utviklet med evne til å «lære» selv, forbedre seg og se og tilpasse seg mønstre basert på tidligere erfaringer. Slike lærende systemer defineres som maskinlæring, og er en del av KI. EUs definisjon av kunstig intelligens inkluderer maskinlæringsteknikker med ulik grad av kompleksitet (inkludert framtreddende algoritmer som hjerneinspirerte nevralt nettverk), men også andre og mindre tekniske regelbaserte systemer. Denne definisjonen av KI kan dermed ses som en samlebetegnelse. Vi legger denne definisjonen til grunn i denne kunnskapsoppsummeringen.

«Sterk» og «svak» kunstig intelligens

Det er vanlig å skille mellom «sterk» og «svak» kunstig intelligens. Såkalt sterk kunstig intelligens viser til datasystemer som er *intelligente* og som har en form for selvbevissthet. Slike systemer finnes ikke i dag, og alle KI-systemer som er tatt i bruk er basert på det som kalles «svak» kunstig intelligens. Svak KI kjennetegnes ved at systemene kan utføre det vi forstår som intelligente oppgaver, som for eksempel se og gjenkjenne mønstre. Dagens KI-systemer er altså å regne som svake eller smale, og innrettet mot å løse oppgaver innenfor et avgrenset område. De er ikke svake fordi de gjør flere feil eller fordi systemene er mindre kraftfulle, men fordi systemet kun gjør akkurat det men-

nesket har instruert det til å gjøre (KMD 2020: 10).

I denne kunnskapsoppsummeringen løftes forskjellen fram for å synliggjøre at KI-systemer varierer med hensyn til grad av intelligens. Intelligens referer her til *autonomi*, evnen til å utføre avanserte oppgaver uten kontinuerlig veiledning fra mennesker, og *adaptivitet*, evnen til å forbedre operasjoner ved å lære av erfaringer. Kunstig intelligens består altså av flere ulike teknikker og metoder som varierer med hensyn til kompleksitet og hvor teknisk avanserte de er, men alle regnes som svak eller smal kunstig intelligens (se faktaboks neste side).

Sterk og svak kunstig intelligens

Begrepene sterk og svak kunstig intelligens har sitt opphav i John Rogers Searle's (1980) skille mellom det å *være intelligent* og det å *oppføre seg intelligent*. Systemer basert på svak KI er simulerte og har evnen til å oppføre seg intelligent, til tross for at de er datamaskiner. Sterk kunstig intelligens er systemer som ikke bare oppfører seg intelligent, de er intelligente og har en form for selvbevissthet.

Smal kunstig intelligens og generell kunstig intelligens

Et annet skille som kan brukes i forbindelse med KI er skillet mellom smal og generell kunstig intelligens. Skillet kan ses som en filosofisk forståelse til om maskiner kan bli i stand til å fungere som mennesker (Russel og Norvig 2010: 1034). Kunstig generell intelligens referer til datamaskiner med intelligens på nivå med menneskelig intelligens (KMD 2020: 9). Slike systemer vil kunne løse intellektuelle problemer. Selv om utviklingen går raskt, er det enighet om at vi som samfunn er langt unna å utvikle generell kunstig intelligens, og alle metoder innen KI i dag er å regne som smal kunstig intelligens.

Maskinlæring

De teknikkene innen KI i dag som handler om å utvikle systemer som kan lære, tilhører det som kalles *maskinlæring*. Det er på dette området de store framskrittene innen kunstig intelligens gjøres i dag, og det er ofte disse teknikkene det henvises til når man snakker om KI (Andreasson og Stende 2019: 11; Teknologirådet 2018). Det er verdt å påpeke at maskinlæring heller ikke er et klart definert område innen kunstig intelligens, og flere av teknikkene er å regne som statistiske operasjoner som for eksempel *regresjon*.

Innen maskinlæring lærer datasystemer av store mengder data, og ikke ved detaljprogrammering av mennesker (Tørresen 2013). Maskin-

læringsalgoritmene kan lære, justere seg selv og se sammenhenger basert på data fra den virkelige verden (Teknologirådet 2018: 16). Slik kan KI-systemer basert på maskinlæring forstås som selvlærende systemer (Tørresen 2013: 14). Maskinlæringsalgoritmer kan lære på ulike måter (se faktaboks). Teknikken som så langt har vist seg å være mest suksessrik er såkalt *veiledet læring* (Teknologirådet 2018: 22).

Hvordan kan maskiner lære?

Maskiner kan lære på mange forskjellige måter. De tre vanligste er:

Veiledet læring:

Gjennom *veiledet læring* trenes en algoritme på et datasett med eksempler fra den virkelige verden. I disse datasettene har eksemplene både såkalte inngangsverdier og utgangsverdier. For eksempel pikselverdiene til et bestemt bilde og informasjon om hva det er bilde av. Deretter lager systemet en modell basert på hva den har lært om sammenhengen mellom inndata og utdata, for å kunne predikere nye verdier på ukjente utdata, altså utfall. Denne tilnærmingen er mest utbredt i dag, og kan for eksempel brukes til å bestemme om en e-post er seriøs eller søppel, eller om en føflekk er godartet eller ondartet.

Ikke-veiledet læring:

Såkalt *ikke-veiledet læring* viser til når en algoritme får et datasett uten en fasit eller et resultat. Algoritmen finner og gjenkjenner selv mønstre i et gitt datasett. Disse kan brukes til å ta beslutninger om nye inndata. Eksempelvis kan dette dreie seg om å oppdage undergrupper av sykdommer, slik at pasienter får en bedre tilpasset oppfølging. Gjennom ikke-veiledet læring kan systemene oppdage mønstre som mennesker ikke selv kan oppdage.

Forsterket læring:

En tredje form er det som kalles *forsterket læring*. Ved forsterket læring bygger algoritmen opp modellen ved å prøve og feile, og gjennom tilbakemeldinger fra brukeren eller systemet den er koblet til. Datamaskinen kan på denne måten for eksempel lære seg å vinne i sjakk ved å hele tiden teste ut trekk og få tilbakemelding på om algoritmen vant partiet. Her er det potensiale til å finne smartere strategier enn mennesker kan.

Kilder: KMD 2020: 11; Teknologirådet 2018: 8.

Eksempler på hvordan maskinlæringsalgoritmer fungerer kan en finne i helsesektoren, et av områdene hvor kunstig intelligens trolig vil kunne løse flere oppgaver i framtida. Maskinlæringsalgoritmer kan for eksempel brukes til å beregne hvilke sykdommer pasienter risikerer å få, eller diagnostisere svært sjeldne forekomster av sykdommer. I 2017 demonstrerte en forskergruppe bestående av medisinere og KI-forskere at en maskinlæringsalgoritme trent opp på bilder av føflekker, var mer treffsikker enn hudlegespesialister (Esteva et al. 2017). Maskinlæring kan også brukes til å anbefale filmer eller musikk basert på kunder som ligner på hverandre. Ved å oppdage avvik fra et forventet mønster, kan maskinlæringsystemer varsle om hendelser

som banksvindel eller datainnbrudd, eller lage risikoprofiler for å forutsi om en person vil være i stand til å nedbetale gjeld (Teknologirådet 2018: 9).

En forutsetning for at systemer basert på maskinlæring skal fungere er data, og jo mer data og informasjon systemet har, desto bedre lærer det. Dataene som læringsalgoritmer kan trenes opp på, er en forutsetning for utviklingen av de kunstig intelligente systemene. Tilgang på store mengder data er ifølge Teknologirådet en av årsakene til framskrittene man har gjort særlig innen maskinlæring de siste årene. I tillegg har en utviklet bedre algoritmer og fått en enkel og rimelig tilgang på økende regnekraft (Teknologirådet 2018: 17).

Hva er algoritmenes «sorte boks»?

En utfordring med kunstig intelligens er at enkelte maskinlæringsalgoritmer kan sammenliknes med en «sorte boks» (KMD 2020: 12). Sorte boks er en metafor for å beskrive problemet om manglende innsyn og åpenhet ved automatiserte avgjørelser. Problemet er at en kun ser data som går inn og data som kommer ut, men ikke *hvordan* eller *hvorfor* de prosesseres. Selv om teknikker innenfor ikke-veiledet læring gjør at maskiner kan se sammenhenger i et datasett, kan ikke systemet nødvendigvis forklare hvordan det har kommet fram til en bestemt konklusjon. Dermed er det vanskelig å for eksempel begrunne en automatisert avgjørelse overfor en bruker.

Hvor viktig det er å kunne forklare hvorfor en beslutning er blitt tatt, vil avhenge på hvilket samfunnsområde KI-systemet blir brukt. Eksempelvis vil offentlig forvaltning være et område hvor forklarbarhet er svært viktig. Utfordringene med mangelen på transparens har gjort at forskere og programmerere i dag jobber med å utvikle det som kalles *forklarbar kunstig intelligens (explainable AI, XAI)* (Datatilsynet 2018: 26). Her handler det om å forstå hvordan avgjørelser er tatt. Gjennom statistiske modeller kan vi for eksempel finne ut at kjønn har hatt betydning for en lånesøknad, men utfordringen ligger i *hvordan* kjønn har betydning.

Regelbaserte systemer for automatisering

Mindre teknisk avanserte datasystemer, som også inkluderes i den samlebetegnelsen for KI som vi legger til grunn her, er såkalte regelbaserte systemer for automatisering. I disse systemene er det mennesker som står bak reglene for hvilke operasjoner datasystemet skal foreta seg. Dette til forskjell fra datasystemer basert på maskinlærings teknikker, der reglene utledes fra dataene som systemet er trent opp på (KMD 2020: 11). Reglene er gjerne basert på forretningslogikk, ekspertise eller regelverk. De regelbaserte systemene er gjerne bygget opp på en «hvis x, gjør y»-logikk, og kan settes sammen til avanserte beslutningssystemer (KMD 2020: 10). Disse systemene er ofte

del av et IT-system som er innrettet mot å effektivt løse oppgaver innenfor et avgrenset område.

Selv om de teknologiske framskrittene innen kunstig intelligens de siste årene er knyttet til maskinlærings teknikker (Teknologirådet 2018: 17), er også de regelbaserte systemene i ferd med å prege sentrale samfunnsområder i Norge. Dette gjelder både i arbeidslivet, helse- og omsorgssektoren og i utdanning. Også flere offentlige forvaltningsorganer i Norge har tatt i bruk denne typen systemer i digital og automatisert saksbehandling (KMD 2020: 11).

Fordomsfulle KI-systemer?

Som nevnt innledningsvis finnes det flere eksempler på at kunstig intelligente systemer har vist seg å ha svært uheldige konsekvenser i et

likestillingsperspektiv. Det er flere årsaker til at KI-systemer kan reprodusere stereotypier og fordommer, og stå i fare for å virke diskriminerende.

Utvalgsskjeve data

Som nevnt forutsetter utvikling av KI-systemer tilgang på en viss mengde data. Jo mer data maskinene har tilgang på, desto mer presise blir de og bedre i stand til å utføre oppgaver på egenhånd (Teknologirådet 2018: 16). Det er ikke bare den kvantitative mengden data som er viktig, dataene må også ha god kvalitet. I denne sammenhengen handler kvalitet om at dataene ikke bør være preget av utvalgsskjevhet. Dette oppstår dersom

datasettet kun inneholder informasjon om en begrenset del av det relevante datagrunnlaget (KMD 2020: 57). Årsaken kan være at datasettet speiler faktiske eller historiske skjevheter i samfunnet. Når utvalgsskjeve data blir grunnlaget for beslutningene til de kunstig intelligente systemene, kan resultatet bli fordomsfulle og diskriminerende beslutninger (Teknologirådet 2018: 10).

Hvem var Tay?

I 2016 lanserte Microsoft chatboten Tay på Twitter. Tay var bygget på et KI-system som skulle lære samtaleforståelse basert på samtaler på Twitter. Ideen var at jo mer twitterbrukere snakket med Tay, desto mer skulle boten lære. Med andre ord var tvitringene Tay mottok, dataene som chatboten skulle lære av. Mindre enn 24 timer etter lansering måtte imidlertid Microsoft fjerne chatboten fra Twitter. Twitterbrukere hadde matet systemet med hatefullt, sexistisk og annet kontroversielt innhold, som resulterte i at Tay hadde utviklet seg til å bli både rasistisk og kvinnehatende.

Kilde: Hunt (2016).

Hvem utvikler KI-systemene?

Et annet forhold som kan forårsake skjevheter i data og tilsynelatende fordomsfulle KI-systemer, er at de er kodet av mennesker. Som alle andre vil programmerere og utviklere på KI-området ha bevisste eller ubevisste fordommer knyttet til for eksempel kjønn og etnisitet, og disse kan reflekteres i modellene de utvikler. Dynamikken er særlig relevant med hensyn til kjønnsperspektivet, da KI er et område som må sies å være tydelig kjønndelt. Ifølge World Economic Forums *Global Gender Gap Report* utgjør menn 78 prosent av ekspertene eller fagpersonene innen kunstig intelligens på verdensbasis. Kun 22 prosent er kvinner (World Economic Forum 2018: 8). Tendensen er den samme i Europa, hvor andelen kvinner utgjorde 17 prosent av IKT-spesialister i EU i 2016 og andelen var synkende (European Institute for Gender Equality, EIGE, 2018). Kvinner utgjør et mindretall i IKT-studier, -yrker og fritidsaktiviteter i Norge og tendensen er stabil over tid (Corneliussen og Prøitz

2016). Dette til tross for at det er et uttalt politisk mål å øke kvinneandelen innenfor IKT-sektoren.

Ifølge Crawford (2016) vil KI i likhet med all annen teknologi reflektere skapernes verdier, og dersom sektoren ikke blir mer mangfoldig, står en i fare for å lage systemer som kun speiler en marginal og privilegert del av samfunnet. utfordringer knyttet til mangfold og likestilling i KI-sektoren handler ifølge West, Whittaker og Crawford (2019) grunnleggende sett om makt. Mangelen på likestilling har betydning for hvordan KI-bedrifter jobber, hvilke produkter som utvikles, hvilke målgrupper produktene er designet for og hvem som drar nytte av utviklingen (West et al. 2019: 5). Flere har også tatt til orde for at kjønnsperspektiver bør inngå i utviklingen av KI for å avdekke maktkonstruksjoner, men også for å hindre diskriminering (D'Ignazio og Klein 2020; Leavy; 2018; Adam 1995).

KI i likestillingens tjeneste?

Eksemplene fra de amerikanske tech-selskaperne og den manglende kjønnsbalansen blant de som utvikler kunstig intelligens, demonstrerer at KI-baserte systemer ikke er kjønnsnøytrale. Det betyr imidlertid ikke at KI-systemer utelukkende vil ha negativ betydning for likestillingen. I 2017 lyktes for eksempel det norske IT-selskapet Evry, som tok i bruk et KI-basert system som

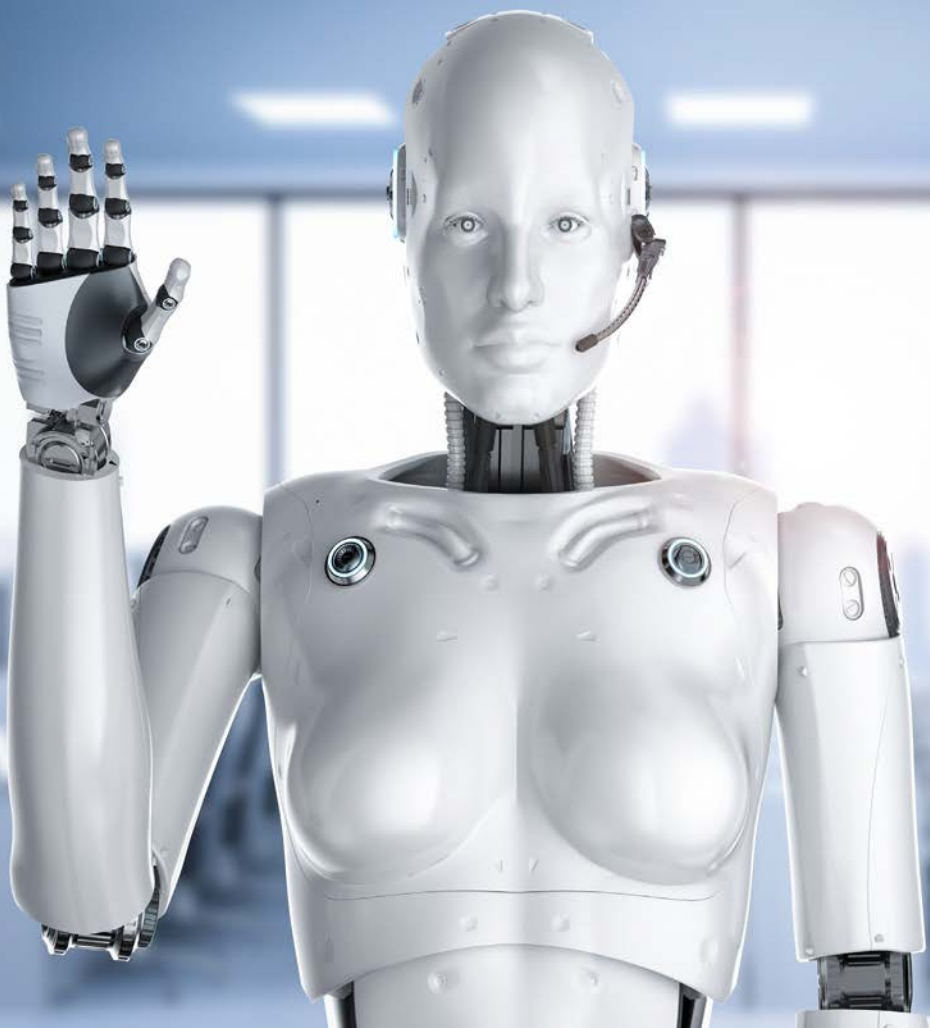
en del av rekrutteringsprosessen, å oppnå sitt mål om å ansette flere kvinner. Kvinneandelen som ble ansatt dette året var 33 prosent, mot 20 prosent et par år tidligere. Blant nyutdannede ansatte var kvinneandelen på 40 prosent. I følge Evry selv skyldes dette at utvelgelsen ved hjelp av systemet ble basert på mer objektive kriterier (Guldbrandsen 2017; Teknologirådet 2019: 49).

KI-teknologi kan dessuten brukes til å fremme likestilling og kvinners deltagelse på sosiale og digitale medier. For eksempel rammer kjønnsbaserte hatytringer og seksuell trakassering på sosiale digitale plattformer i særlig grad kvinnelige journalister, politikere, menneskerettighetsforkjempere og kvinner med offentlige roller, ifølge EUs rådgivningskomité for like muligheter mellom kvinner og menn (EU-kommisjonen 2020: 6). Komiteen peker imidlertid på at KI kan være et verktøy for å oppdage kjønnsbasert

hatefullt innhold på nettet. Et eksempel på dette presenterte en kanadisk forskningsgruppe da de i 2019 viste fram sin «ParityBOT» (Cuthberton et al. 2019). Dette er en KI-basert Twitter-bot (robot) som tar sikte på å direkte motvirke hatefulle tvitringer rettet mot kvinnelige politikere. Et annet eksempel er hvordan programmerere nå er i ferd med å utvikle såkalte «MeTooBots» som kan identifisere innhold som er å regne som seksuell trakassering i bedrifters dokumenter, e-poster og chatter (Woodford 2020).

KAPITTEL 3

METODE OG DATAGRUNNLAG



Kapittel 3

Metode og datagrunnlag

Kunnskapsoversiktens mål er å kartlegge norsk forskning på kunstig intelligens med kjønnsperspektiv, som tar for seg norske forhold relatert til arbeidsliv, tjenester og digitale sosiale plattformer. I samråd med Kilden kjønnsforskning.no og en referansegruppe bestående av fem sentrale forskere innen feltet (se forordet), har førstebibliotekar Anne Sæbø ved Universitetsbibliotekets HumSam-bibliotek i Oslo utført de systematiske søkene i de internasjonale og norske

databasene som utpekte seg som viktige. Søkene har deretter blitt kvalitetssikret av hovedbibliotekar Glenn Karlsen Bjerknes ved samme bibliotek. Kvalitetssikringen innebærer at søkestrategi og søkelogg har blitt vurdert og kommentert før ferdigstilling. Både fra Universitetsbiblioteket og Kildens side har de systematiske søkene blitt supplert med hånd søk og andre iterative søk. Metoden for søkene beskrives i det følgende.

Seleksjon av søkeord, søkestrenger og søkestrategier

Kilden hadde på forhånd satt opp en liste av søkeord med utprøvd søk i Oria, Web of Science og Google Scholar. Referansegruppa kom også med innspill til søkeord. Førstebibliotekaren ved Universitetsbiblioteket i Oslo utførte deretter en videre utprøving og justering i flere omganger av søkeord og søkestrenger med bruk av boolske (logiske) søkeoperatører (OR og AND) for å optimalisere søkestrategiene i de ulike basene. Alle søkeord der det ble utprøvd søk med ulike kombinasjoner og søkestrenger, er dokumentert i Vedlegg 1: Søkeord. De søkestrategiene som faktisk produserte relevante treff, og kun disse, er dokumentert i Vedlegg 2: Søkelogg.

Kunnskapsoversikten favner bredt, men det ble likevel tidlig klart i arbeidet med kartleggingen av relevant litteratur at det er publisert relativt lite fagfelleverdert forskning på dette emnet. Dette til tross for sentrale forskeres innstendige oppfordringer om nødvendigheten for forskning på dette (se f. eks. Lilleslått 2018; Borgan 2019). Vi besluttet derfor å ikke begrense det systematiske søket til tre baser, som opprinnelig bestilt fra Kilden til Universitetsbiblioteket, men til fem baser – to norske og tre internasjonale – for å fange opp alt det vi kunne av norsk forskning utgitt i både norske og internasjonale tidsskrifter. I tillegg så vi det som nødvendig å kartlegge hvem som forsker på dette i Norge, og gjøre manuelle søk på hva de eventuelt har publisert av forskning om dette («hånd søk»).* Disse søkene resulterte i ytterligere treff. «Hånd søk» referer til manuelle søk og kan også vise til søk i enkeltnumre av relevante tidsskrifter og artikler, og gjennomgang av referanselister i sentrale artikler og dokumenter,

også i sentrale rapporter og avhandlinger, såkalt grå litteratur. Grå litteratur er betegnelsen på litteratur som ikke er utgitt i kommersielle kanaler, og som dermed ikke nødvendigvis har samme vitenskapelige standard, men som like fullt kan vise til den ferskeste og mest relevante forskningen før den har gått gjennom den tidkrevende prosessen av fagfellevurderingen og kø for publisering.

I samråd med referansegruppa ble kunnskap om relevant litteratur også etterspurt fra aktuelle forskere og i relevante forskningsfora.**

Til sammen mener vi å ha utført en grundig systematisk kartleggingsoversikt av all relevant litteratur, og kartlagt det som er publisert av norsk forskning på norske forhold om kunstig intelligens med et kjønnsperspektiv, som berører arbeidsliv, tjenester og digitale sosiale plattformer. Systematiske kartleggingsoversikter etterstreber, slik systematiske søk gjør, å få full oversikt over et felt, men har ikke samme systematiske standard og like transparente etterprøvbarehet. En slik kartleggingsoversikt deler likevel funksjoner med et systematisk søk når det gjelder å være nøyaktig, gjennomiktig og reproduserbar (Folkehelseinstituttet 2020).

Hvert søk har blitt tilpasset de tekniske søkefunksjonene i den enkelte database. Dette har i hovedsak handlet om ulikheter i antall mulige søkekombinasjoner, bruk av trunkering og sorteringsmulighet. Trunkering betyr at en søker på stammen av et ord for å få med entalls-/flertallsendelser og ulike varianter av ordet.

Seleksjon av baser

Norart. Nasjonalbibliotekets base for norske og nordiske tidsskrifter: Norart peker seg ut som den sentrale databasen for referanser til norsk forskning, med referanser til artikler fra ca. 380 norske og nordiske årbøker og tidsskrifter, hvorav nærmere halvparten er vitenskapelige. Søk i Norart er helt essensielt i en kartlegging av norsk forskning på norske forhold. Imidlertid har ikke Norart like optimal funksjonalitet for søk som store internasjonale baser har, med begrensninger i både lengde og kompleksitet på søkestrenger. Her har det da vært nødvendig å søke bredt med enkle og åpne søkeord, og så heller gå gjennom lange trefflister for ikke å utelate relevant litteratur.

