

CompLex



Senter for rettsinformatikk / Avdeling for forvaltningsinformatikk

Julie Lossius Husum

Rettslige begrensninger for forvaltningens bruk av data til trening av maskinlæringsalgoritmer

Spesielt om personvernrettslige og forvaltningsrettslige rammer

4/2021



UiO • Det juridiske fakultet

Henvelseler om denne bok kan gjøres til:
Senter for rettsinformatikk
Postboks 6706 St. Olavs plass
0130 Oslo
<http://www.jus.uio.no/ifp/om/organisasjon/seri/>

ISSN 2703-8777 Complex (online)

Grafisk produksjon: 07 Media AS – 07.no

Forord

Denne rapporten baserer seg på min masteroppgave i rettsvitenskap som ble levert våren 2021. Det er gjort noen få korreksjoner i fotnoter og litteraturlisten, men ellers er oppgaven uendret.

Bakgrunnen for temaet er min interesse for kunstig intelligens og de rettslige rammene som har betydning for kunstig intelligens. Interessen stammer fra at jeg i løpet av studiet forstod at det er mange uavklarte rettslige spørsmål som knytter seg til utvikling og bruk av kunstig intelligens. Det er spennende at mange av disse spørsmålene ikke har entydige svar. Jeg valgte å skrive om rettslige rammer for forvaltningens bruk av data til utvikling av kunstig intelligens fordi dette er en aktuell og viktig problemstilling. Forvaltningen sitter på store mengder personopplysninger og det er mange potensielle anvendelsesområder knyttet til disse. Det har derfor stor praktisk betydning å avklare rettslige rammer når opplysninger skal brukes som treningsmateriale for maskinlæringsalgoritmer.

Jeg vil takke Dag Wiese Schartum for tålmodig veiledning under hele prosessen. Jeg har satt stor pris på hans tilgjengelighet og hans gode råd når jeg har stått fast. Videre vil jeg takke pappa for gjennomlesing og nyttige innspill ved ferdigstillingen av oppgaven.

Oslo, 04. juli 2021.

Julie Lossius Husum

Innhold

Forord	3
Sammendrag	6
1 Innledning	7
1.1 Tema og aktualitet	7
1.2 Problemstilling	9
1.3 Rettskilder og metode	11
1.4 Fremstillingen videre	14
2 Nærmere om maskinlæring og betydningen av data	16
2.1 Om maskinlæring og prediksjonsmodeller	16
2.2 Betydningen av data	17
2.2.1 Den faktiske betydningen av data	17
2.2.2 Den rettslige betydningen av data	17
2.3 Særlig om profilering	19
3 Rettslige begrensninger for bruk av data	21
3.1 Innledning	21
3.2 Prinsippet om formålsbegrensning, jf. art. 5(1)(b)	23
3.2.1 Innledning	23
3.2.2 Lovlig behandling under primærformålene, jf. art. 6(1)(e), jf. art. 5(1)(b)	24
3.2.3 Forbud mot viderebehandling av data til «uforenelige» formål, jf. art. 5(1)(b), jf. art. 6(4)	30
3.2.4 Unntak fra viderebehandlingsforbud for «statistiske formål»,jf. art. 5(1)(b), jf. art. 89 nr. 1	36
3.2.5 Oppsummering	37
3.3 Forbud mot å dele taushetsbelagte opplysninger med andre, jf. fvl. §§ 13 flg.	38
3.3.1 Innledning	38
3.3.2 Gjelder taushetsplikten ved bruk av opplysninger til trening av maskinlæring?	39
3.3.3 Unntak fra taushetsplikten	41
3.3.4 Oppsummering	43

3.4	Krav om kvalitet og minimering, jf. PVF art. 5(1)(c) og (d) og det forvaltningsrettslige prinsippet om forsvarlig saksbehandling.	43
3.4.1	Innledning.	43
3.4.2	Rettslige utgangspunkter.	44
3.4.3	Forsvarlighet i ML-baserte beslutningsstøttesystemer.	47
3.4.4	I hvilken utstrekning vil krav om kvalitet og minimering medføre begrensninger for bruk av data?	50
3.4.5	Oppsummering.	53
3.5	Samlet drøftelse.	53
4	Rettspolitiske vurderinger.	56
	Litteraturliste.	58