Videre har ikke basen samme nivå av presisjon som de store internasjonale basene. Et søk på «kunstig intelligens» i basen resulterer i 113 treff. Derav er det 4 treff når søket er avgrenset til vitenskapelige artikler. Men en gjennomgang av alle de 113 treffene viste derimot at 8 av de 113 var vitenskapelige artikler. Dette understreker betydningen av manuell kontroll av trefflister.

I vedlegget for søkeloggen er de søkeordene og søkestrengene som gav treff på relevant litteratur i denne basen, dokumentert. Søk

med de andre norske og engelske søkeordene (Vedlegg 1) resulterte ikke i ytterligere treff, og er derfor ikke med i søkeloggen.

Idunn: Idunn er Universitetsforlagets digitale plattform for fag- og forskningstidsskrifter og åpent tilgjengelige forskningsbøker. Det er mulig å gjøre enkle og avanserte søk i Idunn, men med begrensninger. Da mye norsk forskning er publisert og tilgjengeliggjort i Idunn, er det likevel viktig å gjøre søk her for å supplere treff i Norart og ekskludere dubletter. Heller ikke her resulterte søk med de andre norske og engelske søkeordene til ytterligere treff, og er derfor ikke med i søkeloggen.

Web of Science (levert av Clarivate Analytics): Web of Science dekker verdens ledende tidsskrifter innen naturvitenskap, samfunnsfag og humaniora.

Scopus (levert av Elsevier): Scopus er en stor internasjonal database innen livsvitenskap, sosialvitenskap, helsevitenskap og fysikk.

Sociological Abstracts (levert av ProQuest): Sociological Abstracts er en omfattende database som dekker sosiologi, sosialantropologi, medievitenskap, pedagogikk og andre beslektede fag.***

Inklusjons- og eksklusjonskriterier

Treff har blitt avgrenset til fagfelleverderte publikasjoner, rapporter og avhandlinger som dokumenterer norsk forskning på kunstig intelligens med kjønnsperspektiv, og som tar for seg norske forhold relatert til arbeidsliv, tjenester og digitale sosiale plattformer. Målsettingen har vært å inkludere litteratur hvor kjønn ikke kun er en variabel, men del av en systematisk analyse. Imidlertid har det vist seg nødvendig å også inkludere litteratur hvor kjønn kun berøres i mindre grad. Denne litteraturen åpenbarer dessuten både behov for mer kunnskap og potensialet for nye studier.

Når det gjelder søk i de internasjonale basene, resulterte for eksempel et enkelt søk på «artificial intelligence», avgrenset til forskning i Norge og artikler utgitt de 10 siste årene, betydelig færre treff. Men alle disse treffene, og ikke flere, er inkludert i trefflisten til den mer komplekse søkestrengen vi endte opp med. Når vi likevel dokumenterer det større, mer komplekse søket vi utførte, er dette for å demonstrere hvor omfattende og grundig vi gikk til verks for å «støvsuge» feltet for å ikke gå glipp av relevant litteratur.

Vi har også søkt etter relevant forskning i aktuelle rapporter, conference papers, avhandlinger og masteroppgaver, og har da gjort søk i alle norske vitenarkiv.**** Vi har også søkt i open-access basen arXiv.org der det er opprettet en egen kategori for kunstig intelligens. Søk her resulterte ikke i treff på relevant forskning, og er derfor ikke med i søkeloggen.

Dubletter og litteratur som ikke tilfredsstilte inklusjonskriteriene, ble sortert bort i databasene der det var mulig. For ikke å ekskludere mulig relevant litteratur, har det vært nødvendig å analysere treff ikke bare basert på tittel, emneord og sammendrag, men å gå inn i fulltekst for å skimlese og gjøre søk på emneord der.

Søketreffene ble deretter importert til et referansebibliotek i programvaren EndNote og delt med prosjekt- og referansegruppa, samt andre forskere og forskerforum som ble kontaktet for innspill. Den fullstendige referanselisten ble til slutt delt med referansegruppa online for endelig vurdering og siste innspill.

Det totale antall norske studier som presenteres som egne caser i denne rapporten er 24 (21 vitenskapelige artikler og 3 forskningsrapporter). Ikke alle disse har et tydelig kjønns- eller likestillingsperspektiv, men er likevel inkludert fordi de på mer implisitte måter er relevante i et kjønns-

og likestillingsperspektiv. Det kan for eksempel være fordi studiene belyser sosial ulikhet i møte med myndighetene, eller fordi de undersøker endringer som følge av automatisering i arbeidslivet, som i Norge fortsatt er relativt kjønnsdelt.

Noter:

* Forskere som utpekte seg på fagfeltet, og som det ble gjort manuelle søk på, inkluderer bl.a. Lene Petersen, Taina Bucher, Anne-Britt Gran, Miria Grisot, Margunn Aanestad, Ingunn Moser, Ann Therese Lotherington, Einar Duenger Bøhn, Arnt Maasø, Petter Bae Brandtzæg, Jill Walker Rettberg, Mareile Kaufmann, Lisa Reutter, i tillegg til forskerne i referansegruppa: Hege Kristin Andreassen, Erling Barth, Morten Goodwin, Tereza Østbø Kuldova og Roger A. Søraa.

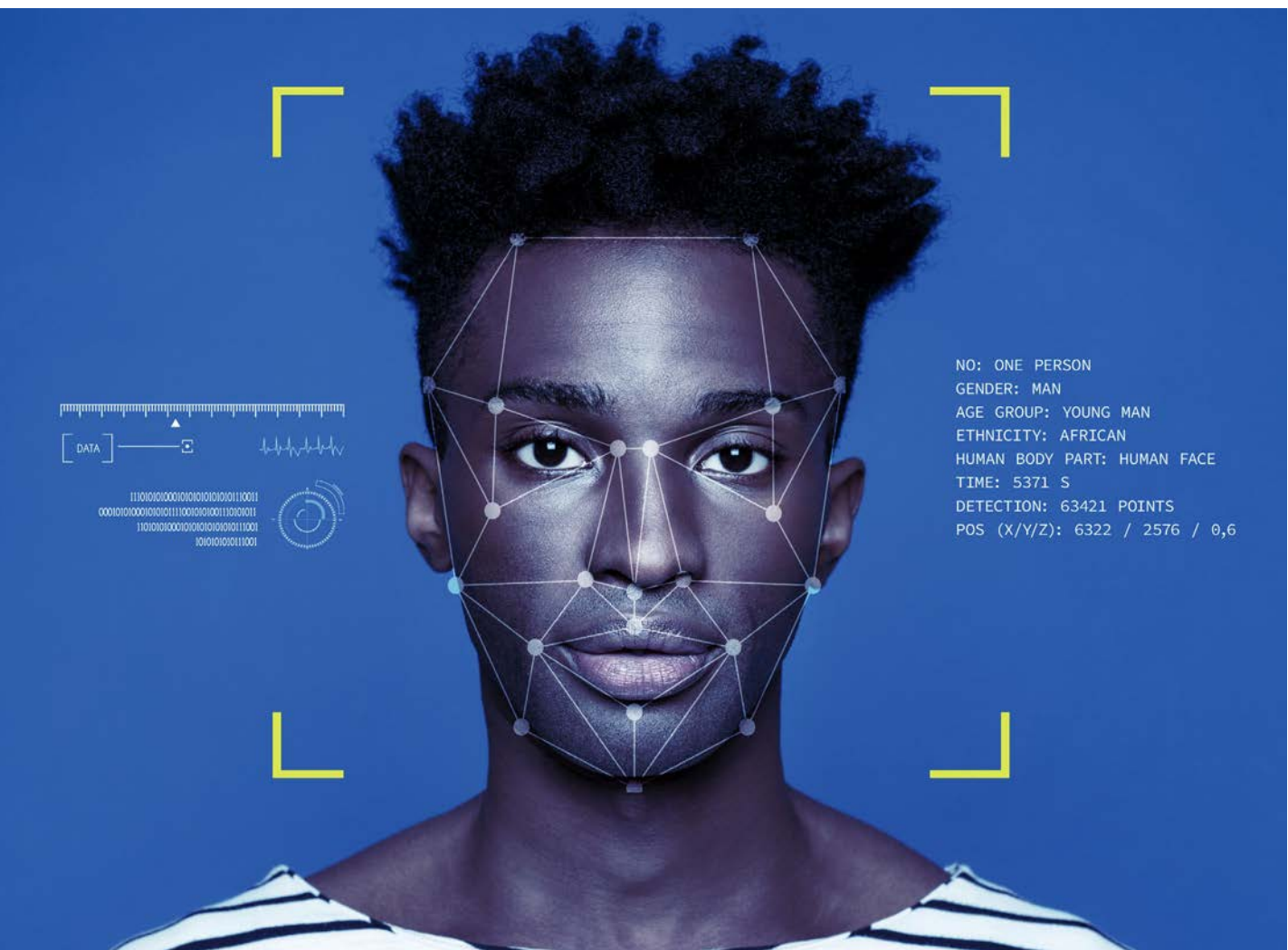
** E-post ble sendt ut til forskere som utpeker seg etter systematiske og manuelle søk, samt til de to sentrale forskningsnettverkene innen feltet i Norge: NORA: Norwegian Artificial Intelligence Research Consortium og Norwegian Open AI Lab ved NTNU. I tillegg postet vi et innlegg på veggen til Facebook-gruppa «Demokrati og digitale medier» hvor vi etterlyste innspill.

*** International Bibliography of the Social Sciences – IBSS (levert av ProQuest) ble også utprøvd, men ble ikke tatt med pga. manglende treff på relevant litteratur her.

**** I tillegg til Nora Open Access – «Alt i norske vitenarkiv i én tjeneste» (<https://nora.openaccess.no/>) – har vi søkt i alle norske universiteters kunnskapsarkiv (se fullstendig oversikt her: <https://www.openaccess.no/SOS%20-%20Sp%C3%B8rsm%C3%A5l%20og%20Svar/norske-institusjonsarkiv/>), da Nora viste seg å ikke dekke disse fullstendig. Disse søkene resulterte kun i treff på én relevant masteroppgave og er derfor ikke dokumentert i søkeloggen. Vi gjorde også søk i arkiv til relevante forskningsinstitutter uten hell, så disse har vi heller ikke tatt med i søkeloggen.

KAPITTEL 4

TJENESTER



NO: ONE PERSON
GENDER: MAN
AGE GROUP: YOUNG MAN
ETHNICITY: AFRICAN
HUMAN BODY PART: HUMAN FACE
TIME: 5371 S
DETECTION: 63421 POINTS
POS (X/Y/Z): 6322 / 2576 / 0,6

Kapittel 4

Tjenester

I Norge er det et uttalt politisk mål at offentlig og privat sektor skal øke bruken av kunstig intelligente systemer for å effektivisere og levere mer treffsikre og brukertilpassede tjenester (KMD 2020: 53). Forventningene er også store til hva kunstig intelligens kan utrette i helsetjenesten når det gjelder raskere og mer presis diagnostisering, bedre behandling og mer effektiv ressursbruk (Helse- og omsorgsdepartementet, HOD, 2019).

Offentlige forvaltningsorganer som NAV, Lånekassen, Husbanken og skatteetaten har allerede tatt i bruk automatisert saksbehandling i form av regelbaserte systemer som bygger på prinsippet «hvis x, gjør y». Også opptak til videregående opplæring

og til høyere utdanning skjer gjennom automatiserte vedtak (NOU 2019: 5). Både private og offentlige tjenesteytere og institusjoner bruker chatbotter for å svare på enkle spørsmål. I helsetjenesten brukes kunstig intelligens i medisinsk utstyr, som for eksempel kirurgiske roboter og bildediagnostikk, og det planlegges å bli brukt for å analysere og systematisere helsedata (KMD 2020: 23).

Dette kapitlet tar for seg forskning om kjønns- og likestillingsaspekter ved innføringen av automatiserte KI-systemer i tjenester som er vesentlige i innbyggernes dagligliv fra et brukerperspektiv. Vår kartlegging avdekker at dette er et utforsket område i Norge.

Forskningsstatus

Som nevnt i kapittel 1 og 2 har internasjonal forskning påvist hvordan algoritmer kan virke diskriminerende fordi verken de som konstruerer systemene, eller de store datamengdene de bygger på, er nøytrale. Vi har ikke funnet empirisk forskning fra norske forskningsinstitusjoner som eksplisitt tar for seg kjønns- og likestillingsperspektiver når det gjelder automatiserte beslutningsprosesser, eller bruk av annen KI-teknologi i helsetjenestene. Men vi har funnet flere artikler som på et kunnskapsteoretisk grunnlag drøfter mulige diskriminerende aspekter ved KI-baserte systemer i vesentlige tjenester for befolkningen, inkludert robotteknologi, og som etterlyser kjønnsperspektiver i framtidige empiriske studier (Bucher 2020; Buhman, Paßmann og Fieseler 2020; Poulsen, Fosch-Villaronga og Søraa 2020; Søraa 2017; Søraa og Fosch-Villaronga 2020).

At algoritmer er preget av fordommer og skjevheter, er et tema i forskning på såkalt intelligent overvåking med bruk av KI. Internasjonalt har politi og grensekontroll begynt å eksperimentere med bruk av algoritmiske systemer for å forutse hvem som vil gjøre kriminelle handlinger, en utvikling som potensielt kan undergrave

enkeltmenneskers rettigheter (Kuldova 2020). Et annet tema som er relevant i en rettighetsammenheng, gjelder forskning på NAVs digitale og automatiserte saksbehandling. Selv om ikke kjønnsperspektivet er sentralt i denne forskningen, er den nyttig fordi den påviser at enkelte grupper ikke får ytelsene de har krav på fordi de ikke mestrer NAVs digitale løsninger (Hansen, Lundberg og Syltevik 2018; Kane 2020).

Fordi forskningslitteraturen på samfunnsmessige konsekvenser av bruken av KI i offentlige og private tjenester i Norge er nokså beskjeden, har vi i dette kapitlet også tatt med grå litteratur. Når det gjelder bruk av KI i offentlig sektor, har en rapport fra Nordisk ministerråd undersøkt nordiske kommuners bruk av KI (Andreasson og Stende 2019). Mulighetene for å bruke KI som verktøy for å bosette flyktninger i norske kommuner, blir diskutert i en NOVA-rapport om strategisk mottaksplassering og treffsikker bosetting og integrering av flyktninger (Seeberg et al. 2020). Heller ikke i disse rapportene diskuteres kjønnsperspektiver i særlig grad, selv om det nevnes at det er viktig å være oppmerksom på dem.

Inkluderende og diskriminerende robotteknologi

NTNU-forsker Roger A. Søraa har i flere artikler argumentert for en mer inkluderende robotteknologi som ikke bygger på maskuline normer og kjønnsstereotyper (Søraa 2017; Poulsen, Fosch-Villaronga og Søraa 2020; Søraa og Fosch-Villaronga 2020). Ut fra en feministisk teori-tradisjon diskuterer Søraa (2017) hvordan roboter påvirker vårt syn på kjønn, og hvordan vårt syn på kjønn manifesterer seg i hvordan roboter blir utformet. Kulturen «kjønner» roboter, i den forstand at jo mer menneskeliknende en robot er, jo mer kjønnnes den, ofte på en stereotypisk måte. I kommentarartikkelen «Queering machines» påpeker Adam Poulsen, Eduard Fosch-Villaronga og Roger A. Søraa (2020) at marginaliserte grupper fra LHBT-miljøer kan bli ekskludert på grunn av heteronormative algoritmer. For eksempel har KI-verktøyet *Perspective*, som flagger «giftig» (*toxic*) innhold på internett, flagget dragartisters Twitter-kontoer som «giftige» oftere enn de har gjort med rasistiske kontoer, fordi systemet ikke forstår i hvilken kontekst ord og uttrykk brukes. Dermed kan ord som i noen LHBT-mil-

jøer brukes i positiv forstand, flagges som hatefullt. Poulsen, Fosch-Villaronga og Søraa (2020) argumenterer for at det er viktig å utvikle kunstig intelligente systemer som inkluderer LHBT-miljøene og andre ofte marginaliserte grupper.

Robotteknologi som blir brukt i helsevesenet kan også ekskludere grupper som faller utenfor normen. For eksempel kan såkalte eksoskjelett, en robot som kan hjelpe personer med lammelser til å reise seg og gå, være konstruert på en måte som gjør at personer som er enten for høye eller for lave, for tunge eller for lette, ikke kan bruke dem (Søraa og Fosch-Villaronga 2020). Mekanikk og robotteknologi er historisk blitt oppfattet som del av en mannlig sfære, og kriteriene som blir brukt i utviklingen av ny teknologi kan fortsatt gjenspeile tidligere tiders kjønnskjevheter, og idealer om «standardmennesket». Inspirert av interseksjonelle perspektiver, foreslår Søraa og Fosch-Villaronga (2020) konkrete verktøy som designere og produsenter kan bruke for å gjøre eksoskjeletter mer inkluderende og tilpasset alle slags kroppsfasonger.

Kunstig intelligens i Kommune-Norge

I offentlig sektor er det altså et mål å øke bruken av kunstig intelligente systemer for å effektivisere og levere mer brukertilpassede tjenester. Kommuner i Norge har begynt å ta i bruk KI, men arbeidet er fortsatt på et tidlig stadium, ifølge en rapport skrevet av Ulf Andersson og Truls Stende fra Nordisk ministerråd (2019). Foreløpig begrenser bruken av KI seg hovedsakelig til chatboter som for eksempel den feminine «Kommune-Kari», som svarer på enkle spørsmål fra innbyggere og henviser til nettsider. Samtidig planlegger og tester mange kommuner ut KI-teknologi. Rapporten ser på kommuner i hele Norden, og de utvalgte norske kommunene er Oslo og Trondheim. Studien er dermed ikke representativ for hele Kommune-Norge, men beskriver generelle tendenser for hvordan nordiske kommuner arbeider med innføring av KI, og hvordan KI-teknologi kan påvirke befolkningens tillit til den offentlige forvaltningen. Ettersom prinsippet om likebehandling holdes høyt i den nordiske statsforvaltningen, påpeker rapporten at tilliten til kommunene forutsetter at den offentlige sektoren er rettfærdig og ikke forskjellsbehandler uten saklig grunn. Like tilfeller skal behandles likt, og innbyggere skal ikke forskjellsbehandles på grunn av kjønn og etnisitet (Andreasson og Stende 2019: 30).

sempel mener kommunene at KI vil være et nyttig verktøy for å forutse og forhindre lekkasjer i vann- og avløpsnettet og for å kunne forutse og identifisere personer eller bedrifter med «høy risiko for uønsket utvikling», slik at kommunen kan iverksette tidlige tiltak (Andreasson og Stende 2019: 8). Sistnevnte bruk av KI-baserte forutseende analyser var imidlertid ennå ikke iverksatt da undersøkelsen ble utført.

Et tredje område der KI-baserte verktøy blir sett som nyttig, er i kommunenes saksbehandling. Ifølge Andreasson og Stende kan KI «hjelpe kommunene til å bli enda bedre til likebehandling enn de er i dag», samtidig som «teknologien også kan resultere i mer urimelig forskjellsbehandling» (Andreasson og Stende 2019: 30). Ettersom algoritmer kan videreføre skjevheter fra dataene de bygger på, for eksempel knyttet til kjønn, kan det forsterke en eksisterende forskjellsbehandling. Ifølge rapporten tok kommunene i undersøkelsen i liten grad opp denne problemstillingen. Samtidig uttrykte mange av kommunene bekymring for at private aktører kan få for stor innflytelse på offentlig forvaltning og demokratiske prosesser, og at dette kan svekke innbyggernes tillit til kommunene (Andreasson og Stende 2019:12).

Kommunene i studien ser for seg at KI kan gi bedre tjenester og mer effektiv drift. For ek-

KI som verktøy for å bosette flyktninger i norske kommuner

Norsk institutt for forskning om oppvekst, velferd og aldring (NOVA) har på oppdrag for Integrerings- og mangfoldsdirektoratet (IMDi) skrevet en rapport om bosetting av flyktninger i kommunene, der KI som verktøy for treffsikker bosetting blir diskutert (Seeberg et al. 2020). IMDi's nasjonale bosettingsteam har allerede tatt i bruk digitale verktøy som IMDinett og Kompass for å bistå i beslutninger om hvor flyktninger skal bosettes, men den endelige avgjørelsen ligger fortsatt hos den enkelte saksbehandlers skjønnsvurdering (Seeberg et al. 2020: 57). Internasjonalt jobbes det med å utvikle algoritmiske verktøy basert på maskinlæringsteknikker for treffsikker bosetting av flyktninger og asylsøkere basert på teorien om matching i markeder. Teorien går i korte trekk ut på å bruke en matematisk matchingsfunksjon for å «optimalisere dannelse av relasjoner mellom umatched individer» (Seeberg et al. 2020: 58), for eksempel å koble arbeidstakere med arbeidsgivere, eller, som i dette tilfellet, koble flyktninger med lokalsamfunn.

Sveits er det første landet som har begynt å bruke KI for å bosette flyktninger, ved hjelp av en maskinlæringsalgoritme som er utviklet av Immigration Policy Lab (IPL). Ut fra historiske data om kjennetegn ved flyktninger og deres videre karriere og bosted, beregner algoritmen sannsynligheten for

at den enkelte flyktningen, med et sett personlige demografiske kjennetegn, skal komme i jobb på et gitt geografisk sted (Seeberg et al. 2020: 59-64). Spørsmålet er om slike algoritmer skal brukes for å bosette flyktninger i norske kommuner. Som Seeberg et al. påpeker, er det ikke nødvendigvis slik at automatiserte beslutningssystemer tar bedre beslutninger selv om de går raskere. Maskinlæringsalgoritmer tar beslutninger på grunnlag av dataene de er matet med, og siden disse kan være skjeve eller mangelfulle, er det fare for at de automatiserte beslutningene gir utilfredsstillende resultat. For eksempel kan en flyktning bli registrert med «lav måloppnåelse» når det gjelder å komme i jobb i en kommune, uten at det verken skyldes flyktningen selv eller kommunen vedkommende er bosatt i (Seeberg et al. 2020: 67).

Fordi arbeidet med bosetting av flyktninger er så sammensatt, anbefaler ikke NOVA-rapporten at IMDi innfører KI-baserte systemer for treffsikker bosetting før de er grundig prøvd ut og evaluert (Seeberg et al. 2020: 72). Modeller for bosetting av flyktninger må være fleksible, og også ta hensyn til at familiemedlemmer har ulike behov. Rapporten anbefaler at hvert familiemedlem må kartlegges hver for seg før beslutninger om bosetting blir tatt, og at det er viktig å ivareta både barneperspektivet og kjønnsperspektivet.

Digitalisering av NAV-tjenester i et brukerperspektiv

Stadig flere av NAVs tjenester er blitt digitalisert og automatisert, i tråd med målet om at offentlig saksbehandling skal bli mer effektiv. Hvert år har om lag halvparten av innbyggerne i Norge befaring med NAV. Mange av disse mottar stønader som er basert på enkle kriterier og som ikke fordrer personlig kontakt, som for eksempel barne-trygd. Andre har mer sammensatte problemer som krever stadig samhandling med NAV over tid.

Hans-Tore Hansen, Kjetil Lundberg og Liv Johanne Syltevik (2018) har studert NAV-brukeres erfaringer med digital teknologi. Studien er en del av evalueringen av NAV-reformen, som ble ledet av Uni Rokkan (2006-2014). Den er basert på surveydata fra og med 2008 til og med 2016 (n=7500 per år med gjennomsnittlig svarprosent på ca. 20), feltarbeid på NAV-kontorer og kvalitative intervjuer med personer som mottar helse- og velferdsgoder. Forskerne undersøkte hvorvidt de digitale løsningene hadde erstattet møter ansikt til ansikt, og hvorvidt erfaringene ble påvirket av alder, kjønn og stønadstype.

Forskerne fant at skjermkommunikasjon hadde erstattet ansikt-til-ansikt-kommunikasjon for noen grupper (som nylige pensjonister og mottakere av familieytelser), og at mange NAV-brukere med digital kompetanse foretrekker upersonlig kontakt med systemet. Samtidig fant de at mange som er i kontakt med NAV har komplekse problemer som ikke nødvendigvis passer inn i de forhåndsdefinerte systemene (Hansen, Lundberg og Syltevik 2018: 84). Enkelte brukere forteller at de opplever at IKT-løsningene gir en følelse av å miste seg selv i systemet, og ikke bli behandlet som et virkelig menneske. Andre opplever å ikke få de utbetalningene de har krav på fordi de av ulike grunner ikke har fylt ut de digitale skjemaene på korrekt måte.

Forskerne ser tendenser til at det oppstår et nytt digitalt skille mellom brukere av velferdstjenester (jf. diskusjonen om *digital divide* i kapittel 1). Det digitale skillet mellom brukere av velferdstjenester handler om hvem som mestrer og ikke mestrer de digitale løsningene. Forskerne finner ikke at det er store kjønnsforskjeller når det gjelder dette, men

at utdanningsnivå, alder og livssituasjon er viktigst. Generelt sett mestrer voksne og middelaldrende med høyere utdanning systemet best, mens lavt utdannede eldre mestrer det minst. Samtidig kan personer med høy digital kompetanse falle ut av systemet når depresjon eller andre psykiske problemer gjør at de ikke klarer å forholde seg til det.

Når det gjelder bruken av NAVs nettsted for å skaffe seg informasjon på egenhånd, fant Hansen, Lundberg og Syltevik at kvinner bruker det noe mer enn menn, uten at kjønnsforskjellen er signifikant. De fant ikke forskjeller mellom norskfødte og utenlandsfødte, men at bruken av de selvbetjente løsningene øker med utdanningsnivå, uansett landbakgrunn (2018: 80). Et interessant funn er at de høyt utdannede respondentene svarte at de var minst fornøyde med informasjonen på NAVs nettsted, og at norskfødte var mindre fornøyde enn utenlandsfødte. Misnøyen med informasjonen var også større blant menn enn blant kvinner (Hansen, Lundberg og Syltevik 2018: 81). Samtidig er ikke synspunkter på NAVs nettside avgjørende for om brukerne faktisk forstår og kan bruke systemet. Problemet oppstår når enkelte grupper ikke får ytelsene de har krav på fordi de ikke mestrer NAVs digitale løsninger.

Jurist Aina Aune Kane utførte i 2017–2018 kvalitative intervjuer med ansatte ved tre NAV-kontorer. NAV-kontorene var i Nord-Norge og blant respon-

dentene var kvinner i flertall. Kane reiser spørsmålet om NAVs digitale løsninger i tilstrekkelig grad sikrer brukernes rettssikkerhet (Kane 2020: 44). Hennes studie viser at NAV-ansatte er bevisst både fordelene og ulempene ved de digitale løsningene. På den ene siden gjør systemet «Ditt NAV» at brukerne lettere kan nå og holde kontakt med NAV-systemet. På den andre siden kan brukere som ikke mestrer digitale løsninger, eller ikke mestrer skjematiske besvarelser på standardiserte spørsmål, komme til kort overfor NAVs krav, og risikere å bli feilvurdert og dermed bli stående uten bistand fra NAV. Forhåndsbestemte svaralternativer uten muligheter for fritekst kan dessuten begrense brukernes mulighet til å gi relevant informasjon som er nødvendig for å gi tilstrekkelig kartlegging av bistandsbehov.

Ifølge Kane kan NAVs digitale løsninger utgjøre en systematisk risiko for rettighetstap for brukerne. Hun viser til at hovedmålet med digitalisering av offentlige tjenester er «en brukerrettet og effektiv offentlig forvaltning» hvor «brukernes behov skal være det sentrale utgangspunktet» (KMD 2016, ref. i Kane 2020: 44). NAV-ansatte i Kanes studie sier at de digitale løsningene ikke i tilstrekkelig grad rommer informasjon om brukerne og deres behov. Kane argumenterer derfor for at målene om effektivitet ikke må gå på bekostning av grunnleggende rettigheter.

Intelligent overvåkning og enkeltmenneskers rettigheter

KI, maskinlæring og analyse av store datamengder blir i økende grad brukt i såkalt intelligent overvåkning, i politiet og i grensekontroller. For å forebygge kriminalitet, blir det utviklet KI-baserte prediktive (eller forutseende) analyser for å forutsi forbrytelser og identifisere kriminelle. Spesielt i USA tar politiet i bruk forutseende analyser for å sannsynliggjøre hvor og når det er risiko for visse typer kriminelle handlinger, hvem som er målet for slike og hvem som har sannsynlighet for å utføre kriminalitet, og som politiet dermed følger ekstra med på (Teknologirådet 2015: 4-5).

Det europeiske grense- og kystvaktbyrået Frontex, som har ansvaret for å beskytte Schengen-landenes yttergrenser og som Norge er tilsluttet, er i front for å utvikle og teste ny KI-basert teknologi. For eksempel har Frontex testet ut et kontroversielt ansiktsgjenkjenningssystem som ut fra tolkning av folks ansiktsmimikk skal avsløre antatt «uærlige» inntrengere ved grensene (Kuldova 2020: 54).

Forsker ved Arbeidsforskningsinstituttet AFI, Tereza Østbø Kuldova (2020), advarer mot at intelligente overvåkingssystemer kan føre til et

mer autoritært og paranoid samfunn. For å sikre en mer effektiv og tilsynelatende «evidensbasert» offentlig administrasjon, og et mer forutseende og «intelligent» politi og militærvesen, samarbeider myndighetene med store private selskaper, og gjør seg avhengige av dem og deres teknologiske løsninger (Kuldova 2020: 53).

Kuldova bruker sine etnografiske studier av kriminelle MC-klubber i Sentral-Europa for å belyse hva overvåkning og mistillit gjør med samfunnet. Ifølge Kuldova gjenspeiler disse MC-klubbene på en ekstrem måte en generell samfunnstrend preget av en sterk opptatthet av sikkerhet, kontroll og overvåkning, og det hun kaller et paranoid verdensbilde der alle ses som potensielle svindlere eller inntrengere.

Kun menn kan bli medlemmer i de kriminelle MC-klubbene, der det dyrkes en særegen maskulin kultur med egne strenge regler og ritualer. Det såkalte brorskapet er hierarkisk og utgjør transnasjonale nettverk. Den mest beryktede MC-klubben av dem alle, Hell's Angels, som i dag har mer enn 450 avdelinger i over 50 land,

er kjent for å drive med narkotikahandel, prostitusjon og illegal våpenhandel, og har i det siste også begynt med cyberkriminalitet (Kuldova 2019). Fordi de kriminelle MC-klubbene oppfattes som representanter for organisert kriminalitet, brukes de for å legitimere nye sikkerhetstiltak, lovendringer og overvåkningsmetoder, samt deling av data og etterretning på tvers av landegrensener.

Samtidig har kriminelle MC-klubber tatt i bruk samme overvåkningsteknologi som brukes mot dem, både mot utenforstående og mot sine egne (Kuldova 2020: 58-59). Kuldova mener MC-klubbenes besettelse av å avsløre svindlere gjenspeiles i storsamfunnet. Myndighetene bruker overvåkning for å avsløre alt fra terrorister, cyberkriminelle og identitetstyver til «illegale» flyktninger med fabrikkerte historier og trygdesvindlere. For å avdekke det myndighetene mener er sannheten, blir KI og algoritmisk styring basert på store datamengder oppfattet som nyttige redskaper. Samtidig er de selvlærende algoritmene innelukket i teknologiselskapenes «sorte boks». For når et «intelligent» overvåkingskamera oppfatter oppførselen til en person som «avvikende» og unormal og slår fast at det er en mulig kriminell, er det ikke nødvendigvis bestemt av personens egne handlinger, men av andres handlinger fordi personen likner på noen som tidligere har gjort et lovbrudd, som for eksempel et kriminelt MC-medlem eller en svart og fattig

mann. Akkurat slik som de målrettede reklamene på Facebook er basert på analyser av hva en gjør på nettet, kan lovverket bli tilpasset gjennom algoritmisk profilering. Når slike algoritmebaserte verktøy overføres til kriminalitetsforebygging i form av forutseende risikovurdering er det en fare for rettssikkerheten, ifølge Kuldova (2020: 53).

Samtidig påpeker Kuldova at overvåkning og opp-tattheten av å sikre seg mot svindlere ikke begrenser seg til grensekontroller, politiet og militæret, men er blitt en viktig del av hverdagen. For eksempel har flere arbeidsplasser innført overvåkning av sine ansatte, og nettstedet Airbnb, som formidler overnattingssteder over hele verden, har patentert KI-baserte prediktive analyser som saumfarer gjestenes profiler på sosiale medier og andre digitale spor de har etterlatt seg på nettet, for å vurdere om de kan tenkes å gjøre suspekte handlinger. Forsikringsselskapet Gjensidige utvikler risikovurderingssystemer for å skille mellom ærlige og uærlige kunder med det de kaller en smart algoritme bygget på objektive data (Kuldova 2020: 55). Kuldova minner om at såkalte objektive data har en tendens til å reproducere samfunns-messige ulikheter, fordommer og stereotypier. Slik kan forutseende systemer basert på algoritmer ikke bare føre til feil konklusjoner og urettferdig behandling, men også bli selvpoppfyllende profetier og føre til et mer autoritært samfunn.

Oppsummering

I dette kapitlet har vi tatt for oss norsk forskning som er relevant for å belyse likestillingskonsekvenser av KI-baserte og automatiserte offentlige og private tjenester. Vi har sett på hvordan samfunnsvitenskapelig forskning problematiserer hvordan roboter «kjønnes» og dermed kan videreføre eller forsterke kjønnsstereotypier. Et annet forskningsfelt ser på følger av automatisert saksbehandling fra et brukersperspektiv. I denne

forskningen påpekes det at det er en fare for at automatiserte beslutninger kan gi feilaktig resultat fordi maskinlæringsalgoritmer tar beslutningene på grunnlag av dataene de er matet med, og disse kan være skjeve, mangelfulle eller fordomsfulle. Dette gjelder både automatisert saksbehandling i NAV og forutseende analyser for å bosette flyktninger eller for å forebygge kriminalitet.

KAPITTEL 5

ARBEIDSLIV



Kapittel 5

Arbeidsliv

Det norske arbeidslivet er inne i en digital omstilling, der KI-baserte systemer og automatisering preger og forandrer stadig flere arbeidstakeres hverdag. Samtidig som arbeidsoppgaver erstattes av maskiner og noen yrker forsvinner, skapes det også nye muligheter og nye jobber. Flere store norske teknologiselskaper og arbeidslivsorganisasjoner har gått sammen om en felles «Erklæring om ansvarlig bruk av kunstig intelligens i arbeidslivet», der det understrekes at beslutnings- og styringsprosesser som er støttet av systemer med innslag av KI må skje på en ansvarlig måte (NITO 2019). For å sikre en ansvarlig digitalisering i det norske arbeidslivet, presiserer erklæringen at beslutninger og anbefalinger med støtte av KI må være rettferdige, ikke-diskriminerende og transparente. Det betyr at KI ikke må

diskriminere mellom personer på grunn av etnisk opprinnelse, religion, kjønn, seksuell orientering eller nedsatt funksjonsevne. Datagrunnlaget, systemkonstruksjonen og bruken av KI må være så nøytral og så fordomsfri som mulig.

I dette kapitlet ser vi på forskning som kan kaste lys over hva den digitale omstillingen med KI-teknologi og automatisering har å si for likestillingen i arbeidslivet. Hvordan blir arbeidsforholdene påvirket av den digitale omstillingen? Fører kunstig intelligens og automatisering av arbeidsoppgaver og beslutningssystemer til mer likestilling i arbeidslivet, eller til økte sosiale forskjeller? Hvilke arbeidstakere er mest utsatt for å bli erstattet av maskiner?

Forskningsstatus

Som tidligere nevnt har internasjonal forskning påvist diskriminerende effekter av KI-systemer som peker ut de beste jobbsøkerkandidatene ut fra hva algoritmer har lært av data fra tidligere ansettelser. Denne forskningen har vist at algoritmer kan lære av fordommer når det gjelder kjønn og hudfarge, og slik konkludere med at hvite menn er de beste kandidatene (se f. eks Dastin 2018; Howcroft og Rubery 2019; West, Whittaker og Crawford 2019: 8). Enkelte forskere argumenterer imidlertid for at algoritmer kan brukes for å motvirke utilsiktede diskriminerende effekter av rekrutteringsalgoritmer, dersom utviklerne er bevisste på mulige fordommer og programmerer algoritmene til å fremme likestilling (Peña, Serna og Morales 2020). Og, som nevnt i kapittel 1, nevner Teknologirådets rapport om kunstig intelligens at det norske IKT-selskapet Evry oppnådde sitt mål om flere kvinnelige ansatte etter at de begynte å bruke et KI-system som del av rekrutteringsprosessen. Ifølge Evry skyldtes økningen at systemet baserer utvelgelsen på mer objektive kriterier. Rapporten utdyper imidlertid ikke hva disse objektive kriteriene består i (Teknologirådet 2018: 49).

Vår kartlegging tyder på at det ennå ikke er publisert vitenskapelige artikler som utforsker mulige diskriminerende effekter av rekrutterings-

algoritmer når det gjelder norske forhold. Det finnes imidlertid norsk forskning som undersøker endringer i yrkessammensetningen som følge av automatisering av arbeidsoppgaver, der kjønn er én av flere variabler (Hessel, Christiansen og Skirbekk 2017; Pettersen 2019; Øye 2019). Selv om ikke kjønns- eller likestillingsperspektivet er sentralt i denne forskningen, er den relevant for å analysere hvordan KI og automatisering slår ut i typiske kvinne- eller mannsdominerte yrker.

Et annet relevant forskningsfelt utforsker hvordan arbeidstakere opplever omstilling på arbeidsplassen som følge av automatisering og økt bruk av KI. Med unntak av en rapport som tar for seg likestillingskonsekvenser av melkeroboter på norske gårdsbruk, har vi ikke funnet forskning med kjønnsperspektiv som konkret handler om dette. Derfor har vi utvidet søket og inkludert forskningslitteratur om digitalisering og automatisering av norsk høyteknologisk industri, omtalt som den fjerde industrielle revolusjon eller Industri 4.0, der kjønn er nevnt som variabel (Thun et al. 2019), samt forskning som tar for seg arbeidstakeres opplevelse av digitalisering av arbeidsoppgaver i videre forstand, og der kjønnsperspektivet er mer eller mindre uttalt (Halford et al. 2010; Kane 2020; Dahlstrøm og Hognestad 2016).

I den digitale omstillingen av arbeidslivet er IKT-sektoren viktig, og i dette kapitlet har vi også inkludert forskning som problematiserer at yrkene i

denne sektoren er av de mest kjønnsdelte i Norge (Corneliussen og Seddighi 2019; Mark et al. 2019).

Endringer i yrkessammensetningen som følge av automatisering

Endringer i yrkessammensetningen som følge av automatisering og KI møtes med både optimisme og bekymring, og denne diskusjonen gjenspeiles i forskningen. For mens automatiseringen bidrar til økt produktivitet og skaper nye arbeidsplasser, overtas andre jobber av roboter og maskiner. En annen følge av automatisering er en polarisering av arbeidsmarkedet, dvs. at det blir flere som jobber i både høytlønnsyrker og lavtlønnsyrker og færre i mellomlønnede yrker (Autor, Katz og Kearney 2006; Goos, Manning og Salomons 2009; Asplund et al. 2011). Tall fra OECD (2017) viser at også i Norge har sysselsettingsandelen falt i

mellomlønnede yrker, og at forskjellene mellom høyt- og lavtlønnede økte i perioden 1995-2015.

Forskningen gir ikke entydige svar på hvilke konsekvenser den teknologiske utviklingen har for likestillingen. Internasjonal forskning antyder imidlertid at lønnsgapet mellom kvinner og menn øker som følge av digitalisering av arbeidsoppgaver, enten fordi det er kjønnsforskjeller når det gjelder digital kompetanse (OECD 2018; Howcroft og Rubery 2019) eller endringer i arbeidsforhold som følge av såkalt plattformøkonomi (Kullmann 2018; Altenried 2020).

Plattformøkonomi

Plattformøkonomi, også kalt delingsøkonomi, gig-økonomi (oppdragsøkonomi) eller digital økonomi, går i korte trekk ut på at digitale plattformer, som for eksempel apper eller nettsider, brukes til å formidle ytelser mellom tilbydere og kunder. Ytelsene kan være tjenester, arbeidskraft, eiendeler eller eiendom (Jesnes et al. 2016). Mens noen av ytelsene plattformene formidler er stedbundne, som for eksempel transport eller rengjøringstjenester, foregår andre kun på internett, og kalles i forskningslitteraturen for *crowdwork*. Sistnevnte digitale arbeidsform, som for eksempel kan gå ut på å trene KI-systemer til å gjenkjenne bilder, utføres gjerne i hjemmet eller på internettkafeer verden over, er dårlig betalt og midlertidig. Moritz Altenried (2020) viser hvordan dette «skjulte arbeidet bak KI» innebærer en utnyttning av digitale arbeidere, og særlig kvinner i det globale sør. Annen internasjonal forskning viser at det særlig er kvinner som påtar seg deltidsoppdrag innenfor plattformøkonomien (Webster 2016). Dessuten er det beregnet at kvinner i gjennomsnitt tjener 20 prosent mindre enn menn når de

jobber gjennom Amazons Mechanical Turk plattform, fordi ansvar for hjem og barn gjør arbeidsdagen mer fragmentert (Adams-Prassl 2020).

Når det gjelder den stedbundne delen av plattformøkonomien, har internasjonal forskning vist at noen av de kvinnelige sjåførene som formidler sine tjenester gjennom plattformen Uber får mindre betalt enn mannlige sjåførere. Dette fordi algoritmer bestemmer betalingen ut fra beregning av faktorer som for eksempel tilgjengelighet og rangering fra kunder (Kullmann 2018).

Kjønnsperspektiver blir foreløpig ikke diskutert i forskningen på hva plattformøkonomien har å si for det norske arbeidslivet (Jesnes et al. 2016; Alsos, Jesnes og Øistad 2018). Men vi har identifisert vitenskapelige artikler som undersøker effekten av automatisering for forskjellige yrkesgrupper. Ettersom det norske arbeidsmarkedet fortsatt er relativt kjønnsdelt, gir disse kunnskapsgrunnlag for analyser med kjønnsperspektiv.

Er det arbeidsoppgaver eller hele yrker som forsvinner?

Et sentralt bidrag i forskningen om hvilke yrkesgrupper som vinner og taper på automatisering, er metoden som Oxford-forskerne Carl Benedikt Frey og Michael A. Osborne (2013) har utviklet for å beregne sannsynligheten for automatisering. Med utgangspunkt i 702 yrker, fant de at 47 prosent av jobbene i USA sto i fare for å bli automatisert bort. Ifølge OECD er dette en over-

estimering fordi beregningsmetoden har som premiss at hele yrkesgrupper, snarere enn arbeidsoppgaver, blir erstattet av teknologi (Arntz et al. 2016). Mange arbeidsoppgaver er allerede blitt automatisert, mens andre oppgaver, som for eksempel fordrer sosial samhandling og komplekse problemløsning, ikke lar seg automatisere. Med sin tilnærming basert på arbeidsoppgaver,

anslo Arntz et al. (2016) at 10 prosent av yrkene i 21 OECD-land er i risikozonen for å forsvinne som følge av automatisering. Andelen på 10 prosent gjaldt også for Norge. En ny OECD-rapport fra 2018 ned-

justerte anslaget for hvor mange norske arbeidsplasser som vil forsvinne som følge av automatisering til 6 prosent, som er det laveste antallet for alle landene i OECD (Nedelkoska og Quintini 2018).

Hvem er mest utsatt for automatisering?

Frey og Osbornes metode er også blitt brukt for å beregne automatiseringssannsynligheten for yrker i Norge. Mika Pajarinen, Petri Rouvinen og Anders Ekeland (2015) tok utgangspunkt i 374 forskjellige yrker og anslo at én av tre jobber i Norge er i høyrisikograppa for automatisering de neste 10-20 årene. Mest truet er lavtlønnsyrker og yrker som krever liten utdanning, og ifølge forskerne er yrker i privat sektor mer utsatt enn i offentlig sektor. Studien inkluderer også kjønn som variabel, og finner at menn har noe større risiko for automatisering enn kvinner (36 prosent av menn kontra 30 prosent av kvinner er i høyrisikograppa). Studien forklarer ikke hvorfor det er slik, men ettersom den dokumenterer at yrker innen helse og utdanning er minst utsatt for å forsvinne som følge av automatisering, er det nærliggende å anta at overrepresentasjonen av kvinner i offentlig sektor og helse- og omsorgssektoren kan spille inn. Tilsvarende er flere menn enn kvinner ansatt i privat sektor,

der automatiseringen ifølge Pajarinen, Rouvinen og Ekeland utgjør en trussel mot flere yrker.

Påstanden om at lavt utdannede er mest utsatt for å forsvinne som følge av automatisering, må nyanseres. Maskinlæringsalgoritmer blir brukt til å utvikle spesialisert kunnskap på et fagfelt, som for eksempel juss. Internasjonalt diskuteres det problemstillinger knyttet til såkalt *robo-lawyering*, både når det gjelder effekter på sysselsettingen av (høyt utdannede) jurister og brukernes rettsikkerhet (Simshaw 2019). Samtidig påpeker forskningen at det oppstår nye serviceyrker som følge av automatisering. Digitale plattformer gjør at privatpersoner kan tilby varer og tjenester og nå et større marked til lave kostnader. Den såkalte plattformøkonomien kan føre til at andelen selvstendig næringsdrivende og frilansere vil vokse, og at faste ansettelser, og dermed fagorganisering, blir mindre vanlig (Øye 2019: 23).

Færre jobber i automatiserbare yrker

Typisk kvinnedominerte yrker i helse- og omsorgssektoren er i liten grad truet av KI og automatisering. Mens andre kvinnedominerte yrker, som for eksempel kontor- og regnskapsmedarbeidere, er blant de mest utsatte. Det gjelder også tradisjonelt mannsdominerte yrker som for eksempel lagermedarbeidere og anleggsmaskinførere (Asplund et al. 2011; Pajarinen et al. 2015).

Dana Darja Øye (2019) har brukt Frey og Osbornes metode for å beregne forholdet mellom automatiseringssannsynligheter og yrkessammensetningen i Norge. Med utgangspunkt i 348 yrker og med kjønn og utdanning som variabel, har Øye sett på hvilke yrkesgrupper som er mest og minst truet av automatisering. Hun finner at yrker med høy sannsynlighet for automatisering har hatt en nedgang i sysselsetting i perioden 2009 til 2016, og at denne sammenhengen har vært sterkest for kvinner og folk med mellomlang utdanning, slik som kontormedarbeidere. Yrkesgrupper med lav risiko for automatisering har hatt en gjennomsnittlig positiv endring i sysselsettingsandelen, dvs. at det er en økning i sysselsettingen i yrker som i liten grad er truet av automatisering.

På yrkesgruppenivå er hovedkategoriene *Kontor- og serviceyrker* og *Prosess- og maskinoperatører, transportarbeidere mv.* mest utsatt for automatisering, mens kategoriene *Administrative ledere og politikere* og *Akademiske yrker* er minst utsatt. Samtidig er det store variasjoner mellom yrkene under de forskjellige yrkesgruppekategoriene.

Ifølge Øye er de mest automatiseringsutsatte yrkene kontor- og regnskapsmedarbeidere, lagermedarbeidere, anleggsmaskinførere, operatører i næringsmiddelproduksjon og lastebilsjåfører. Minst utsatt er administrerende direktører, et yrke med en mannsandel på 77 prosent, ifølge tall fra SSB (2019). Men også typisk kvinnedominerte yrker som sykepleiere, førskolelærere og grunnskolelærere er lite truet av automatisering.

For å unngå arbeidsledighet og større ulikhet som følge av automatisering, KI og robotteknologi, trekker Øye fram borgerlønn og robotskatt som mulige løsninger for å omskolere arbeidere og utvide utdannings- og helse-sektoren, og dermed skape sikre arbeidsplasser.

Langtidssyke mest utsatt for automatisering

Forskere fra Folkehelseinstituttet har også tatt i bruk Frey og Osbornes beregningsmetode i en studie av sammenhengen mellom helse og risiko for å miste jobben på grunn av automatisering (Hessel, Christiansen og Skirbekk 2018). Med utgangspunkt i den norske levkårsundersøkelsen (n=6393) fra 2015 finner de at langtidssyke har mye større risiko enn andre for å bli «automatisert bort». Kjønnforskjeller har betydning, og langtidssyke menn med lav utdanning ser ut til å være hardere rammet enn kvinner. Selv om færre menn enn kvinner i studien rapporterte at de hadde langvarige helseplager (28,5 prosent menn sammenlignet med 37 prosent kvinner), hadde menn en signifikant større sjanse for å bli erstattet som følge av automatisering (56 prosent gjennomsnittlig sannsynlighet) sammen-

lignet med kvinner (49 prosent gjennomsnittlig sannsynlighet). Langvarig sykdom økte signifikant risikoen for å bli erstattet med 13 prosent for menn, og 11 prosent for kvinner. Kontrollert for utdanning og inntekt ble den relative risikoen for menn redusert, men var fortsatt signifikant. Kontrollert for de samme variablene for kvinner var risikoen ikke lenger signifikant.

Hessel, Christiansen og Skirbekk påpeker at det fra et folkehelsesynspunkt er bekymringsfullt at automatisering kan bidra til å skape større sosial ulikhet, fordi mennesker med dårlig helse og lav utdanning kan bli ekskludert fra arbeidsmarkedet, med den usikkerheten det medfører både økonomisk, sosialt og helsemessig.

KI kan ikke erstatte komplekst kunnskapsarbeid

Mens enkelte arbeidsoppgaver forsvinner som følge av automatisering, er det andre arbeidsoppgaver som ikke kan erstattes. Oppgaver som krever komplekse problemløsninger som ikke har universelle eller generiske løsninger, kan vanskelig erstattes av KI, hevder Lene Pettersen (2019). KI kan ikke programmeres til å etterlikne den menneskelige hjernen fordi KI forutsetter logisk tenking («hvis x, gjør y») og universelle eller generiske svar. På den andre siden forutsetter kompleks problemløsning en form for kunnskap som ikke nødvendigvis er uttalt, men som er intuitiv, relasjonell og kontekstavhengig, og som

vanskelig lar seg programmere. Pettersen er inspirert av filosofen Hubert Dreyfus' kritikk av KI slik den er formulert i klassikeren *What Computers Still Can't Do* (1972/1992). Hun tar utgangspunkt i OECDs beregningsmåte for automatisering, som altså ser på arbeidsoppgaver snarere enn yrkesgrupper. Selv om Pettersen ikke anlegger et kjønnsperspektiv i sin analyse av KI og yrkes sammensetning, er hennes understreking av at yrker som krever sosial samhandling ikke er truet, men snarere kan nyttiggjøre seg av KI-teknologi som arbeidsverktøy, relevant for å forstå omstillingen i arbeidslivet i et likestillingsperspektiv.

Arbeidstakeres erfaringer med automatisering av arbeidsoppgaver

Hvordan erfarer arbeidstakere omstillinger på arbeidsplassen som følge av automatisering og økt bruk av KI? Som nevnt ovenfor, er mulige likestillingskonsekvenser av den digitale omstillingen med vekt på KI-teknologi fortsatt et utforsket tema. Derfor har vi inkludert forskning

på digitalisering av arbeidslivet der kjønn kun nevnes som variabel. Vi ser både på arbeidsforhold i privat sektor, dvs. høyteknologiske industribedrifter, mediebedrifter og landbruket, og i offentlige virksomheter som sykehus og NAV.

Industri 4.0 som ovenfra-ned-revolusjon

Konseptet Industri 4.0 ble lansert av den tyske regjeringen i 2011 som en strategi for å etablere Tyskland som en globalt ledende høyteknologisk innovasjonsnasjon (BMBF [2020]; Nærings- og fiskeridepartementet 2017: 78-79). Industri 4.0, eller «den fjerde industrielle revolusjon», knyttes til storstilt digitalisering og algoritmestyrte automatisering av industrien, der alt

produksjonsutstyr, salg, logistikk og markedsføring er koblet sammen gjennom internett.

Mens den første industrielle revolusjonen på 1800-tallet var karakterisert av arbeidseffektiverende maskineri drevet av vannkraft og dampmaskin, var den andre industrielle revolusjonen på 1900-tallet basert på elektrisitet

og forbrenningsmotoren, mekanisk masseproduksjon og samlebånd på fabrikker. Den tredje industrielle revolusjon fra annen halvdel av 1900-tallet pågår ennå, og knyttes hovedsakelig til digitalisering, elektronikk og informasjonsteknologi. Integrasjon av informasjonsteknologi i såkalt cyberfysiske systemer innebærer at produksjonen kan foregå med betydelig færre ansatte enn før (Garnes og Heggernes 2019).

Den fjerde industrielle revolusjonen er ennå ikke et faktum, men norsk høyteknologisk industri har allerede satt i gang såkalt Industri 4.0-omstilling med digitalisering, automatisering, robotisering og datadeling. Ifølge en forskergruppe fra SINTEF er de digitale verktøyene som er basert på Industri 4.0-konseptet først og fremst tilpasset behovene til ledelsen snarere enn til

arbeiderne. De hevder at Industri 4.0 foreløpig er en ovenfra-ned-revolusjon (Thun et al. 2019).

SINTEF-forskerne har undersøkt 10 norske industribedrifter og analysert svarene fra 417 informanter, hvorav 246 produksjonsledere (kvinneandel på 8,5 prosent) og 171 produksjonsoperatører (12,3 prosent kvinneandel). Studien analyserer ikke eventuelle kjønnsforskjeller blant informantene, men finner at ledere var mer fornøyde med digitale systemer og verktøy enn operatørene. Industri 4.0-visjonen er for viktig til å overlates til mellomlederes og analytikerers skrivebord, hevder forskerne. De argumenterer for at omstillingene må forankres blant operatørene «på gulvet», ved at de digitale verktøyene støtter operatørenes arbeid og gir dem mer medbestemmelse og autonomi i arbeidet (Thun et al. 2019: 54).

Journalisters erfaringer med nedbemanning som følge av digitalisering

På mange arbeidsplasser fører den teknologiske utviklingen til nedbemanning. I mediebransjen har digitalisering og kostnadsuttømt ført til en endring av journalistikken i retning såkalt nettverksjournalistikk (Heinrich 2012), som innebærer at brukerne har større innflytelse, at journalistikken styres av skreddersøm, involvering og tilrettelegging, og at markedsføring i stor grad skjer gjennom Facebook og Googles algoritmer. Ved at publikum har tilgang til nyheter på telefonen når som helst, forventes det at nyhetene blir produsert hurtig. Digitaliseringen gjør at journalistene kan gjøre jobben mer effektivt, samtidig har nedbemanning gjort at færre journalister skal fylle flere spalter, sendinger og nettsider (Dahlstrøm og Hognestad 2016). I 2013 viste Medieundersøkelsen at de aller fleste journalister og redaktører fryktet at nedbemanningen vil gi dårligere kvalitet på journalistikken, både når det gjaldt kilde sjekk, språk, faktasjekk og mengden egenprodusert stoff (Nordiske Mediedager 2013: 15).

I 2015 undersøkte Hilde Kristin Dahlstrøm og Liv Iren Hognestad hvordan journalister i nedbemannede redaksjoner i Agder og Sør-Rogaland (n=282) opplevde at nedbemanning påvirker og utfordrer journalistikken fra idé til ferdig produkt. Et stort flertall av journalistene som svarte, opplevde at arbeidsmiljøet ble negativt påvirket, at det var blitt mer press, dårligere økonomiske rammer og mindre tid til grundig kildearbeid (Dahlstrøm og Hognestad 2016). Studien inkluderte kjønn som variabel, og det viste seg at langt flere kvinner (81 prosent) enn menn (61 prosent) mente at arbeidsmiljøet hadde blitt dårligere. Undersøkelsen sier ikke noe om hvorfor det er slik, om det for eksempel skyldes at kvinnelige journalister er mer sårbare for nedbemanning eller er mer misfornøyde med at journalistikken er blitt mer markedsstilpasset og styrt av «klikkvinnere» enn det mannlige journalistene er. Dahlstrøm og Hognestad påpeker at kjønnsforskjellene er såpass store at det trengs mer forskning for å identifisere årsaker.

Melkerobot gir ikke mer likestilling på gården

Forsker Renate Marie Butli Hårstad fra Ruralis – Institutt for rural- og regionalforskning har undersøkt hvordan melkeroboten påvirker bondens hverdag og likestillingen i familiene (Hårstad 2019). Tidligere forskning har vist at ny teknologi har ført til en maskulinisering av bondeyrket (Brandth 1993; Bjørkhaug og Blekesaune 2007; Heggem 2014). Tradisjonelt var for eksempel melking kvinnearbeid, men med melkemaskinens inntog i etterkrigstida ble det mannfolkarbeid.

Melkerroboten er et automatisk melkesystem som brukes i løsdriftsfjøs, der kyrne går fritt og lar seg melke av en robotarm når de selv vil. Roboten erstatter melkemaskinen som krever manuell betjening av bonden, lar kua bli melket opptil fire ganger i døgnet og er i drift døgnet rundt. Et avansert datasystem registrerer all helseinformasjon om kua og melka som blir produsert. Melkeroboten ble tatt i bruk på 1990-tallet, og bruken har økt etter 2005. I dag

har omtrent halvparten av melka som blir produsert i Norge vært gjennom en melkerobot.

Hårstads studie er basert på data fra en nasjonal spørreundersøkelse utført av Ruralis i årsskiftet 2017/2018 og følger opp tidligere kvalitative studier om melkeroboter. Spørreskjemaet ble sendt til alle de 1700 melkebøndene som i 2017 hadde melkerobot i Norge, samt til en kontrollgruppe med ca. 1700 melkebønder som ikke hadde det. Svarprosenten var 38 prosent, og kjønnsfordelingen blant respondentene var den samme som blant melkeprodusenter for øvrig, nemlig 15 prosent kvinner og 85 prosent menn (Hårstad 2019: 19). Installasjon av melkerobot fordelte seg omtrent likt mellom kjønnene.

Rapporten viser at bønder investerer i melkerobot for å få mer fleksibilitet, lette det fysiske arbeidet og redusere arbeidsmengden. Fordi melkeroboten gjør at bønder slipper å melke og føre til faste tider, får de en mer fleksibel og «moderne livsstil». For eksempel kan de til-

bringe tid med familien på tider som ikke var mulig før, og følge opp barnas fritidsaktiviteter.

Frigjøringen av tid som melkeroboten i fjøset gir, kan gi mer kjønnslikestilling når det gjelder arbeidsdelingen på gården. Men, ifølge studien, bruker mannlige bønder den frigitte tida på merarbeid på gården og ikke på mer husarbeid og familieliv. Kvinnelige bønder oppgir på sin side at fleksibiliteten som følger med melkeroboten er viktig for at de skal greie å kombinere barn og familieliv med arbeid som bonde. De gjør ikke mindre husarbeid selv om de har melkerobot.

Melkeroboten er ikke nødvendigvis en drivkraft for likestilling på norske gårder, verken i fjøset eller i hjemmet, ifølge Hårstad (2019: 11). Selv om melkeroboten har endret gårdsdriften, har den ikke endret kjønnsrollene på gården. Ny teknologi som gir mer fleksibilitet fører ikke automatisk til mer likestilling. Uansett om bøndene har melkerobot eller ikke, fortsetter kvinnelige bønder å gjøre mer husarbeid enn mannlige bønder, ifølge Hårstad.

Kjønnsrelasjoner i møte med ny teknologi på arbeidsplassen

Hvordan digitalisering av arbeidsoppgaver kan bidra til både å reprodusere og å utfordre kjønne maktforhold på arbeidsplassen, er tema for en kvalitativ studie som ble utført på et norsk universitetssykehus i forbindelse med innføringen av elektroniske pasientjournaler (Halford et al. 2010). Studien er basert på observasjoner og intervjuer med 32 sykehusansatte i perioden 2004 til 2008. Elektroniske pasientjournaler (EPJ) lagrer opplysninger om pasienter digitalt i egne IT-systemer med høye sikkerhetskrav, og er altså et digitalt verktøy som ikke faller inn under definisjonen av kunstig intelligens. Studien er tatt med her fordi kjønne aspekter ved implementering av ny teknologi på arbeidsplassen er i sentrum for analysen. Arbeidsoppgaver og

hierarkier på sykehus har tradisjonelt vært kjønne, der sykepleie er assosiert med kvinnelighet og legegjerningen med maskuline verdier. I studien fant forskerne at implementeringen av EPJ kunne forsterke kjønnsnormer ved at sykepleiere ble sett som dumme og med behov for grundig opplæring i det nye verktøyet, mens legene ble tiltrodd kompetanse til å lære seg det på egenhånd. Samtidig ble kjønnsnormene utfordret ved at sykepleiere i mange tilfeller ble mer kompetente enn legene i EPJ. Teknologisk kompetanse assosieres med maskulinitet, og ved at sykepleiere tilegnet seg denne kunnskapen bedre enn legene, kunne de kjønne maktforholdene på sykehuset destabiliseres, ifølge forskerne.

Skjermbyråkrati i NAV gjør at ansatte ikke gir brukere god nok hjelp

I Norge har automatiseringen av velferdsforvaltningen kommet lenger enn i de fleste andre land (Hansen, Lundberg og Syltevik 2018). Arbeids- og velferdsforvaltningen NAV har allerede innført automatisert saksbehandling der KI-systemer fattar vedtak. I forrige kapittel om tjenester så vi på forskning om dette fra et brukerperspektiv, mens vi her ser på studier om hvordan ansatte i NAV opplever digitaliseringen av arbeidshverdagen. Vi har ikke identifisert forskningslitteratur som legger vekt på kjønnsperspektivet når det gjelder NAV-ansattes erfaringer med digitalisering.

Likevel er to studier om utfordringer knyttet til økt digitalisering av velferdsforvaltningen fra et ansattperspektiv, og der kjønn nevnes som variabel, relevante (Røhnebæk 2016; Kane 2020).

Veiledere på NAV er etter loven pliktige til å gi brukerne individuell tilpasning i velferdstjenesteytingen. Gjennom en etnografisk studie av informasjonssystemenes rolle i utvikling av veilederrollen i NAV, finner Maria Røhnebæk (2016) at de ansattes handlingsrom for individuelle vurderinger er blitt begrenset av deres digitale arbeidshverdag. Ikke

fordi digitaliseringen i seg selv undergraver de ansattes utøvelse av skjønn og individuell tilpasning av tjenestene for brukerne, men fordi «skjermyråkratiet» gir ledelsen økte muligheter for rigid detaljstyring. Ledelsens digitale detaljstyring gjør at de ansatte mister kontroll over arbeidet og undergraver den profesjonelle skjønnsutøvelsen (Røhnebak 2016: 202).

At digitaliseringen av NAVs brukerrettede arbeid kan føre til at enkelte brukere ikke får de ytelsene de har krav på, bekymrer også NAV-ansatte. I en kvalitativ undersøkelse av hvordan NAV-ansatte

opplever sitt arbeid med å fatte vedtak om brukernes bistandsbehov, finner Aina Aune Kane (2020) at idealene om forsvarlige kartleggings- og beslutningsprosesser, brukermedvirkning og målrettet og adekvat bistand, blir utfordret av begrensede tidsressurser og digitale løsninger som ikke i tilstrekkelig grad gir rom for relevant informasjon. De digitale skjemaene med forhåndsbestemte spørsmål og svaralternativer dekker ikke alltid den enkelte brukerens sammensatte bistandsbehov, slik at brukerne risikerer å bli satt i feil innsatsnivå og dermed blir stående uten bistand de har rett til.

Manglende kjønnslikestilling i IKT-bransjen som problem

Utviklingen av kunstig intelligente systemer er ikke nøytral, men påvirkes både av kvaliteten på datasettene systemene trenes på og hvem det er som koder datasettene, programmerer algoritmene og lager systemene. IKT-sektoren er et av de mest kjønnsdelte fagområdene i Norge, og den lave kvinneandelen er blitt problematisert ettersom samfunnet blir stadig mer digitalisert (Corneliusen og Seddighi 2019; Mark et al. 2019; OECD 2018).

Ifølge en studie gjort av en forskergruppe fra Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU) er det nødvendig å øke kvinneandelen i IKT-sektoren for å styrke IKT-sikkerhetskompetansen i møte med utfordringene som Tingenes internett (*Internet of Things*, IoT) bringer med seg (Mark et al. 2019). IoT kobler sammen den fysiske og digitale verden gjennom såkalt smarte enheter, overvåkingsutstyr, nettforbindelser, data og analyse. IoT gir praktiske løsninger, men utgjør også en sikkerhetsrisiko fordi enhetene gir mulige innganger for cyberkriminelle. IKT-kriminalitet skjer i stadig større omfang fordi informasjonen kan manipuleres og misbrukes. Eksempelvis kan kriminelle ta kontrollen over selvkjørende biler og bruke dem i terrorhandlinger, eller de kan ta kontroll over smarthustek-

nologi for å kreve løsepenger. Andre eksempler er identitetstyveri og overtakelse av bankkonti.

Ifølge NIFU-forskernes studie er det norske samfunnet langt fra godt nok rustet til å møte sikkerhetsrisikoene knyttet til IoT-enheter (Mark et al. 2019). Studien, som bygger på en kvantitativ framskriving av arbeidslivets framtidige behov og kvalitative intervjuer med aktører på feltet IKT-sikkerhet, beregner at det i 2030 vil være en underdekning på 4100 personer med adekvat IKT-sikkerhetskompetanse i Norge.

For å heve kompetansen i IKT-sikkerhet, argumenterer forskerne for at det er nødvendig å styrke utdanningen og øke kvinneandelen på fagfeltet. Tall fra NSDs Database for statistikk om høgre utdanning (DHB) viser imidlertid at kvinneandelen blant studentene i IKT faktisk har gått ned de siste årene. I 2012 var 17,8 prosent av studentene i IKT-sikkerhet kvinner, mens kvinneandelen i 2017 hadde sunket til 13,4 prosent (Mark et al. 2019: 184). NIFU-forskerne argumenterer for at utdanningene i IKT-sikkerhet må gjøres mer attraktive for kvinner fordi en bedre kjønnsbalanse bidrar til å heve IKT-sikkerhetskompetansen i Norge.

Lite motivasjon for likestillingstiltak i IKT-sektoren

Selv om det er et uttalt mål med en høyere andel kvinner i IKT-yrker, viser en kvalitativ studie av 12 norske bedrifter, institusjoner og organisasjoner som arbeider med IKT i forskning, utvikling og innovasjon, at det skorter på praktiske tiltak for å øke kvinneandelen (Corneliusen og Seddighi 2019). I denne kvalitative studien fant Hilde G. Corneliusen og Gilda Seddighi at det i de 12 bedriftene og organisasjonene finnes kulturelle forestillinger og stereotyper om at menn passer best til tekniske sider ved IKT-arbeid, mens kvin-

ner assosieres med kommunikasjon og sosiale ferdigheter. Selv om organisasjonene i studien anerkjente at det var et behov for å rekruttere kvinner til IKT-arbeid, bidro slike stereotypiske forestillinger til at ingen av organisasjonene hadde strategier for praktisk likestillingsarbeid.

I de undersøkte bedriftene og organisasjonene var det en gjengs holdning om at den lave kvinneandelen i IKT var dels samfunnsskapt, dels skyldtes naturlige og medfødte forskjeller mellom

kvinner og menn, og at praktisk likestillingsarbeid var enten unødvendig, uønsket eller umulig. Corneliussen og Seddighi konkluderer med at disse tilnærmingene til likestilling i bedriftene og

organisasjonene «bidrar til å bremse likestilling i IKT-arbeid, gjennom å uthule motivasjon, vilje og tro på at det nytter å engasjere seg for å skape likestilling i IKT.» (Corneliussen og Seddighi 2019: 284).

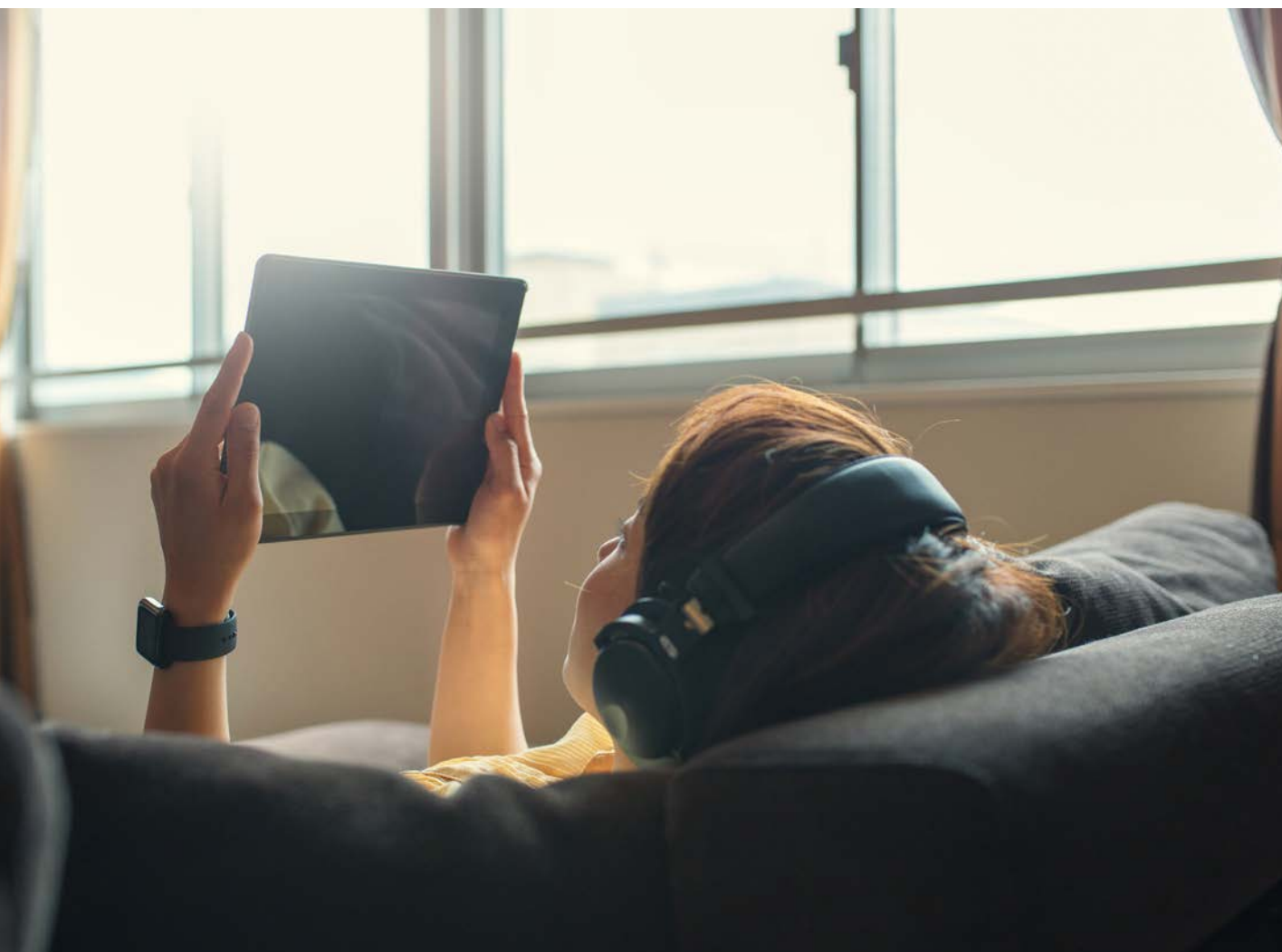
Oppsummering

I dette kapitlet har vi kartlagt norsk forskning som undersøker sosiale konsekvenser, herunder likestillingskonsekvenser, av økt bruk av kunstig intelligente systemer som del av omstillingen i arbeidslivet. Vi har sett på forskning om endringer i yrkessammensetningen og av arbeidsoppgaver som følge av automatisering, og på forskning om medbestemmelse på arbeidsplassen. Forskningen viser at tradisjonelt kvinnedominerte yrker som kontormedarbeidere allerede er automatisert bort, mens andre kvinnedominerte jobber innen-

for helse, utdanning og omsorg kan nyte godt av KI-teknologi som hjelpemidler. Det kommer også fram at en del tradisjonelt mannsdominerte jobber i transport og industri er mest utsatt for å forsvinne som følge av automatisering. Flere studier viser at automatisering fører til en polarisering av arbeidslivet, både når det gjelder inntekt og arbeidstakeres rettigheter. Digitale løsninger kan gi ledelsen økte muligheter for rigid detaljstyring, og den såkalte plattformøkonomien innebærer utfordringer for den norske arbeidslivsmodellen.

KAPITTEL 6

DIGITALE SOSIALE PLATTFORMER OG UNDERHOLDNING



Kapittel 6

Digitale sosiale plattformer og underholdning

De fleste av oss benytter KI-basert teknologi bevisst eller ubevisst i hverdagen. Gjennom ansikts-gjenkjenning får vi enkelt åpnet smarttelefoner, vi får relevante treff i et Google-søk, og vi kan stille spørsmål og gi kommandoer til en smarthøytaler. Algoritmer gjør dataprogrammer i stand til å se mønstre i både bilder, tekst og video som brukere deler på for eksempel sosiale digitale plattformer som Facebook, Instagram, Twitter og TikTok, og kan deretter anbefale personlig tilpasset innhold. Dermed påvirker KI-systemene også hva vi blir anbefalt å lese, lytte til og kjøpe. Også flere digitale spill er understøttet av automatiserte systemer.

Mange, særlig unge, bruker mye tid på sosiale plattformer og digitale underholdningstjenester. 73 prosent av befolkningen i aldersgruppa 16 til 79 år bruker sosiale medier som Facebook, Messenger, Snapchat og Instagram daglig eller

nesten daglig, og omtrent halvparten i samme aldersgruppe har brukt internett til å kjøpe og lytte til musikk (SSB 2020a). Om lag 35 prosent spiller digitale spill hver dag (SSB 2020b). Som diskutert tidligere i rapporten, finnes flere eksempler på at denne typen teknologi kan ha diskriminerende effekter, og dette gjør den interessant og relevant i et likestillingsperspektiv.

Denne delen av kartleggingsoversikten tar for seg forskningen som sier noe om mulige likestillingskonsekvenser av stadig økt bruk av KI-teknologi i hverdagen, blant annet på digitale sosiale plattformer. Sier forskningen noe om hvorvidt KI-teknologi kan reprodusere konvensjonelle kjønnsroller og kjønnsstereotyper? Eller om KI kan være et verktøy for å identifisere kjønnsbasert hatefullt innhold på nettet og dermed ha positive likestillingskonsekvenser?

Forskningsstatus

Vår kartlegging viser at norsk forskning på kunstig intelligens i sosiale digitale plattformer og underholdning i liten grad vektlegger et kjønns- og likestillingsperspektiv. Likevel har det lyktes oss å finne vitenskapelige publikasjoner som til en viss grad belyser forholdet mellom likestilling og kunstig intelligens på dette samfunnsområdet. Dette inkluderer en empirisk undersøkelse av den norske befolkningens kjennskap og holdninger til algoritmer (Gran, Booth og Bucher 2020), en artikkel om KI i musikk- og strømmetjenester (Kiberg 2018) og en artikkel som kartlegger menn og kvinners ulike holdninger til roboter og sexroboter (Nordmo, Næss, Husøy og Arnestad 2020).

Vi har også valgt å inkludere to artikler som handler om kjønn i dataspill. Ulike digitale spill bygger i varierende grad på former for KI-teknologi, i hovedsak regelbaserte «hvis x, gjør y»-systemer, og er en svært kjønnnet aktivitet (Ask, Svendsen og Karlstrøm 2016; Arneberg og Hegna 2018). Årsaken til at vi inkluderer publikasjonene om digitale spill er at internasjonale studier har vist at personer som spiller video- eller dataspill har større sjanse for å utvikle en interesse for IKT

(Gil-Juárez, Feliu og Vitores 2018). Dermed kan spill forstås som en sentral læringsarena som kan ha betydning for rekruttering til IKT og teknologiske bransjer hvor kunstig intelligens utvikles – samt at dataspill i seg selv er en viktig del av den digitale samtiden vi lever i, og er større enn både film- og musikkbransjen kombinert.

Som nevnt innledningsvis har forskere innen *kritiske algoritmestudier* argumentert for at samfunnet trenger mer innsikt i algoritmene fordi de påvirker hva vi som brukere møter på nett (Beer 2017; Diakopoulos 2015; Gillespie og Seaver 2016). Også norske forskere har latt seg inspirere av disse studiene, og ifølge Taina Bucher (2012; 2018) er en i denne tradisjonen opptatt av å forstå hva algoritmene *gjør*, framfor en teknisk tilnærming til hva de *er*.

Det framstår som en sentral faglig diskusjon å vurdere på hvilken måte og gjennom hvilke metoder forskere kan frambringe kunnskap om algoritmer og hvordan de påvirker samfunnet. Også studiene som tematiserer KI på sosiale digitale plattformer retter et kritisk blikk på

algoritmenes sorte boks og studerer for eksempel hvordan algoritmene bidrar til å utvelge informasjon som anses for å være relevant for enkeltpersoner, og dermed setter rammene for informasjon, forbruk og deltagelse i det offentlige liv (Bucher 2018, Gillespie 2013). Flere studier påpeker at kunstig intelligente systemer ikke ek-

sisterer i et vakuum, men i en kommersiell logikk (Kalsnes og Pettersen 2019, Kiberg 2019). Med hensyn til metoder har Bucher (2012, 2018) her tatt til orde for såkalt *reverse engineering* (som kan oversettes til omvendt konstruksjon), hvor søkeresultater dekonstrueres trinn for trinn, og at dette kan bidra til bedre innsyn i de sorte boksene.

Mystiske og utilgjengelige algoritmer

I et temanummer av *Norsk medietidsskrift* om algoritmer, automatisering og analysedata i det digitale medielandskapet, argumenterer redaktørene Bente Kalsnes og Lene Pettersen (2019: 3) for at til tross for at algoritmer preger store deler av vår internettbaserte tidsalder, framstår algoritmer, analysedata og automatisering i vår dagligdagse forståelse som mystiske og ikke-tilgjengelige. Redaktørene foretar det de kaller en «avkleddingsmanøver» for å imøtegå noen myter om algoritmer.

Kalsnes og Pettersen understreker at det ikke bare finnes én type, man mange former for algoritmer som har forskjellige funksjoner, og som aldri kan være nøytrale. Ettersom algoritmer er programmert av mennesker, er de også basert på deres normer, erfaringer og kunnskap. Artikkel-

forfatterne viser også til den manglende kjønnsbalansen i IKT-sektoren, og påpeker at de menneskene som vanligvis står bak algoritmene er «unge, hvite menn» (Kalsnes og Pettersen 2019: 3).

Forfatterne vektlegger et kritisk perspektiv på at algoritmer ofte inngår i en kommersiell logikk, og hevder at «selv om algoritmer i bunn og grunn er trinn-for-trinn-instruksjoner for hvordan noe skal gjennomføres, får de først mening, verdier eller konsekvenser når de bygges inn i virksomheters forretningsmodeller» (Kalsnes og Pettersen 2019: 2). De viser til at algoritmer ofte, men ikke nødvendigvis, er bygget på kommersielle plattformer eller tjenester, hvor målet er å få forbrukerne til å bli på plattformen så lenge som mulig.

Har kjønn betydning for kjennskap til algoritmer?

Algoritmer er som nevnt innebygget i nærmest alle de digitale plattformene vi bruker i dag. Dermed kan kjønnsforskjeller når det gjelder kunnskap om disse algoritmene være ett av flere elementer som kan belyse forholdet mellom kjønn, likestilling og kunstig intelligens. Internasjonal forskning har vist at mange forbrukere ikke nødvendigvis er klar over at aktører, som for eksempel Facebook, bruker algoritmer til å velge ut innhold til deres personlige nyhetsstrøm (Eslami et al. 2015; Smith 2018).

Anne-Britt Gran, Peter Booth og Taina Bucher (2020) undersøker i hvilken grad et representativt utvalg (N=1624) av den norske befolkningen har kjennskap til algoritmer, hvilke holdninger de har til algoritmer, og hvordan disse forholdene er påvirket av kjønn, alder, utdanningsnivå og bosted. Respondentene blir bedt om å svare på i hvilken grad de har kjennskap til algoritmer, og hva de mener om at algoritmer for eksempel brukes til å gi musikk anbefalinger, tilpassede annonser og annet innhold på nett. Svaralternativene gis på en skala med fem verdier fra «ingen kjennskap» til «veldig høy kjennskap». Respondentene fikk også muligheten til å utdype svarene med egne ord (Gran et al. 2020: 5).

Resultatene av undersøkelsen viser at 61 prosent av den norske befolkningen rapporterer liten grad eller ingen kjennskap til algoritmer (Gran et al. 2020: 2), og kjennskap til algoritmer har en relativt tydelig kjønnsdimensjon. Det er en signifikant sammenheng mellom kjønn og kjennskap til algoritmer, der menn oppfatter at de har høyere grad av kjennskap enn kvinner. Blant dem som oppgir at de ikke har noen kjennskap til algoritmer, er kvinner i stort flertall. Samtidig svarer flere menn enn kvinner at de mener de har kjennskap og uttalte synspunkter til algoritmiske funksjoner på nett (Gran et al. 2020: 7).

En mer detaljert lesning av studiens funn viser at nesten halvparten (48 prosent) av de kvinnelige respondentene oppga at de ikke hadde noen kjennskap til algoritmer overhodet. Til sammenlikning var det 33 prosent menn som svarte det samme. Når det gjelder den andre siden av skalaen, oppga flere menn (17 prosent) enn kvinner (8 prosent) at de hadde høy eller veldig høy grad av kjennskap til algoritmer. Utdanningsforskjeller ser ikke ut til å kunne forklare disse kjønnsforskjellene. Likevel er den generelle tendensen at utdanning henger tett sammen med

grad av kjennskap til algoritmer, og jo mer kjennskap man har, desto mer kritisk blir man.

Med hensyn til kjønnsforskjellen er det verdt å understreke at studien er basert på selvrapportering, altså at respondentene selv oppgir hvordan de vurderer egne kunnskaper. Slik kan det umiddelbart antas at menn har høyere grad av selvtillit og dermed rapporterer høyere grad av kjennskap enn de reelt har. Samtidig har internasjonale studier undersøkt om kjønn er en avgjørende faktor for overdrevet selvtillit (*overconfidence*) i selvrapportert kunnskap. Gran et al. (2019: 7) viser til at det ikke er et entydig grunnlag som peker mot at dette er tilfellet, og spekulerer ikke i om menn overdriver egen kjennskap til algoritmer.

Forskerne viser til et skille mellom generasjoner med hensyn til kjennskap til algoritmer. Flere eldre enn yngre oppgir at de ikke har noen kjennskap til algoritmer. I tillegg viser studien at folk som bor på landet oppgir å ha mindre kjennskap til algoritmer enn de som er bosatt i urbane strøk (Gran et al. 2019: 8).

Forskerne gjør også interessante funn når det gjelder holdninger til tre typiske algoritmebaserte fenomener; anbefalinger på digitale underholdningstjenester som Spotify og YouTube, algoritmedrevet annonsering og reklame, og algoritmedrevet innhold gjennom personlige nyhetsstrømmer. De finner en statistisk sammenheng mellom kjønn og holdninger til disse algoritmene, hvor menn tenderer mot å ha mer bestemte og

uttalte holdninger til algoritmer, mens kvinner uttrykker seg mer nøytralt eller svarer oftere «*vet ikke*». Tendensen er sterkest når det gjelder holdning til annonsering og reklame, og svakest for algoritmedrevet innhold (Gran et al. 2020: 8-9).

Gjennom en klyngeanalyse, hvor forskerne finner likhetstrekk og mønster i et datamateriale, utvikles også en typologi for internettbrukere. Her klassifiser forskerne respondentene inn i fem ulike grupper basert på både deres kjennskap og holdninger til algoritmer.

Klart flest, 40 prosent, av alle respondentene tilhører den gruppa som forskerne kaller «ubeviste». Kvinner er overrepresentert i denne gruppa og utgjør nærmere 60 prosent av de «ubeviste» (Gran et al. 2019: 11). Samtidig er menn overrepresentert i gruppene som oppgir at de har noe kjennskap, 60 prosent blant «de usikre» og 63 prosent blant de «anerkjennende». Til gjengjeld er disse gruppene en langt mindre del av det totale utvalget. Kjønnfordelingen er lik blant de som klassifiseres som «nøytrale» eller «skeptiske» til algoritmer, mens menn dominerer klart blant de som oppgir at de har god kjennskap til algoritmer og er kritiske (Gran et al. 2020: 12). Gran et al. (2020) drøfter sine funn i lys av den faglige diskusjonen rundt digitale skiller og kritiske perspektiver på algoritmer. De argumenterer for at dersom kjennskap til algoritmer er en forutsetning for deltagelse i norsk samfunnsliv, representerer den manglende kjennskapen et demokratisk problem (Gran et al. 2020: 13).

«Kjønnsbobler» i den algoritmiske musikkulturen

Som nevnt brukes algoritmiske funksjoner til å gi persontilpassede anbefalinger på internett. Strømmetjenester er relativt utbredt i Norge, og mange har erfart å få anbefalt en bestemt artist, film, serie, spilleliste eller låt. 26 prosent av Norges befolkning bruker Spotify til å strøme musikk daglig, og 21 prosent bruker Netflix til å se filmer og serier daglig. YouTube er særlig populært blant unge, og i aldersgruppa 18 til 29 år har syv av ti en brukerprofil, og én av to bruker netjtjenesten daglig (Ipsos 2020).

Selv om musikk- og strømmetjenester kan sies å ha gjort musikk mer tilgjengelig, er de også i ferd med å forandre hvordan vi som konsumenter presenteres for musikk. Samfunnsmessige og kulturelle konsekvenser av en mer digitalisert musikkbransje er blitt et eget forskningsfelt innenfor medievitenskapen. Sentrale spørsmål er hvordan ulike strømmetjenester påvirker hvordan vi forholder oss til musikk og kultur, og hvordan forretnings-

modellene til strømmetjenestene påvirker verdiskapningen i sektoren (se for eksempel IMK 2018).

Forskerne på strømmetjenester er opptatt av at en algoritmisk musikkultur, en musikkultur der algoritmer i stor grad styrer hva publikum blir anbefalt å lytte til, også må ses i et kjønnsperspektiv. Håvard Kiberg (2019) diskuterer hvordan strømmetjenestenes anbefalingsalgoritmer kan stå i fare for å produsere såkalte «kjønnsbobler» der det skilles mellom «gutte- og jentemusikk». Han tar utgangspunkt i hvordan algoritmer påvirker sjangervariasjon og kjønnsrepresentasjon, og viser hvordan de algoritmiske anbefalingssystemer kan skade musikk mangfoldet og forsterke konvensjonelle kjønns mønstre.

Kibergs diskusjon om potensielle kjønnsbobler baserer seg på logikken rundt strømmetjenestenes anbefalingsmekanismer. Disse drives av algoritmiske systemer som er basert på store

datamengder og statistikk. Strømmetjenesten Spotify anbefaler for eksempel både innhold som er rettet mot det brede publikum gjennom for eksempel topplister, men også innhold som er rettet mot hver enkelt bruker. Utfordringen oppstår ifølge Kiberg fordi mennesker har en tendens til å favorisere det som ligner oss selv. Når dette kombineres med at algoritmene er konstruert som prediksjonsmaskiner som skal anta hva brukeren vil høre, kan dette i sum skape informasjonsbobler tilpasset hver enkelt av oss (Kiberg 2019: 6).

Kiberg viser til at det ved opprettelsen av et Spotify-abonnement, en såkalt «cold-start», oppstår et mulig problem. Fordi de algoritmiske anbefalingssystemene ikke har nok informasjon om brukeren, som for eksempel avspillingshistorikk, må systemet ta utgangspunkt i informasjon om alder,

kjønn og bosted som oppgis ved etablering av abonnementet. Basert på de generelle tallene tjenesten allerede har, forsøker systemet å predikere hva en mannlig og kvinnelig lytter liker og ikke liker (Kiberg 2019: 6). Teknisk kan det å registrere eget kjønn ved opprettelsen av et abonnement «dytte» lytteren inn i en såkalt avgrenset informasjonsboble, eller i dette tilfellet, en kjønnsboble.

Forskere på strømmetjenester har anslått at to av tre anbefalingsalgoritmer er basert på menns lytting, og at menn i større grad enn kvinner lytter til artister av eget kjønn. For eksempel mener strømmeforsker Arne Maasø at dette kan lede til at tjenesten påvirker musikkulturen også på et strukturelt nivå, ved at menn får større definisjonsmakt over hva som blir synlig på strømmetjenestene (Kiberg 2019: 6).

Kjønnsfrihet i dataspill: Er alle menn til det motsatte er bevist?

Internasjonal forskning har vist at kvinnelige spillere av dataspill kan møte andre forventninger og holdninger blant medspillere, og at digitale spill er en svært kjønnert teknologi (Eklund 2011, Kafai et al. 2008). I tillegg peker forskning mot at visuell eller verbal seksuell trakassering kan være mer utbredt på internett enn i andre sammenhenger hvor man er ansikt til ansikt (Barak 2005).

Kristine Ask, Stine Helena Bang Svendsen og Henrik Karlstrøm (2016) har undersøkt hvilken rolle seksuell trakassering spiller i menn og kvinners ulike spillopplevelser, kartlagt hvor utbredt seksuell trakassering i spill er, og hvordan det oppfattes blant spillerne selv. De stiller spørsmål om alle spillere kan uttrykke sin kjønnsidentitet fritt, og om det oppstår kjønnsbestemte muligheter eller begrensninger i en spillsituasjon.

Datagrunnlaget for studien er resultatene fra en nettbasert spørreundersøkelse (N=935) og dybdeintervjuer med ekspertspillere (N=9). Alle respondentene var aktive spillere og mellom 18 og 40 år, og utvalget besto av 2/3 menn og 1/3 kvinner. Analysen hadde fem hovedtema; respondentenes spillvaner, bruk av skjellsord forbundet med konkrete diskrimineringsgrunnlag, erfaring med og oppfatninger om seksuelle tilnærmelser i spillet, strategier for å unngå seksuell trakassering og oppfatninger om betydningen av kjønn og seksualitet i onlinespill (Ask et al. 2016: 7).

Studien viser at seksuell trakassering er noe som spillere, og spesielt de kvinnelige, aktivt forholder seg til. Ulike skjellsord er svært utbredt, og nesten 80 prosent av respondentene hadde sett ord som «retard» og «gay» blitt brukt i ulike spillkontekster.

I tillegg er skjellsord forbundet med kvinnekroppen og seksualisert vold også utbredt. Om lag halvparten av respondentene hadde sett ord som «hore», «pussy» og «fytte» bli brukt, og like mange hadde sett bruk av ordet «rape» og uttrykk som inneholdt voldtektstrusler (Ask et al. 2016: 8-9).

Ifølge studien har kvinner oftere sett og opplevd seksuell trakassering i spill, og kvinner reagerer i større grad enn menn negativt på seksualisert språk. Kvinner har også oftere opplevd oppmerksomheten om kjønn og seksualitet som plagsom. Kvinnelige spillere forteller også at de tar ulike forholdsregler for å unngå seksuell trakassering i spillsammenheng. En slik forholdsregel kan for eksempel være å spille sammen med en mannlig venn eller bekjent (Ask et al. 2016: 16). En annen forholdsregel er å holde kjønn eller seksuell orientering skjult for å unngå uønsket oppmerksomhet fra andre spillere.

Forskerne viser til at den sosiale kontrollen som kvinnelige spillere forholder seg til, kan sammenlignes med hvordan noen homofile og lesbiske skjuler sin seksuelle identitet for å unngå negative reaksjoner. Seksuell trakassering kan dermed ses som å tvinge kvinnelige spillere «inn i skapet», der en unngår å dele informasjon om eget kjønn for å minimere sjansen for sanksjoner. Fordi onlinespill fortsatt gjerne oppfattes som en «guttegreie», er normen at en spiller er mann, inntil det motsatte er bevist (Ask et al. 2016: 17). Dette bidrar til at kvinnelige spillere blir usynlige eller framstår som «rariteter». Blant respondentene hadde om lag halvparten av de kvinnelige respondentene erfaring med å holde kjønn og eller seksuell orientering skjult

for å unngå uønsket oppmerksomhet fra andre. Til tross for disse forholdsreglene, rapporterte flere kvinnelige enn mannlige spillere at de hadde sluttet å spille som følge av uønsket oppmerksomhet.

Seksuelt trakasserende språkbruk er altså utbredt i spill, men spillerne i studien er uenige om språkbruken er et problem, eller om det er et element som skiller spillsituasjonen fra den virkelige verden (Ask et al. 2016: 18). Flere respondenter beskrev den trakasserende språkbruken som en relativt harmløs del av spillet, som en form for «bakgrunnsstøy», og at ord som «gay» ikke hadde samme betydning online som offline. Samtidig opplevde andre spillere, og i særlig grad kvinner, hvordan seksualisert trakassering og

skjellsord forhindret dem fra å «koble av» i spillet.

Ifølge Ask, Svendsen og Karlstrøm er det klare kjønnsforskjeller når det gjelder synet på om seksuell trakassering er et problem i spillkulturen. Mannlige spillere ser i større grad ut til å vektlegge de gode sidene ved aksepten for å bryte sosiale normer i samhandling med andre i spillet, mens flere kvinner forstår verbal og visuell trakassering som et problem. Forskerne konkluderer med at selv om det kan være variasjoner mellom ulike spilltyper, spillkulturer og spillfelleskap, slår de fast at kjønnsbasert og seksuell trakassering ser ut til å kunne begrense kvinners kjønnsfrihet i dataspill (Ask et al. 2016: 18).

Hvem er en ekte gamer?

Som Ask, Svendsen og Karlstrøm (2016) har belyst, eksisterer det ulike forståelser av betydningen av kjønn, seksuell oppmerksomhet og trakassering blant brukere av onlinespill. Edda J. Arneberg og Kristinn Hegna (2018) ser nærmere på forståelsen av kjønn og identitet i onlinespillet *League of Legends*. Arneberg og Hegna tar utgangspunkt i kjønn som en symbolsk grense som ofte innebærer en over- og underordning.

I studien ble et spillnettverk bestående av fem spillere og av begge kjønn (tre menn og to kvinner) valgt som case. Basert på data fra semistrukturerte dybdeintervjuer med de unge voksne spillerne, observasjon, uformelle samtaler og deltagelse i spillet, finner Arneberg og Hegna at det er klare kjønnsstereotyper rundt det å være en «ekte gamer» (2018: 266). De viser til hvordan ulike distinksjoner, stereotyper og klassifiseringer i spillet bidrar til å underbygge en grense som assosierer det kvinnelige negativt.

Ulike hierarkier av spilltyper, spillestiler og ferdigheter definerer hva som anses som en ekte gamer. Blant informantene var noen spill ansett som ekte gamerspill, og disse ga status innad i gruppa. *Le-*

ague of Legends var ansett som et slikt spill. Ulike spillestiler er også, ifølge forskerne, preget av stereotypiske kjønnsrollemønstre, der spillkarakterer med støttefunksjoner ble ansett som passende for kvinnelige spillere (Arneberg og Hegna 2018: 266).

I likhet med Ask et al. (2016) finner også Hegna og Arneberg at flere kvinnelige spillere valgte å skjule sitt kjønn, blant annet gjennom nøytrale kallenavn. Synlig kvinnelige spillere, altså kvinnelige spillere som var åpne om å være kvinne eller valgte tydelig kjønnetrekk ved valg av avatar og/eller navn, ble gjerne forstått som «jentegamere». Det at de viser fram kvinnelige karakteristikk ble i seg selv ansett som å aktivt velge å vise at man var jente ved gå ut av en nøytral rolle (2018: 267).

De grensedragningene og sanksjonene som særlig kvinnelige spillere utsettes for, kan både begrense deres handlingsrom innad i spillet, men kan også ha betydning for deres selvforståelse og mentale helse utenfor spillet. Dermed konkluderer Hegna og Arneberg med at det ikke bør trekkes klare skiller mellom de symbolske kodene som eksisterer i og utenfor spillet.

Menn og kvinners holdninger til sexroboter

Ordet «robot» kan brukes om både fysiske og ikke-fysiske løsninger, og er i ferd med å innta stadig flere samfunnsområder. En fysisk robot kan for eksempel være en industriell robot som kan gjennomføre fysiske oppgaver mennesker ikke kan, eller gjøre dem mer effektivt eller presist. En ikke-fysisk robot kan for eksempel være en aksje-robot, et ikke-fysisk dataprogram som automatisk handler med aksjer. Chatboter er blant de ikke-fysiske robotene som blir stadig vanligere.

Flere forskere har tatt til orde for at roboter kommer til å kunne tilby personlige tjenester og vil inngå i ulike relasjoner med mennesker. Morten Nordmo, Julie Øverbø Næss, Marte Folkestad Husøy og Mads Nordmo Arnestad (2020) reiser spørsmålet om hvordan vi i framtida vil forholde oss til ulike former for intimitet mellom mennesker og roboter, ettersom roboter nå innrettes mot å kunne tilfredsstille mennesker både seksuelt og følelsesmessig. Dette

reiser flere etiske, psykologiske og sosiale problemstillinger om menneske-maskin-relasjoner.

Gjennom et psykologisk eksperiment undersøker Nordmo et al. kjønnsforskjeller i deltakernes holdninger til to typer roboter: Roboter som bare kan utføre seksuelle handlinger og roboter som bare kan være platoniske kjærlighetsroboter (Nordmo et al. 2020: 1). Utvalget besto av 164 kvinner og 114 menn.

Eksperimentet gikk ut på at deltagerne fikk se en illustrasjon med tekst og bilde som presenterte mulighetene for å utvikle enten en fysisk sexrobot eller en ikke-fysisk platonisk kjærlighetsrobot i år 2035. Den ene halvdel av utvalget ble presentert for en tekst som beskrev mulighetene for å utvikle en sexrobot, mens den andre halvdel ble presentert en platonisk robot. Begge tekstene ble illustrert med et bilde. Sexroboten ble

presentert med et bilde av en mannlignende og en kvinnelignende seksualisert robot. Fordi den platoniske roboten var ikke-fysisk, men en digital samtalepartner, var illustrasjonen av den platoniske roboten i form av et bilde av ørepropper. Det ble gjort klart at sexrobotene kun var ment for seksuelt bruk, mens kjærlighetsrobotene ikke kunne stimulere seksuelt, men utforme meningsfylte romantiske eller vennskapelige relasjoner med et menneske (Nordmo et al. 2020: 4).

Nordmo et al. finner at både kvinner og menn i utvalget mente utviklingen av både platoniske og seksuelle roboter var realistisk. Generelt var holdningene til robotene negative, uavhengig av kjønn og type robot. Likevel var kvinner mer skeptiske til roboter, og spesielt sexroboter, enn menn. Kvinner ville også føle større grad av sjalusi om partneren deres anskaffet sexrobot framfor en platonisk kjærlighetsrobot (Nordmo et al. 2020: 8).

Oppsummering

I dette kapitlet har vi sett nærmere på norsk forskning om bruk av kunstig intelligens i uformelle og hverdagslige sammenhenger med et kjønns- og likestillingsperspektiv. Forskningen viser at befolkningens kjennskap og holdninger til algoritmer er påvirket av demografiske variabler som kjønn, alder, utdanning og bosted. Det store flertallet har liten eller ingen kjennskap til algoritmer, og aller minst kjennskap har eldre kvinner med lav utdanning. Innenfor medievitenskapen har forskere utforsket algoritmers påvirkning på

digitale medier, og hvordan de kan påvirke musikk mangfold og kjønnsrepresentasjon. Forskningen tematiserer også seksualitet og seksuelle tilnærmelser i digitale spill og holdninger til trakassering og kjønn i spillmiljø, og viser hvordan seksuell trakassering ser ut til å begrense spillerens kjønnsfrihet og hvordan ulike stereotyper i spill assosierer det «kvinnelige» med noe negativt. Forskningen belyser også kjønnsforskjeller når det gjelder holdninger til seksuell og platonisk intimitet mellom mennesker og roboter.

KAPITTEL 7

DISKUSJON OG AVSLUTNING



Kapittel 7

Diskusjon og avslutning

De etiske prinsippene i regjeringens nasjonale strategi for kunstig intelligens slår fast at KI-systemer skal legge til rette for inkludering, mangfold og likebehandling, og at KI-baserte løsninger skal respektere menneskets selvbestemmelse og kontroll:

«Utvikling og bruk av kunstig intelligens skal bidra til et demokratisk og rettferdig samfunn ved å styrke og fremme enkeltmenneskets grunnleggende friheter og rettigheter. Den enkelte skal ha rett til ikke å være underlagt en automatisert behandling hvis beslutningen systemet fatter berører dem i vesentlig grad. Mennesker skal være inne i beslutningsprosessene for å kvalitetssikre og gi tilbakemelding i alle ledd i prosessen («human-in-the-loop».)» (KMD 2020: 59).

I dette avsluttende kapitlet knytter vi hovedfunnene fra forskningen vi har presentert i de forutgående kapitlene til de politiske målene om at KI-teknologi skal bidra til likebehandling og til et demokratisk og rettferdig samfunn. Som kartleggingen viser er det ikke forsket mye på forholdet mellom kunstig intelligens og likestilling, og det er behov for flere empiriske studier for å tette kunnskapshullene. For hvert av de tematiske områdene vi har tatt for oss; tjenester, arbeidsliv og digitale sosiale plattformer, reiser vi noen problemstillinger som kan inspirere til videre forskning.

KI-baserte tjenester: mer likebehandling eller forskjellsbehandling?

Regjeringen ønsker å legge til rette for økt bruk av kunstig intelligens i offentlig og privat saksbehandling. I forslaget til ny forvaltningslov går Forvaltningslovutvalget inn for mer automatisering av beslutningsprosesser (NOU 2019: 5). Automatisert saksbehandling gir etter utvalgets vurdering effektiviseringsgevinster og bidrar til økt likebehandling. Dette fordi alle som etter systemets kriterier er i samme situasjon, automatisk behandles likt. Systemet kan for eksempel automatisk treffe vedtak som innvilger ytelser når det har registrert at vilkårene er oppfylt. Ifølge utvalget er automatisert saksbehandling særlig til fordel for de «ressurssvake i samfunnet som da ikke behøver å orientere seg om og følge opp det de har krav på» (NOU 2019: 5: 259).

Nasjonal strategi for kunstig intelligens deler Forvaltningsutvalgets syn om at automatisering kan bidra til mer likebehandling og konsistent gjennomføring av regelverket. Samtidig heter det at innføring av saksbehandlingsløsninger med elementer av KI forutsetter at algoritmenes vurderinger er «minst like gode og tillitvekkende som det menneskelige skjønnet de erstatter» og at KI-systemene derfor må være transparente og forklarbare (KMD 2020: 26). Innbyggerne har krav på å vite hvilken informasjon vedtak bygger på, og begrunnelsene må være forståelige. Rettssikkerhet og personvern må ivare-

tas. Det gjelder også bruken av KI-systemer til kontrollformål, for eksempel ved å bruke algoritmiske systemer for å identifisere personer som kan tenkes å bryte loven (KMD 2020: 53).

Regjeringen erkjenner at KI reiser noen dilemmaer (KMD 2020: 2). For hvem har ansvaret for konsekvensene av en beslutning som er truffet av KI? Hva skjer når autonome systemer tar egne beslutninger som ikke er gode og som i verste fall fører til skade? Og hvem skal ta ansvar for at teknologien ikke viderefører og forsterker bevisst og ubevisst diskriminering og forutinntatthet basert på kjønn, alder og etnisitet?

Foreløpig er det gjort få empiriske studier i Norge av samfunnsmessige konsekvenser av at tjenester blir drevet av automatiske løsninger med elementer av KI. Enda færre empiriske studier har et tydelig kjønns- og likestillingsperspektiv. Vi har likevel identifisert forskningsfunn som er relevante når en skal vurdere likestillingskonsekvenser av KI-teknologi i offentlige og private tjenester.

Som nevnt hevder både regjeringen og Forvaltningsutvalget at automatisering av saksbehandling bidrar til mer likebehandling fordi like saker da automatisk behandles likt. Denne positive holdningen finnes også i Kommune-Norge, der foreløpige undersøkelser tyder på at problemstillinger

knyttet til algoritmestyrte forskjellsbehandling i liten grad blir problematisert (Andreasson og Stende 2019). Forskningen legger derimot vekt på at maskinlæringsalgoritmer kan videreføre skjevheter fra dataene de mates med, for eksempel knyttet til kjønn, og at det kan forsterke forskjellsbehandlingen (Kuldova 2020; Seeberg et al 2020; Søraa 2017).

Mens Forvaltningsutvalget framstiller automatisert saksbehandling som særlig fordelaktig for «ressurssvake i samfunnet som da ikke behøver å orientere seg om og følge opp det de har krav på» (NOU 2019: 5: 259), legger den kritiske algo-

ritmeforskningen vekt på at mennesker som berøres av automatiserte beslutninger ikke har innsyn i algoritmene som beslutningene bygger på, og dermed ikke kan vurdere om de er rettfærdige eller ikke. Innsynet i algoritmenes sorte boks kan begrenses med vilje, enten av sikkerhetsmessige grunner eller av kommersielle årsaker, fordi de private selskapene som har utviklet algoritmene ikke vil avsløre forretningshemmeligheter. Det kan utgjøre et demokratisk problem for kommunale og statlige virksomheter når kommersielle aktører får for stor innflytelse på forvaltningen (Andreasson og Stende 2019; Seeberg et al. 2020).

Prediktive analyser, overvåking og kontroll

Et område innenfor offentlig sektor der KI-teknologien forventes å fungere effektiviserende, er i bosetting av flyktninger. Prediktive analyser basert på historiske data om flyktninger vurderes å bli tatt i bruk for å bosette flyktninger i kommuner der algoritmene har beregnet at det er størst sannsynlighet for at de kommer raskt i jobb. På grunn av risikoen for urettferdig behandling, anbefaler forskere at myndighetene avventer å ta i bruk KI-systemer for å bosette flyktninger (Seeberg et al. 2020: 72). I stedet anbefaler de fleksible modeller for bosetting av flyktninger som tar hensyn til alle familiemedlemmers behov og som ivaretar både barneperspektivet og kjønnsperspektivet.

Prediktive analyser basert på KI-systemer vurderes også å bli tatt i bruk av kommuner for å forutse hvilke personer eller bedrifter som har «høy risiko for uønsket utvikling», og som kommunene dermed kan iverksette tidlige tiltak overfor (Andreasson og Stende 2019: 8). Lånkassen har allerede

gjort forsøk med KI for å avsløre studenter som får borteboerstipend selv om de bor hos foreldrene sine. Forsøket viste at algoritmene fant flere falske borteboere enn det Lånkassens saksbehandlere gjorde med tradisjonell stikkkontroll (KMD 2019: 25). Også i politiet, grensekontroll og militæret utvikles KI-basert sikkerhetsteknologi, både i form av overvåking og prediktive analyser for å identifisere mulige kriminelle. Kuldova (2020) advarer mot at samfunnet, i sin iver etter å sikre seg mot svindlere og kriminelle gjennom kontroll og overvåking, går i autoritær retning. Vanlige borgere blir utsatt for overvåking på arbeidsplassen og blir vurdert som potensielle svindlere av algoritmestyrte risikovurderingssystemer av forsikringsselskaper så vel som av utleiefirmaer. Algoritmiske styringssystemer reproduserer etter Kuldovas oppfatning sosiale ulikheter, fordommer og stereotypier, og utgjør en trussel mot enkeltmenneskers frihet og rettigheter.

Digitalt skille mellom brukere av velferdstjenester

At automatiserte beslutningssystemer kan utgjøre en trussel mot enkeltmenneskers rettigheter, blir også tatt opp i forskningen på NAV. NAVs digitale systemer er regelbaserte med forhåndsbestemte svaralternativer, der svarene avgjør hvilket innsatsnivå brukerne skal settes i. For mange NAV-brukere er de digitale løsningene effektive og tidsbesparende, mens brukere som ikke mestrer dem, risikerer å ikke få de ytelsene de etter loven har krav på (Kane 2020). Det er oppstått et nytt såkalt digitalt skille mellom NAV-brukere som mestrer og ikke mestrer de digitale løsningene (Hansen, Lundberg og Syltevik 2018). Dette digitale skillet dreier seg, ifølge forskningen, i liten grad om kjønn, men om utdanningsnivå og alder. Eldre med lav utdanning er mest sårbare for rettighetstap som følge av digitale løsninger,

ifølge Hansen, Lundberg og Syltevik (2018). Men også personer med høy digital kompetanse kan falle ut av systemet på grunn av psykiske problemer eller annen sykdom. Forskningen sier lite om hvorvidt språkkunnskaper også spiller inn, men siden NAVs sider kun er på norsk og engelsk, kan man anta at også personer som ikke behersker disse språkene risikerer å ikke få de ytelsene de har krav på. Ifølge Kane (2020) må ikke NAVs mål om effektiv saksbehandling gå på bekostning av grunnleggende rettigheter.

I NAVs innspill til nasjonal strategi for kunstig intelligens påpekes det at de etiske retningslinjene for statstjenesten (KMD 2017), som alle statsansatte er pålagt å følge, i svært liten grad tar opp utfordringer som kan oppstå når en saksbehandler

enten skal utvikle eller bruke IT-systemer for automatisert saksbehandling (NAV 2019). Statsforvaltningen skal etterstrebe likebe-

handling, men ifølge NAV konkretiserer ikke de etiske retningslinjene hva likebehandling faktisk betyr når det gjelder KI-løsninger.

Kunstig intelligens i helsetjenestene

Helsetjenestene er et av de områdene der det knyttes store forventninger til hva kunstig intelligens kan utrette. Ifølge Nasjonal helse- og sykehusplan 2020-2023 mener regjeringen at bruk av KI sammen med store mengder helsedata kan bidra til bedre og mer persontilpassede helse- og omsorgstjenester og mer effektiv ressursbruk (HOD 2019: 96). I spesialisttjenesten kan KI-baserte systemer brukes i forebygging, diagnostisering, behandling og oppfølging av pasienter, i administrative støttesystemer og logistikk. Kartlegging av risiko for sykdom, tolkning av bilder, kirurgroboter, automatisk generering av pasientjournal, pasientlogistikk og flåtestyring av ambulansetjenesten er oppgaver som kan utføres av autonome KI-baserte verktøy (KMD 2020: 26).

KI-basert helseteknologi kan utvilsomt gi store gevinster, men innføringen av slik teknologi reiser også etiske problemstillinger. Fra å tidligere ha blitt oppfattet som verdinøytrale virkemidler, har utviklingen gått i retning av å forstå helse- og velferdsteknologi som verdibærende, og dermed som gjenstand for etiske og kritiske refleksjoner (Hofmann 2019). Det innebærer også å spørre hvilke pasientgrupper som vil nyte godt av helseteknologien. Ikke bare når det gjelder sykdomstype, men også når det gjelder sosioøkonomiske og kulturelle forhold, samt alder og kjønn.

Kjønnsperspektivet nevnes ikke i verken Nasjonal helse- og sykehusplan 2020-2023 (HOD 2019) eller i *Nasjonal e-helsestrategi* (Direktoratet for e-helse 2019). Vår kartlegging viser at foreløpig er kjønnsperspektiver også relativt fraværende i norsk forskning på KI-basert helseteknologi. Samtidig viser Søraa og Fosch-Villaronga (2020) at kjønnsperspektiver er relevant for robotteknologi i helsevesenet. For eksempel må eksoskjelett, som er beregnet på å hjelpe folk til å gå, være in-

kluderende og ikke ekskludere mennesker som faller utenfor normer om «standardmennesket».

Hege Andreassen, som har forsket på IKT i helse- og omsorgssektoren, hevder at store helsedatabaser kan bidra til å overkomme utfordringene med skjeve utvalg i medisinske forskning, både når det gjelder kjønn, alder, klasse og etnisitet. Samtidig understreker hun at det er en fare for å forsterke eksisterende bevisste og ubevisste skjevheter når avgjørelser overlates til systemer som kun bygger på matematikk og statistikk. Andreassen etterlyser mer forskning på hva digitalisering av helsesektoren får å si for kjønnsforskjeller i helse, for kjønnsdelingen på arbeidsmarkedet, og for yrkesdeltakelsen blant kvinner med omsorgsoppgaver i hjemmet (Andreassen 2019). Kjønnsperspektivet er viktig både når det gjelder konsekvenser av at persontilpasset e-helsetilbud overfører mer ansvar for oppfølging og behandling til den enkelte pasient (Andreassen et al. 2018), og i analysen av helsedata basert på hele populasjonen av pasienter.

Vår kartlegging viser at det er behov for flere empiriske studier av hvordan utviklingen av tjenester med elementer av KI-teknologi påvirker innbyggernes hverdagsliv, og om dette har noe å si for likestilling og sosial ulikhet. Hvilke grupper inkluderes og ekskluderes når KI-teknologi utvikles, utprøves og iverksettes i samfunnet? Kan innføring av nye regelbaserte automatiseringsprosesser og algoritmiske styringssystemer forsterke sosiale forskjeller, fordommer og stereotyper, eller bidra til å styrke likebehandlingsprinsippet i saksbehandling? Videre kan det være relevant å utforske hva KI-teknologi har å si for tilliten i samfunnet og demokratiske prinsipper, som ikke bare inkluderer rettssikkerhet og personvern, men også likestilling mellom kjønnene.

Kunstig intelligens og likestilling i arbeidslivet

Ifølge «Erklæring om ansvarlig bruk av kunstig intelligens i arbeidslivet» trengs det mer kunnskap for å motvirke uønskede effekter av digitaliseringen i arbeidsmarkedet (NITO 2019). Det trengs mer forskning på etiske, samfunnsmessige og økonomiske konsekvenser av KI i arbeidslivet. Internasjonalt etterlyses det mer forskning som undersøker forholdet mellom manglende like-

stilling i arbeidslivet og implementering av KI-baserte systemer (Collett og Dillon 2019; Howcroft og Rubery 2019). Innenfor EU-jussen er en opp-tatt av hvordan anti-diskrimineringslovene kan brukes for å motvirke algoritmisk diskriminering i arbeidslivet, så vel som på andre samfunnsområder (Kullmann 2018; Xenidis og Senden 2020).

Denne kartleggingsoversikten viser at sosiale konsekvenser av KI i arbeidslivet, som også inkluderer likestillings- og kjønnsperspektiver, fortsatt er relativt utforsket i Norge. Det er ikke oppsiktsvekkende ettersom bruken av KI i arbeidslivet fortsatt er i sin spede barndom.

Polarisering av arbeidslivet

Automatisering kan føre til en polarisering i arbeidslivet ved at det blir større økonomiske forskjeller. Allerede nå er det blitt færre mellomlønte ansatte og flere høytlønte og lavtlønte i Norge (Asplund et al. 2011; OECD 2017). Den såkalte plattformøkonomien, der privatpersoner kan tilby tjenester på digitale plattformer, gjør at nye serviceyrker oppstår, men har innvirkning på ansettelsesforhold ved at det blir færre faste anset-

Imidlertid er det gjort relevante forskningsfunn når det gjelder omstilling i arbeidslivet som følge av automatisering, både når det gjelder hvilke yrkesgrupper som er mest og minst utsatt og for medbestemmelse på arbeidsplassen.

telser og flere frilansere, selvstendig næringsdrivende og løsarbeidere, og at fagorganiseringen dermed svekkes (Øye 2019). Usikre arbeidsforhold og ekskludering av personer med lav utdanning og dårlig helse fra arbeidsmarkedet, skaper større sosial ulikhet som igjen har negativ innvirkning på folkehelse (Hessel, Christiansen og Skirbekk 2018), og trolig også for likestillingen.

Likestilling og IKT

Ettersom algoritmer og KI-systemer blir stadig viktigere, øker behovet for flere fagfolk i IKT-sektoren, et fagfelt der kvinner utgjør et mindretall. For å øke IKT-kompetansen i Norge, også når det gjelder sikkerhet, hevder forskere at

det er nødvendig å styrke IKT-utdanningen og å rekruttere flere kvinner til fagfeltet (Mark et al. 2019). Foreløpig ser det imidlertid ikke ut til at konkrete likestillingstiltak blir prioritert i IKT-sektoren (Corneliussen og Seddighi 2019).

Yrkesgruppers utsatthet for automatisering

Når det gjelder spørsmålet om hvilke yrkesgrupper som er mest eller minst utsatt for automatisering og hva dette har å si for likestillingen, er det klart at kjønn alene ikke er en tilstrekkelig variabel. En må også trekke inn variabler som utdanning, kompetanse og alder. I typisk mannsdominerte yrker som administrerende direktør og lagerarbeider har automatiseringen diametralt forskjellig effekt: Mens administrerende direktører tjener på automatisering, står lagerarbeidere i fare for å miste jobben. Typisk kvinnedominerte yrker i helse-, omsorgs- og utdanningssektoren er blant de minst automatiserbare, og for dem kan robotteknologi og KI fungere som nyttige hjelpemidler i jobben. Samtidig har mange av arbeidsoppgavene til kontormedarbeidere, et annet typisk kvinnedominert yrke, allerede blitt automatisert bort (Øye 2019).

Høyere utdanning, der kvinner er i flertall, framheves i forskningen som et vern mot å miste jobben som følge av automatisering (Pajarinen et al. 2015). Men uansett kjønn vil yrker som fordrer kompleks problemløsning og sosial samhandling, ikke kunne erstattes av KI og ro-

botteknologi med det første (Pettersen 2019).

For å motvirke negative konsekvenser av automatiseringen av arbeidslivet, pågår det en diskusjon om å innføre borgerlønn for å sikre inntekten for grupper som faller utenfor arbeidslivet (Øye 2019). Fagbevegelsen går imidlertid imot borgerlønn i form av en flat grunninntekt for alle innbyggere, blant annet fordi den vil være for lav til å kunne leve av og fordi ordningen strider mot den norske arbeidslinja. En annen innvending fra LO-hold er at borgerlønn er en individualistisk løsning på kollektive problemer, som innebærer å tilpasse seg, i stedet for å kjempe imot, at arbeidslivet blir stadig mer usikkert for stadig flere (Bals 2020). I stedet kan kortere arbeidstid være en løsning på å inkludere flere i arbeidslivet (Vigerust 2019). En annen løsning for å unngå arbeidsledighet og større ulikhet som følge av automatisering, KI og robotteknologi, er robot-skatt som kan brukes til å omskolere arbeidere og utvide utdannings- og helsesektoren, og dermed skape sikre arbeidsplasser (Øye 2019: 33).

Gir automatisering mer makt til sjefene?

Hvilke konsekvenser automatisering med KI-baserte systemer har for arbeidsmiljøet er fortsatt relativt uutforsket, men den forskningen vi har kartlagt som studerer dette, peker i retning av at automatisering gagnar ledelsen mer enn de ansatte. I den høyteknologiske industrien ser det ut til at omstilling med digitale verktøy basert på Industri 4.0-konseptet foreløpig er mer tilpasset ledelsen enn arbeidernes behov (Thun et al. 2019). I NAV opplever veilederne at «skjermbyråkratiet» gir ledelsen økte muligheter for rigid detaljstyring. Den digitale detaljstyringen gjør at de ansatte mister kontroll over eget arbeid og muligheten til å utøve skjønn i saksbehandlingen (Rønnebak 2016). I mediebransjen har digitaliseringen ført til nedbemanning og lavere kvalitet på journalistikken, som kvinnelige journalister mer enn mannlige rapporterer at de er misfornøyd med (Dahlstrøm og Hognestad 2016).

En fersk AFI-rapport undersøker hvordan nye teknologiske styrings- og ledelsesformer virker inn på ansattes medbestemmelse og muligheter for å utøve faglig skjønn i fire sektorer: helseforetak, politiet, universitets- og høyskolesektoren samt olje- og gasselskaper (Kuldova et al. 2020). AFI-forskerne finner at standardiserte og automatiserte styringsstrategier fører til at arbeidslivet går i en mer autoritær retning der både ansatte og mellomlederes medbestemmelse på arbeidsplassen svekkes, og dermed mulighetene for bruk av faglig skjønn (se også Underthun og Steen 2018). Internasjonalt er begrepet *algorithmic governance*

blitt tatt i bruk for å karakterisere en styringsmodell som er basert på automatisering, digitalisering og kvantifisering (Katzenbach og Ulbricht 2019). AFI-forskerne finner at i kombinasjon med prinsippene fra New Public Management, utgjør de algoritmiske styringssystemene en trussel mot demokratiet på arbeidsplassen. Fordi algoritmiske styringssystemer ennå ikke er anerkjent som en maktfaktor av partene i arbeidslivet, til tross for at de i økende grad styrer beslutningene der, utgjør de en trussel mot den norske arbeidslivsmodellen og trepartssamarbeidet. Automatiserte beslutninger blir presentert som teknologisk nøytrale og er ikke underlagt demokratiske prosesser, og er dermed vanskeligere for de ansatte å opponere mot. For at trepartssamarbeidet skal fungere, er det nødvendig at fagbevegelsen og de tillitsvalgte har kompetanse om hva digital omstilling innebærer slik at de kan ivareta arbeidernes rettigheter (se også Hagen og Oppegaard 2020).

I forskningen om automatisering og digitalisering av arbeidslivet blir sosial ulikhet ofte trukket fram som negative konsekvenser, og det argumenteres for at omstillingene må forankres blant de ansatte. Likestilling drøftes i liten grad i denne forskningen. Det er behov for mer forskning om hva effekten av automatisering og digitalisering får for kjønnsdelingen på arbeidsmarkedet. Vår gjennomgang tyder også på at det er behov for å undersøke hvordan den såkalte plattformøkonomien virker inn på kjønnsrelasjoner, ulikhet og likestilling i arbeidslivet.

Kunstig intelligens i hverdags- og underholdningsteknologi

Store deler av befolkningen bruker bevisst eller ubevisst KI-teknologi basert på maskinlæringsteknikker og regelbaserte systemer daglig. Vi finner som nevnt teknologien blant annet på sosiale medier, i strømnetjenester, og i digitale spill, og algoritmer styrer i stor grad hva vi blir anbefalt å lese, lytte til og kjøpe. Internasjonalt tematiserer både forskere og politiske aktører som EU og FN hvordan den KI-baserte teknologien er tydelig kjønnet, og hvordan stereotypiske framstillinger av kvinner og menn kan bidra til å reproducere tradisjonelle kjønnsrollemønstre (Adams og Loideain 2019; Woods 2019; West, Kraut og Chew 2019; Gendered Innovations [2020]).

Et eksempel på dette er hvordan digitale assistenter, som for eksempel Siri og Alexa, i hovedsak er utformet med tradisjonelle kvinnelige trekk, og har vist seg å være dårlig rustet til å svare på sexistiske kommentarer som de tolket

som komplimenter (Adams og Loideain 2019; West, Kraut og Chew 2019). EUs rådgivningskomité for like muligheter mellom kvinner og menn, har påpekt at selv om dette kan synes relativt harmløst, bidrar det til å opprettholde og legitimere sexistiske holdninger, noe som også i særlig grad rammer kvinner og jenter på internett (Europarådet 2019: 18; EU-kommisjonen 2020: 6).

Et relatert aspekt er hvordan dette rammer barn og unge som er aktive på digitale sosiale plattformer. For eksempel viste en rapport fra Medietilsynet i 2020 at 43 prosent av ungdommer mellom 13 og 18 år har sett innhold eller diskusjoner på sosiale medier med skremmende eller voldelig innhold, eller hatmeldinger. En høyere andel jenter enn gutter har sett innhold der folk viser fram ulike måter å være tynn på eller måter å skade seg selv (Medietilsynet 2020: 3).

Samtidig påpeker flere at det finnes muligheter innenfor KI-teknologien som kan bidra til å sikre trygg digital deltagelse for alle. For eksempel er det utviklet ulike roboter som

kan identifisere og filtrere ut seksualisert eller hatefullt innhold på sosiale digitale plattformer (Cuthberston et al. 2019; Woodford 2020).

Kjønnsforskjeller i kjennskap til algoritmer

Befolkningens digitale kompetanse er en viktig forutsetning for å kunne nyttiggjøre mulighetene som ligger i digitalisering. Til tross for at Norge sammenlignet med andre land i Europa har høy digital kompetanse (EU-kommisjonen 2016), viser Gran et al. (2020) hvordan store deler av Norges befolkning har svært liten kjennskap til algoritmer. Resultatene viser relativt store kjønnsforskjeller hva angår befolkningens kjennskap til algoritmer og holdninger til hverdagslig bruk av kunstig intelligens. Menn oppgir oftere enn kvinner at de har

kjennskap til og synpunkter på algoritmer. En stor andel kvinner, spesielt eldre og med lite utdanning, oppgir at de ikke har noen kjennskap til algoritmer overhodet. Et annet interessant funn er at jo mer folk vet om algoritmer, som er tett forbundet med utdanning, desto mer kritiske er de til algoritmer (Gran et al. 2020: 12). Manglende kunnskap om algoritmer kan være et demokratisk problem ettersom algoritmer stadig får større betydning for tilgang på informasjon og deltagelse på internett.

Kunstig intelligens i media og musikk-kulturen

Kunstig intelligens og algoritmer har vekket forskningsmessig interesse innenfor medievitenskapen. Sentralt i de faglige diskusjonene er kunnskapsteoretiske (epistemologiske) utfordringer knyttet til algoritmenes sorte boks (Kalsnes og Pettersen 2019). Enkelte forskere foreslår løsninger i form av nye former for eksperimentell humaniora og konkrete framgangsmåter som reverserer dataingeniørers konstruksjoner (*reverse engineering*) (Bucher 2018; Kiberg 2019; Kalsens og Pettersen 2019). Særlig relevant i et likestillingsperspektiv er diskusjonen om hvor-

dan algoritmiske anbefalingssystemer i strømmetjenester kan stå i fare for å standardisere og reproducere konvensjonelle kjønnsrollemønstre. Konkret viser Kiberg (2019) til at KI-funksjonene kan føre til såkalte «kjønnsbobler» hvor systemet forsøker å predikere hva som er typisk jente- og guttemusikk. Ifølge Kiberg er det behov for mer forskning om hvorvidt anbefalingsalgoritmene skader musikk mangfoldet og potensielt påvirker tilveksten av artister innen enten mannsdominerte eller kvinnedominerte sjangre (Kiberg 2019: 8).

Seksualitet og trakassering

Den raske teknologiske utviklingen aktualiserer nye spørsmål om forholdet mellom menneske og maskin, inkludert forholdet mellom menneskelig seksualitet og ny teknologi. Nordmo et al. (2020) diskuterer denne tematikken i lys av KI-basert robotteknologi, og viser til at ettersom roboter som er designet for å tilfredsstille mennesker seksuelt og emosjonelt vil bli kommersialisert, reiser det etiske og psykologiske problemstillinger. Forskerne finner tydelige kjønnsforskjeller i sin undersøkelse, der kvinner er mer skeptiske til roboter generelt, og til sexroboter spesielt.

En annen tilnærming til seksualitet og teknologi er å se på hvordan mennesker uttrykker sin seksualitet, sitt kjønn eller opplever uønsket seksuell oppmerksomhet på digitale sosiale plattformer. Ifølge Arneberg og Hegna (2018) blir såkalte «jentegamere» negativt vurdert, fordi kjønnsrelasjoner i dataspill har en klar over- og

underordning. Arneberg og Hegna hevder at dynamikken rundt kjønn som forekommer i spill, trolig også påvirker spillere i det virkelige liv.

Kjønnsbasert seksuell trakassering forekommer i digitale spill. Forskningen viser at kvinner tar ulike forholdsregler for å unngå uønsket seksuell oppmerksomhet, blant annet ved å skjule at de er kvinner (Ask et al. 2016: 16). Dataspill kan fungere som en digital læringsarena. Blant annet viser en internasjonal studie at jenter som var blitt eksponert for dataspill i ung alder, hadde fire ganger så stor sjanse for å se for seg en karriere innen programmering og koding enn jenter som ikke var det (West, Kraut og Chew 2019: 41). Når Ask et al. (2016: 18) konkluderer med at kjønnsbasert og seksuell trakassering ser ut til å kunne begrense kvinners kjønnsfrihet i dataspill, kan det være nærliggende å anta at dette også påvirker digitale spill som en mulig læringsarena.

Kjønns- og likestillingsperspektiver er ikke fram-tredende i forskningen på hvordan befolkningen tar i bruk KI-baserte verktøy på digitale sosiale plattformer, inklusive underholdningstjenester. Vår litteraturgjennomgang tyder på at det er behov for å forske mer på kjønnsforskjeller når det gjelder kjennskap og holdninger til kunstig intelligens, og hvilke konsekvenser dette har for likestilling. Videre kan det i et likestillingsperspektiv

være nyttig med tverrfaglige studier som utforsker hvorvidt anbefalingssystemer på sosiale digitale plattformer bidrar til å forsterke eller utfordre konvensjonelle kjønnsmønstre og stereotypiske forestillinger om kvinner og menn. Det kan også være hensiktsmessig å utforske og utvikle metoder for hvordan KI-baserte systemer kan brukes til å fange opp kjønnsbasert hatefullt innhold, og dermed skape tryggere digitale rom.

Avslutning

Denne kunnskapsoppsummeringen har kartlagt hva norsk forskning sier om likestillingskonsekvenser av at KI-teknologi får en stadig større innvirkning på befolkningens hverdag, både når det gjelder offentlige og private tjenester, på jobben og i fritida. Målet har vært å svare på følgende problemstillinger: Står kunstig intelligente systemer i fare for å reprodusere, og i verste fall forsterke, de kjønnsforskjellene som eksisterer i samfunnet i dag? Eller kan framskritt innenfor KI-teknologien være et virkemiddel som kan bidra til å oppnå større grad av likestilling og forhindre diskriminering?

Som kartleggingen viser, er det vanskelig å gi entydige svar på disse spørsmålene. Men selv om forskningen vi har kartlagt belyser ulike aspekter ved kunstig intelligens på forskjellige samfunnsområder, går en gjennomgående kritikk ut på at maskinlæringsalgoritmer kan videreføre skjevheter fra dataene de mates med, for eksempel knyttet til kjønn, og at det kan forsterke forskjellsbehandling. Det eksperimenteres med å utvikle KI-baserte verktøy som skal

motvirke diskriminering, men ennå foreligger det ikke norsk forskning som kan underbygge forhåpningene til KI som likestillingsverktøy.

For at teknologi basert på kunstig intelligens skal kunne ha den positive samfunns effekten som mange forventer, viser denne kunnskapsoversikten at det trengs mer kunnskap om hvordan teknologien får konsekvenser for sosial ulikhet og likestilling. Ettersom regjeringen ønsker at Norge skal gå foran i utvikling og bruk av kunstig intelligens med respekt for den enkeltes rettigheter og friheter, bør det utforskes hvordan man kan sikre at teknologien blir et gode for alle.

Generelt peker funnene i denne kartleggingsoversikten mot at det på de tre tematiske områdene vi har behandlet – tjenester, arbeidsliv og digitale sosiale plattformer – er behov for mer tverrfaglig forskning og metodeutvikling for å styrke kunnskapsgrunnlaget om likestillingskonsekvenser av kunstig intelligens. Vår ambisjon er at denne kartleggingen skal inspirere til dette.

Referanser

Artikler og rapporter merket * henviser til de norske studiene som presenteres som egne caser i rapporten.

Adam, Alison (1995). «Artificial intelligence and women's knowledge». *Women's Studies International Forum*, 18(4), pp.407-415. DOI: [https://doi.org/10.1016/0277-5395\(95\)80032-K](https://doi.org/10.1016/0277-5395(95)80032-K)

Adam, Alison (1998). *Artificial knowing: Gender and the thinking machine*. London: Routledge.

Adams Rachel og Nora Ni Loideain (2019). «Addressing Indirect Discrimination and Gender Stereotypes in AI Virtual Personal Assistants: The Role of International Human Rights Law», (22. mai 2019) Annual Cambridge International Law Conference 2019, New Technologies: New Challenges for Democracy and International Law. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3392243%20%20

Adams-Prassl, Abi (2020). «The Gender Wage Gap on an Online Labour Market: The Cost of Interruptions». CEPR Discussion Paper No. DP14294. URL: <https://ssrn.com/abstract=3518636>

Alsos, Kristin, Kristin Jesnes og Beate Sletvold Øistad (2018). «Når sjefen er en app – delingsøkonomi i et arbeidsperspektiv», *Praktisk økonomi & finans* 34(29): 101-110. DOI: <https://doi.org/10.18261/issn.1504-2871-2018-02-03>

Altenried, Moritz (2020). «The platform as factory: Crowdwork and the hidden labour behind artificial intelligence», *Capital & Class* 44(2): 145-158. DOI: <https://doi.org/10.1177/0309816819899410>

Andreassen, Hege (2019). «Digitalisering kan gi ny kunnskap om kvinnehelse – hvis vi gjør det riktig». *Kilden kjønnsforskning*. no, 15.august 2019. URL: <http://kjonnsforskning.no/nb/2019/08/digitalisering-kan-gi-ny-kunnskap-om-kvinnesykdom-mer-hvis-vi-gjor-det-riktig>

Andreassen, Hege K., Kari Dyb, Carl R. May, Catherine J. Pope og Line L. Warth (2018). «Digitized patient-provider interaction: How does it matter? A qualitative meta-synthesis», *Social Science & Medicine* 215: 36-44. DOI: <https://doi.org/10.1016/j.socsci-med.2018.08.036>

Andreasson, Ulf og Truls Stende (2019). *Nordiske kommuners arbeid med kunstig intelligens*. København: Nordisk Ministerråd. URL: <http://norden.diva-portal.org/smash/get/diva2:1365285/FULLTEXT01.pdf> *

Arneberg, Edda J. og Kristinn Hegna (2018). «Virtuelle grenseutfordringer. Symbolske grenser i spill i The League of Legends», *Norsk sosiologisk tidsskrift* 2(3): 259-274. DOI: <https://doi.org/10.18261/issn.2535-2512-2018-03-05> *

Arntz, Melanie, Terry Gregory og Ulrich Zierahn (2016). *The risk of automation for jobs in OECD countries: a comparative analysis*. OECD Social, Employment and Migration Working Papers, No. 189, Paris: OECD Publishing. DOI: <https://doi.org/10.1787/5jlz9h56dvq7-en>

Ask, Kristine, Stine Helena Bang Svendsen og Henrik Karlstrøm (2016). «Når jentene må inn i skapet: Seksuell trakassering og kjønnsfrihet i online dataspill», *Norsk medietidsskrift* 23(1): 1-21. DOI: <https://doi.org/10.18261/issn.0805-9535-2016-01-03> *

Asplund, Rita, Erling Barth, Per Lundborg og Kjersti Misje Nilsen (2011). «Polarization of the Nordic Labour Markets», *Finnish economic papers*, 24(2): 87-110. URL: <https://ideas.repec.org/a/fep/journal/v24y2011i2p87-110.html>

Autor, David H., Lawrence F. Katz og Melissa S. Kearney (2006). «The Polarization of the U.S. Labor Market», *American Economic Review* 96(2): 189-194. DOI: <https://doi.org/10.3386/w11986>

Bals, Jonas (2020). «Borgerlønn og løsarbeid». *Dagsavisen*, 19. august 2020. URL: <https://www.dagsavisen.no/debatt/kommentar/borgerlonn-og-losarbeid-1.1759484>

Beer, David (2016). «The social power of algorithms», *Information Communication and Society* 20(1): 1-13. DOI: <https://doi.org/10.1080/1369118X.2016.1216147>

Bjørkhaug, Hilde og Arild Blekesaune (2007). «Masculinisation or Professionalisation of Norwegian Farm Work - A gender neutral division of work on Norwegian family farms?», *Journal of Comparative Family Studies* 38(3): 423-434. URL: <https://www.jstor.org/stable/41604167?seq=1>

Borgan, Eldrid (2019). «Kunstig intelligens blir mannssjåvinistiske rasister. Hva kan vi gjøre for å stoppe det?». *Forskning*. no, 29. november 2019. URL: <https://forskning.no/arbeid-it-juridiske-fag/kunstig-intelligens-blir-mannssjavinistiske-rasister-hva-kan-vi-gjore-for-a-stoppe-det/1599018>

Brandth, Berit (1993). «Kvinner er ikke lenger hva de var. Kvinnebønder, teknologi og konstruksjon av kjønn». I Brandth, Berit og Berit Verstad (red.), *Kvinneliv i landbruket*, Landbruksforlaget.

Bucher, Taina (2012). «Want to be on the top? Algorithmic power and the threat of invisibility on Facebook», *New Media & Society* 14(7): 1164-1180. DOI: <https://doi.org/10.1177/1461444812440159>

Bucher, Taina (2018). *IF...THEN: Algorithmic power and politics*. New York: Oxford University Press.

Bucher, Taina (2020). «Pensumkvaler: Om avkolonisering, mangfold og akademisk frihet», *Norsk medietidsskrift*, 107(3): 1-7. DOI: <https://doi.org/10.18261/ISSN.0805-9535-2020-03-06>

Buhmann, Alexander, Johannes Paßmann og Christian Fieseler (2020). «Managing Algorithmic Accountability: Balancing Reputational Concerns, Engagement Strategies, and the Potential of Rational Discourse», *Journal of Business Ethics* 193(2): 265-280. DOI: <https://doi.org/10.1007/s10551-019-04226-4>

Bundesministerium für Bildung und Forschung (BMBF) (2020). «Industrie 4.0.» URL: <https://www.bmbf.de/de/zukunftsprojekt-industrie-4-0-848.html>

Buolamwini, Joy og Timnit Gebru (2018). «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification», *Proceedings of Machine Learning Research* 81:1–15. URL: <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

Collett, Clementine og Sarah Dillon (2019). *AI and Gender: Four Proposals for Future Research*. Cambridge: The Leverhulme Centre for the Future of Intelligence. URL: [https://www.repository.cam.ac.uk/bitstream/handle/1810/294360/AI_and_Gender__4_Proposals_for_Future_Research_210619_p8qAu8L%20\(1\).pdf?sequence=1](https://www.repository.cam.ac.uk/bitstream/handle/1810/294360/AI_and_Gender__4_Proposals_for_Future_Research_210619_p8qAu8L%20(1).pdf?sequence=1)

CORE – Senter for likestillingsforskning (2020). «Fremtidens arbeidsliv i et likestillingsperspektiv». Oppdatert 18. februar 2020. URL: <https://www.samfunnsforskning.no/prosjekter/fremtidens-arbeidsliv-i-et-likestillingsperspektiv/index.html>

Corneliusen, Hilde G. og Gilda Seddighi (2019). «Må vi egentlig ha flere kvinner i IKT?» Diskursive forhandlinger om likestilling i IKT-arbeid», *Tidsskrift for Kjønnforskning* 43(4): 273-287. DOI: <https://doi.org/10.18261/issn.1891-1781-2019-04-03> *

Crawford, Kate (2016). «Artificial Intelligence's White Guy Problem», *The New York Times*, 25. juni 2016. URL: <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html>

Cuthbertson, Lana, Alex Kearney, Riley Dawson, Ashia Zawaduk, Eve Cuthbertson, Ann Gordon-Tighe og Kory W. Mathewson (2019). «Women, politics and Twitter: Using machine learning to change the discourse», *arXiv.org*. URL: <https://arxiv.org/abs/1911.11025>

D'Ignazio, Catherine og Lauren Klein (2020). *Data Feminism*. Cambridge, Massachusetts: The MIT Press. URL: <https://data-feminism.mitpress.mit.edu/>

Dahlstrøm, Hilde Kristin og Liv Iren Hognestad (2016). «Nedbemanning som utfordring for journalistikken», *Norsk Medietidsskrift* 22(2): 1-19. DOI: <https://doi.org/10.18261/issn.0805-9535-2016-02-02> *

Dastin, Jeffrey (2018). «Amazon scraps secret AI recruiting tool that showed bias against women», *Reuters*, 10. oktober 2018. URL: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

Datatilsynet (2018). *Kunstig intelligens og personvern*. Oslo: Datatilsynet. URL: <https://www.datatilsynet.no/globalassets/global/dokumenter-pdf/er-skjema-ol/rettigheter-og-plikter/rapporter/rapport-om-ki-og-personvern.pdf>

Datta, Amit, Michael Carl Tschantz og Anupam Datta (2015). «Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination», *arXiv.org*. URL: <https://arxiv.org/abs/1408.6491v2>

Diakopoulos, Nicholas (2015). «Algorithmic Accountability», *Digital Journalism* 3(3): 398-415. DOI: <https://doi.org/10.1080/21670811.2014.976411>

Direktoratet for e-helse (2019). *Nasjonal e-helsestrategi 2017-2022*. Oslo: Direktoratet for e-helse. URL: <https://ehelse.no/Strategi>

Eklund, Lina (2011). «Doing Gender in Cyberspace: The Performance of Gender by Female World of Warcraft Players», *Convergence: The International Journal of Research into New Media Technologies* 17(3): 323–342. DOI: <https://doi.org/10.1177/1354856511406472>

Eriksen, Ingunn Marie, og Idunn Seland (2019). *Ungdom, kjønn og fritid*. NOVA Rapport 6/2019. Oslo: Velferdsforskningsinstituttet NOVA / OsloMet. URL: <https://www.oslomet.no/forskning/forskningsprosjekter/kjoenn-og-unges-fritid>

Eslami, Motahhare, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andu Vuong, Karrie Karahalios, Kevin Hamilton og Christian Sandvig (2015). «I always assumed that I wasn't really that close to [her]: Reasoning about invisible algorithms in news feeds», *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 153–162. ACM Digital Library. DOI: <https://doi.org/10.1145/2702123.2702556>

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau og Sebastian Thrun (2017). «Dermatologist-level classification of skin cancer with deep neural networks», *Nature* 542: 115-118. DOI: <https://dx.doi.org/10.1038/nature21056>

EU-kommisjonen (2020). *Opinion on Artificial Intelligence – opportunities and challenges for gender equality*. Advisory Committee on Equal Opportunities for Women and Men. URL: https://ec.europa.eu/info/sites/info/files/aid_development_cooperation_fundamental_rights/opinion_artificial_intelligence_gender_equality_2020_en.pdf

Europarådet (2019). Preventing and Combating Sexism. Recommendation CM/Rec(2019)1. URL: <https://rm.coe.int/prems-055519-gbr-2573-cmrec-2019-1-web-a5/168093e08c>

European Institute for Gender Equality (EIGE) (2018). Women and men in ICT: A chance for better work-life balance: Research note. Luxembourg: Publications Office of the European Union. URL: <https://op.europa.eu/en/publication-detail/-/publication/dc40b38e-c6b8-11e8-9424-01aa75ed71a1/language-en>

Folkhelseinstituttet (2020). «Systematisk kartleggingsoversikt». Oppdatert 27.02.2020. URL: <https://www.fhi.no/kk/oppsummert-forskning-for-helsetjenesten/kunnskapsoppsummering-produkter/#systematisk-kartleggingsoversikt>

Frey, Carl Benedikt og Michael A. Osborne (2013). The future of employment: how susceptible are jobs to computerization?, Publication Oxford Martin School, University of Oxford. URL: https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Garnes, Åge og Tarjei Alvær Heggernes (2019). «Vi kan ikke spise algoritmer», Beta 33(1): 113-130. DOI: <https://doi.org/10.18261/issn.1504-3134-2019-01-07>

Gendered Innovations [2020]. «Case Study: Machine Translation: Analyzing Gender». URL: <http://genderedinnovations.stanford.edu/case-studies/genderingsocialrobots.html>

Gil-Juárez, Adriana, Joel Feliu og Anna Vitores (2018). «Mutable technology, immutable gender: Qualifying the ‘co-construction of gender and technology’ approach», Women’s Studies International Forum 66: 56-62. DOI: <https://doi.org/10.1016/j.wsif.2017.11.014>

Gillespie, Tarleton (2013). «The relevance of algorithms», i Gillespie, Tarleton, Pablo J. Boczkowski og Kirsten A. Foot (red.) Media technologies: Essays on communication, materiality, and society. MIT Press.

Gillespie, Tarleton og Nick Seaver (2016). «Critical algorithm studies: A reading list». Oppdatert 12. mai 2016. URL: <https://socialmediacollective.org/reading-lists/critical-algorithm-studies/>

Goos, Maarten, Alan Manning og Anna Salomons (2009). «Job Polarization in Europe», American Economic Review, American Economic Association 99(2): 58-63. DOI: <https://doi.org/10.1257/aer.99.2.58>

Gran, Anne-Britt, Peter Booth og Taina Bucher (2020). «To be or not to be algorithm aware: a question of a new digital divide», Information, Communication & Society DOI: <https://doi.org/10.1080/1369118X.2020.1736124> *

Guldbrandsen, Christine (2017). «Rekrutterer ved hjelp av robot – ansetter flere kvinner». Dagens Næringsliv. 7. desember 2017. URL: <https://www.dn.no/arbeidsliv/robotteknologi/utdannelse/evry-evry/rekrutterer-ved-hjelp-av-robot-ansetter-flere-kvinner/2-1-227334>

Hagen, Inger Marie og Sigurd M. Nordli Oppegaard (2020). Partssamarbeid og digital omstilling. Fafo-rapport 2020:07. URL: <https://www.fafo.no/images/pub/2020/20742.pdf>

Halford, Susan, Ann Therese Lotherington, Kari Dyb og Aud Obstfelder (2010). «Un/doing Gender with ICT?», NORA—Nordic Journal of Feminist and Gender Research 18(1): 20-37. DOI: <https://doi.org/10.1080/08038741003626791> *

Hansen, Hans-Tore, Kjetil Lundberg og Liv Johanne Syltevik (2018). «Digitalization, Street-Level Bureaucracy and Welfare Users’ Experiences», Social Policy & Administration 52(1): 67-90. DOI: <https://doi.org/10.1111/spol.12283> *

Harwell, Drew (2018). «Fake-porn videos are being weaponized to harass and humiliate women: ‘Everybody is a potential target’», The Washington Post, 30. desember 2018. URL: <https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/>

Heggem, Reidun (2014). «Diversification and Re-feminisation of Norwegian Farm Properties», Sociologia Ruralis, 54(4): 439-459. DOI: <https://doi.org/10.1111/soru.12044>

Heinrich, Ansgard (2012). «News making as an interactive practice: global news exchange and network journalism». I C. Peters & M. Broersma (red.), Rethinking journalism. Trust and participation in a transformed news landscape. London: Routledge.

Helse- om omsorgsdepartementet (HOD) (2019). «Nasjonal helse- og sykehusplan 2020 – 2023». St. meld. 7 (2019 – 2020). Oslo: Helse- om omsorgsdepartementet. <https://www.regjeringen.no/contentassets/95eec808f0434acf942fca449ca35386/no/pdfs/stm201920200007000dddpdfs.pdf>

Hessel, Phillip, Solveig Christiansen og Vegard Skirbekk (2018). «Poor health as potential risk factor for job loss due to automation: the case of Norway», Occup Environ Med 75(3): 227-230. DOI: <https://doi.org/10.1136/oemed-2017-104349> *

Hofmann, Bjørn (2019). «Hvordan vurdere etiske aspekter ved moderne helse- og velferdsteknologi?», Tidsskrift for omsorgsforskning 5(3): 99-116. DOI: <https://doi.org/10.18261/issn.2387-5984-2019-03-09>

Howcroft, Debra og Jill Rubery (2019). «‘Bias in, Bias out’: gender equality and the future of work debate», Labour & Industry: a journal of the social and economic relations of work, 29(2): 213-227, DOI: <https://doi.org/10.1080/10301763.2019.1619986>

Hunt, Elle (2016). «Tay, Microsoft’s AI chatbot, gets a crash course in racism from Twitter». The Guardian, 24. mars 2016. URL: <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>

Hårstad, Renate Marie Butli (2019). Bonden, familien og melkeroboten - en ny hverdag. Trondheim: Ruralis – Institutt for rural- og regionalforskning. URL: https://ruralis.no/wp-content/uploads/2019/01/rapport-2_19-bonden-familien-og-melkeroboten-en-ny-hverdag--r-m-b--hrstad.pdf *

Institutt for medier og kommunikasjon (IMK) (2018). «Forskningsspørsmål». Oppdatert 25. januar 2018. URL: <https://www.hf.uio.no/imk/forskning/prosjekter/stromming-av-kulturindustriene-stream/forskningsspørsmål.html>

Ipsos (2020). «SoMe-tracker 20Q». Oppdatert 28.april 2020. URL: <https://www.ipsos.com/nb-no/ipsos-some-tracker-q120>

Jesnes, Kristin, Beate Sletvold Øistad, Kristin Alsos og Torstein Nesheim (2016). Aktører og arbeid i delingsøkonomien. Delrapport. Fafo-notat 2016: 23. URL: <https://fafo.no/images/pub/2016/10247.pdf>

Kafai, Yasmin B., Carrie Heeter, Jill Denner og Jennifer Y. Sun (2008). Beyond Barbie and Mortal Kombat. New Perspectives on Gender and Gaming. Cambridge, Massachusetts: The MIT Press.

Kalsnes, Bente og Lene Pettersen (2019). «Algoritmer, analysedata og automatisering i det digitale medielandskapet – en avledningsmanøver», Norsk medietidsskrift 26(1): 1-10. DOI: <https://doi.org/10.18261/ISSN.0805-9535-2019-01-01> *

Kane, Aina Aune (2020). «Navs vurdering av brukeres bistandsbehov», Tidsskrift for velferdsforskning 23(1):35-48. DOI: <https://doi.org/10.18261/issn.2464-3076-2020-01-03> *

Katzenbach, Christian og Lena Ulbricht (2019). «Algorithmic Governance», Internet Policy Review 8(4): 1-18. DOI: <https://doi.org/10.14763/2019.4.1424>

Kiberg, Håvard (2019). «Hvordan forstå og undersøke populærmusikkens «svarte bokser»?», Norsk medietidsskrift 26(1): 1 – 11. DOI: <https://doi.org/10.18261/ISSN.0805-9535-2019-01-05> *

Kommunal- og moderniseringsdepartementet (KMD) (2016). Meld. St. 27 (2015–2016). Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet. Oslo: Departementenes sikkerhets- og serviceorganisasjon. URL: <https://www.regjeringen.no/no/dokumenter/meld.-st.-2720152016/id2483795/>

Kommunal- og moderniseringsdepartementet (KMD) (2017). Ethiske retningslinjer for statstjenesten. Oslo: Kommunal- og moderniseringsdepartementet. URL: https://www.regjeringen.no/contentassets/6febadef60054700aadd61535e979198/no/pdfs/etiske_retningslinjer_rev_2017.pdf

Kommunal- og moderniseringsdepartementet (KMD) (2019). Én digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025. Oslo: Departementenes sikkerhets- og serviceorganisasjon. URL: https://www.regjeringen.no/contentassets/db9bf2bf10594ab88a470db40da0d10f/no/pdfs/digitaliseringsstrategi_for_offentlig_sektor_rettet.pdf

Kommunal- og moderniseringsdepartementet (KMD) (2020). Nasjonal strategi for kunstig intelligens. Oslo: Departementenes sikkerhets- og serviceorganisasjon. URL: <https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf>

Kuldova, Tereza (2019). How Outlaws Win Friends and Influence People. Cham: Palgrave Macmillan.

Kuldova, Tereza (2020). «Imposter Paranoia in the Age of Intelligent Surveillance», Journal of Extreme Anthropology 4(1): 45-73. DOI: <https://doi.org/10.5617/jea.7813> *

Kuldova, Tereza Østbø, Ida Drange, Heidi Enehaug, Eivind Falkum, Anders Underthun og Christin Thea Wathne (2020). Faglig skjønn under press: Fire case-studier og en sammenfatning. AFI Rapport 2020. Oslo: Arbeidsforskningsinstituttet / OsloMet. URL: https://www.academia.edu/44107174/Faglig_skj%C3%B8nn_under_press_Fire_case_studier_og_en_sammenfatning

Kullmann, Miriam (2018). «Platform work, algorithmic decision-making, and EU gender equality law», International Journal of Comparative Labour Law and Industrial Relations 34(1): 1-21. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3195728

Lagesen, Vivian Anette og Knut H. Sørensen (2009). «Walking the line? The enactment of the social/technical binary in software engineering», Engineering Studies 1(2): 129-149. DOI: <https://doi.org/10.1080/19378620902911584>

Leavy, Susan (2018). «Gender bias in artificial intelligence: the need for diversity and gender theory in machine learning», GE '18: Proceedings of the 1st International Workshop on Gender Equality in Software Engineering: 14-16. DOI: <https://doi.org/10.1145/3195570.3195580>

Lilleslåtten, Mari (2018). «Intelligente roboter kan forsterke kjønnsnormer», Kilden kjønnsforskning.no, 9. november 2018. URL: <http://kjonnsforskning.no/nb/2018/11/intelligente-roboter-kan-forsterke-kjonnsnormer>

Lutz, Christoph (2019). «Digital inequalities in the age of artificial intelligence and big data», Human Behaviour & Emerging Technologies 1(2):141-148. DOI: <https://doi.org/10.1002/hbe2.140>

Løvgren, Mette og Julia Orapabo (2018). «En sosiologi om symbolske grenser», Norsk sosiologisk tidsskrift 2(3): 203-207. DOI: <https://doi.org/10.18261/issn.2535-2512-2018-03-01>

Mark, Michael Spjelkavik, Cathrine Edelhards Tømte, Terje Næss og Trude Røsdal (2019). «Leaving the windows open – økt mangdel på IKT-sikkerhetskompetanse i Norge», Norsk sosiologisk tidsskrift 3(3): 173-190. DOI: <https://doi.org/10.18261/issn.2535-2512-2019-03-02> *

- Nedelkoska, Ljubica og Glenda Quintini (2018). Automation, skills use and training. OECD Social, Employment and Migration Working Papers, No. 202. Paris: OECD Publishing. DOI: <https://doi.org/10.1787/2e2f4eea-en>
- NITO (2019). Erklæring. URL: <https://www.nito.no/contentassets/72800175dcc745d98e328f35eb72f1f2/ansvarlig-bruk-av-kunstig-intelligens-i-arbeidslivet-erklaring.pdf>
- Nordiske Mediedager (2013). Medievaner blant journalister: Undersøkelse blant journalister 7. – 25. februar 2013. URL: <http://kyber.blob.core.windows.net/nmd/1689/medievaner-journalister-2013.pdf>
- Nordmo, Morten, Julie Øverbø Næss, Marte Folkestad Husøy og Mads Nordmo Arnestad (2020). «Friends, Lovers or Nothing: Men and Women Differ in Their Perceptions of Sex Robots and Platonic Love Robots», *Frontiers in Psychology* 11(355): 1-10. DOI: <https://doi.org/10.3389/fpsyg.2020.00355> *
- NOU 2019: 5. Ny forvaltningslov. Lov om saksbehandlingen i offentlig forvaltning (forvaltningsloven). Oslo: Departementenes sikkerhets- og serviceorganisasjon. URL: <https://www.regjeringen.no/contentassets/8d681c58c7a94929954ca9538a2476a6/no/pdfs/nou201920190005000dddpdfs.pdf>
- Nærings- og fiskeridepartementet (2017). Meld. St. 27 (2016 – 2017). Industrien – grønnere, smartere og mer nyskapende. Oslo: Nærings- og fiskeridepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-27-20162017/id2546209/sec1>
- OECD (2017). OECD Employment Outlook 2017. Paris: OECD Publishing. DOI: https://doi.org/10.1787/empl_outlook-2017-en
- OECD (2018). Bridging the Digital Gender Divide. Include, Upskill, Innovate. Paris: OECD Publishing. URL: <http://www.oecd.org/internet/bridging-the-digital-gender-divide.pdf>
- Pajarinen, Mika, Petri Rouvinen og Anders Ekeland, Anders (2015). «Computerization Threatens One-Third of Finnish and Norwegian Employment», ETLA Brief 34, The Research Institute of the Finnish Economy. URL: <https://www.etla.fi/wp-content/uploads/ETLA-Muistio-Brief-34.pdf> *
- Peña, Alejandro, Ignacio Serna, Aythami Morales, Julian Fierrez (2020). «Bias in Multimodal AI: Testbed for Fair Automatic Recruitment», IEEE CVPR Workshop on Fair, Data Efficient and Trusted Computer Vision, Washington, Seattle, USA, arXiv.org. URL: <https://arxiv.org/abs/2004.07173>
- Pettersen, Lene (2019). «Why Artificial Intelligence Will Not Outsmart Complex Knowledge Work», *Work, Employment and Society* 33(6): 1058 –1067. DOI: <https://doi.org/10.1177/0950017018817489> *
- Poulsen, Adam, Eduard Fosch-Villaronga og Roger Andre Søraa (2020). «Queering machines», *Nature Machine Intelligence* 2(152): 152-152. DOI: <https://doi.org/10.1038/s42256-020-0157-6> *
- Rangul, Ingeborg Vigerust (2019). «Ikke flere i arbeid etter finsk forsøk: – Borgerlønn betyr hele folket på kontantstøtte, sier Mette Nord», *Fagbladet*. 10. april 2019. URL: <https://fagbladet.no/nyheter/ikke-flere-i-arbeid-etter-finsk-forsok--borgerlo-nn-betyr-hele-folket-pa-kontantstotte-sier-mette-nord-6.91.611518.ac2371396c>
- Robinson, Laura, Shelia Cotten, Hiroshi Ono, Anabel Quan-Haase, Gustavo Mesch, Wenhong Chen, Timothy M. Hale og Michael J. Stern (2015). «Digital inequalities and why they matter», *Information, Communication & Society* 18(5): 569–582. DOI: <https://doi.org/10.1080/1369118X.2015.1012532>
- Russel, Stuart og Peter Norvig (2010). *Artificial Intelligence. A Modern Approach*. 3. utgave. London: Pearson.
- Røhnebæk, Maria (2016). «Fra bakkebyråkrati til skjermbyråkrati», *Tidsskrift for velferdsforskning* 19(04): 288–304. DOI: <https://doi.org/10.18261/issn.2464-3076-2016-04-01>
- Searle, John R. (1980). «Minds, brains, and programs», *Behavioral and Brain Sciences* 3(3): 417-457. URL: <https://www.law.upenn.edu/live/files/3413-searle-j-minds-brains-and-programs-1980pdf>
- Seeberg, Marie Louise, Berit Aasen, Justyna Bell, Christer Hyggen og Ida Tolgensbakk (2020). Strategisk mottaksplassering og treffsikker bosetting: tidlige grep for integrering av flyktninger i kommunene. NOVA Rapport 2/2020. Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring / OsloMet. URL: <https://www.imdi.no/contentassets/4e5e942e51a444c7af60ced8e-b10fb1c/strategisk-mottaksplassering-og-treffsikker-bosetting-tidlige-grep-for-integrering-av-flyktninger-i-kommunene.pdf> *
- Simshaw, Drew (2019). «Ethical Issues in Robo-Lawyeering: The Need for Guidance on Developing and Using Artificial Intelligence in the Practice of Law», *Hastings Law Journal* 70(1): 173-213. URL: https://repository.uchastings.edu/hastings_law_journal/vol70/iss1/4
- Smith, Aaron (2018). «Many Facebook users don't understand how the site's news feed works». Pew Research Centre, Fact tank: News in The Numbers, 5. september 2018. URL: <https://www.pewresearch.org/fact-tank/2018/09/05/many-facebook-users-dont-understand-how-the-sites-news-feed-works/>
- Statistisk sentralbyrå, SSB (2020a). «Bruk av IKT i husholdningene». Oppdatert 16. september 2020. URL: <https://www.ssb.no/ikthus>
- Statistisk sentralbyrå, SSB (2020b). Norsk mediebarometer 2019. Oslo/Kongsvinger: Statistisk sentralbyrå. URL: <https://www.ssb.no/kultur-og-fritid/artikler-og-publikasjoner/norsk-mediebarometer-2019>

- Søraa, Roger Andre (2017). «Mechanical genders: How do humans gender robots?», Gender, Technology and Development 21(1-2): 99-115. DOI: <https://doi.org/10.1080/09718524.2017.1385320> *
- Søraa, Roger Andre og Eduard Fosch-Villaronga (2020). «Exoskeletons for all: The interplay between exoskeletons, inclusion, gender, and intersectionality», Paladyn, Journal of Behavioral Robotics 11(1): 217-227. DOI: <https://doi.org/10.1515/pjbr-2020-0036> *
- Teknologirådet (2015). Forutseende politi. Kan dataanalyser hjelpe politiet til å være på rett sted til rett tid? Rapport 04 2015. Oslo: Teknologirådet. URL: https://teknologiradet.no/wp-content/uploads/sites/105/2018/05/ForebyggendeAnalyse_ending_WEB.pdf
- Teknologirådet (2018). Kunstig intelligens – muligheter, utfordringer og en plan for Norge. Oslo: Teknologirådet. URL: <https://teknologiradet.no/wp-content/uploads/sites/105/2018/09/Rapport-Kunstig-intelligens-og-maskinlaering-til-nett.pdf>
- Thun, Sylvi, Pål F. Kamsvåg, Birgit Kløve, Eva A. Seim og Hans Y. Torvatn (2019). «Industri 4.0: Whose revolution? The Digitalization of Manufacturing Work Processes», Nordic journal of working life studies 9(4). DOI: <https://doi.org/10.18291/njwls.v9i4.117777> *
- Tørresen, Jim (2013). Hva er kunstig intelligens. Oslo: Universitetsforlaget.
- Underthun, Anders og Arild H. Steen (2018). Digital omstilling i arbeidslivet. En rapport for fire bransjer. AFI rapport 2018: 05. Oslo: Arbeidsforskningsinstituttet / OsloMet. URL: <https://s32603.pcdn.co/wp-content/uploads/2018/10/Digital-omstilling-i-arbeidslivet.pdf>
- Vigdor, Neil (2019). «Apple Card Investigated After Gender Discrimination Complaints». The New York Times, 10. November 2019. URL: <https://www.nytimes.com/2019/11/10/business/apple-credit-card-investigation.html>
- Wajcman, Judy (2007). «From Women and Technology to Gendered Technoscience». Information, Community and Society 10(3): 287-298. DOI: <https://doi.org/10.1080/13691180701409770>
- Webster, Juliet (2016). «Microworkers of the Gig Economy: Separate and Precarious», New Labor Forum 25(3): 56-64. DOI: <https://doi.org/10.1177/1095796016661511>
- West, Mark, Rebecca Kraut og Han Ei Chew (2019). I'd blush if I could. Closing Gender Divides in Digital skills Through Education. EQUALS og UNESCO. GEN/2019/EQUALS/1 REV 3. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1>
- West, Sarah Myers, Meredith Whittaker og Kate Crawford (2019). Discriminating Systems: Gender, Race and Power in AI, AI Now Institute. URL: <https://ainowinstitute.org/discriminatingystems.pdf>
- Woodfield, Ruth (2000). Women, work and computing. Cambridge: Cambridge University Press.
- Woodford, Isabel (2020). «Rise of #MeToBots: scientists develop AI to detect harassment in emails». The Guardian, 3. januar 2020. URL: <https://www.theguardian.com/technology/2020/jan/03/metoobots-scientists-develop-ai-detect-harassment>
- Woods, Heather, Suzanne (2018). «Asking more of Siri and Alexa: Feminine persona in service of surveillance capitalism», Critical Studies in Media Communication 35(4): 334-349. DOI: <https://doi.org/10.1080/15295036.2018.1488082>
- World Economic Forum (2018). Global Gender Gap Report. Cologny/Geneva: World Economic Forum. URL: http://www3.weforum.org/docs/WEF_GGGR_2018.pdf
- Xenidis, Raphaële og Linda Senden (2020). «EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination». I Ulf Bernitz et al. (red.), General Principles of EU law and the EU Digital Order. Zuidpoolsingel: Kluwer Law International. URL: https://cadmus.eui.eu/bitstream/handle/1814/65845/Pre-print%20version%20Chapter%20Xenidis_Senden.pdf?sequence=2
- Øye, Dana Darja (2019). «Robotene er allerede her. En empirisk vurdering av automatisering og endringer i yrkessammensetningen i det norske arbeidsmarkedet», Søkelys på arbeidslivet 1(2): 21-35. DOI: <https://doi.org/10.18261/issn.1504-7989-2019-01-02-02> *

Vedlegg 1

Søkeord

Søkeord det har blitt utført søk med i utprøving av ulike søkestrenger.

Norsk	Engelsk
<p>KI</p> <p>“kunstig intelligens” OR KI</p> <p>algoritm*</p> <p>robot*</p> <p>automat*</p> <p>«maskin* AND læring» OR maskinlæring</p>	<p>AI</p> <p>«artificial intelligence» OR AI</p> <p>algorithm*</p> <p>robot*</p> <p>automat*</p> <p>“machine learning”</p> <p>chatbot*</p> <p>bot*</p> <p>«black box*»</p>

<p>kjønn</p> <p>kjønn</p> <p>kvinne*</p> <p>jente*</p> <p>diskrimin*</p> <p>trakass*</p> <p>mobbe* OR mobbing</p> <p>mangfold</p> <p>likestil*</p> <p>demokrati*</p>	<p>gender</p> <p>gender OR sex</p> <p>women OR woman OR female*</p> <p>girl*</p> <p>discrimin*</p> <p>harrass*</p> <p>bully* OR bullies</p> <p>equal*</p> <p>divers*</p> <p>democra*</p>
<p>tema 1:</p> <p>tjeneste*</p> <p>velferd*</p> <p>offentlig</p> <p>helse</p> <p>samfunn</p>	<p>tema 1:</p> <p>service*</p> <p>welfare*</p> <p>public*</p> <p>health</p> <p>society</p>

<p>tema 2:</p> <p>arbeid*</p> <p>jobb*</p> <p>yrke*</p>	<p>tema 2:</p> <p>work*</p> <p>labour*</p> <p>labor*</p> <p>job*</p>
<p>tema 3:</p> <p>«sosiale medier»</p> <p>«sosiale plattform*»</p> <p>cyber*</p> <p>online*</p> <p>internett</p> <p>digital*</p>	<p>tema 3:</p> <p>«social media»</p> <p>«social platform*»</p> <p>cyber*</p> <p>online*</p> <p>internet</p> <p>digital*</p>

Vedlegg 2

Søkelogg

Dette er søkeloggen med de søkestrengene som gav treff på relevant litteratur, og som samtidig viser hvor bredt vi gikk ut for å saumfare feltet.

Norart	17.06.20	«kunstig intelligens»	Enkle ordsøk for å treffe bredt og fange opp mest mulig av relevans.	113		1
		algoritmer		63		4
	18.06.20	digital* AND kjønn		6		1
	18.06.20	(algoritm* OR robot* OR automat* OR cyber OR online) AND kjønn		5		2
Idunn	17.06.20	kjønn OG "kunstig intelligens" OG Norge	6 relevante, 3 av disse dubletter fra treff i Norart	13		3
	23.06.20	algoritmer OG kjønn OG Norge		13	3 (dublett)	0
Web of Science	28.06.20	("artificial intelligence" OR algorithm* OR robot* OR automat* OR "machine learning") AND (gender OR women OR girl*)	Avgrenset til norsk affiliasjon og artikler utgitt etter 2010.	190		3
	28.06.20	algorithm* AND ("health care" OR work* OR labour* OR welfare*) AND Norway		23	1	1

Scopus	28.06.20	("artificial intelligence" OR algorithm* OR robot* OR automat* OR "machine learning") AND (gender OR women OR girl*)	Avgrenset til norsk affiliasjon og artikler utgitt etter 2010.	199	3 (1 dublett)	2
	29.06.20	algorithm* AND ("health care" OR work* OR labour* OR welfare*) AND Norway		76	1 (bare) nevner kjønn	1
Sociological abstracts	28.06.20	("artificial intelligence" OR algorithm* OR robot* OR automat* OR "machine learning") AND (gender OR women OR girl*)	Avgrenset til norsk affiliasjon og artikler utgitt etter 2010.	26		1
	29.06.20	algorithm* AND ("health care" OR work* OR labour* OR welfare*) AND Norway		10	0	



Postboks 564, 1327 Lysaker
post@kilden.forskningsradet.no
www.kjonnsforskning.no
@kjonnsforskning