Sammendrag

Denne rapporten undersøker i hvilken utstrekning personvernrettslige og forvaltningsrettslige reguleringer begrenser forvaltningens mulighet til å anvende data ved utvikling av maskinlæringsalgoritmer som skal benyttes i beslutningsstøtteverktøy. Rapporten behandler kun rettslige rammer for *utviklingen* av maskinlæringsalgoritmer, det vil si bruk av data som treningsmateriale for en maskinlæringsalgoritme. De rettslige grunnlagene som behandles er formålsbegrensningsprinsippet i personvernforordningen (PVF) art. 5(1)(b), reglene om taushetsplikt i forvaltningsloven §§ 13 flg., de personvernrettslige dataminimerings- og riktighetsprinsippene i henholdsvis PVF art. 5(1)(c) og (d) og det forvaltningsrettslige prinsippet om forsvarlig saksbehandling. Rapporten avdekker at alle grunnlagene vil kunne begrense forvaltningens mulighet til å benytte data til å utvikle løsninger basert på maskinlæring. Da regelverkene ikke er utarbeidet med tanke på utvikling av kunstig intelligens er det en mulighet for at disse begrensningene ikke er tilsiktete. I tillegg oppstår utfordringer som en følge av at systemutvikling i liten grad er regulert i forvaltningsretten. Dette medfører at ulovfestede forvaltningsrettslige prinsipper får stor betydning når de rettslige rammene skal klarlegges. Det er derfor behov for klarere rettslige rammer for utvikling av maskinlæringsalgoritmer på forvaltningsrettens område. Dette er nødvendig for å imøtekomme de utfordringene som følger med forvaltningens bruk av personopplysninger til utvikling av beslutningsstøtteverktøy basert på maskinlæring.

1 Innledning

1.1 Tema og aktualitet

Som del av den pågående digitale transformasjonen har den norske stat uttrykt klart at utvikling og bruk av kunstig intelligens i offentlig sektor er en viktig nasjonal interesse i årene som kommer.¹ Det finnes ingen fast definisjon av kunstig intelligens, men en mulig definisjon er at kunstig intelligens er teknologier som på ulike måter utfører oppgaver som ellers ville kreve menneskelig intelligens.² Av de ulike teknologiene som hører under kunstig intelligens, fremheves ofte maskinlæring (ML) for å være den som har størst potensial til å utføre avanserte oppgaver.³ Maskinlæring kjennetegnes ved at systemet er i stand til å lære av data på egenhånd, uten menneskelig innblanding.⁴ Maskinlæring skjer ved at en maskinlæringsalgoritme⁵ analyserer et datamateriale og bygger en modell med utgangspunkt i mønstre som den finner i datasettet.⁶ En type modell som kan utvikles ved anvendelse av maskinlæringsteknikker er prediksjonsmodeller.⁷ Prediksjonsmodeller kan forutsi hva som er et sannsynlig utfall med utgangspunkt i statistiske korrelasjoner maskinlæringsalgoritmen har funnet i dataene.

Det finnes mange anvendelsesområder for prediksjonsmodeller. Et mulig anvendelsesområde er som komponenter i beslutningsstøttesystemer som anvendes i forvaltningens saksbehandling.⁸ Beslutningsstøttesystemer er dataprogrammer som på ulike måter kan forenkle og gi bidrag ved manuell utførelse av oppgaver knyttet til saksbehandling.⁹ Beslutningsstøttesystemer gir kun støtte ved saksbehandlingen, og er ikke det samme som *rettslige beslutningssystemer* der hele eller deler av enkeltsaksbehandlingen automatiseres.¹⁰ Ved anvendelse av prediksjonsmodeller i beslutningsstøttesystemer vil man kunne utvikle saksbehandlingsverktøy som kan gjøre prediksjoner, eksempelvis om en

1 Kommunal og moderniseringsdepartementet (2020) s. 2.

2 Bendiksen og Norman Hansen (2019) s. 11-12.

3 Datatilsynet (2018) s. 5.

4 Datatilsynet (2018) s. 5.

5 En algoritme er en kombinasjon av instruksjoner som forteller en datamaskin hvordan den skal utføre en oppgave, se Bendiksen og Norman Hansen (2019) s. 18.

6 STOA (2020) s. 11.

7 Valevatn (2017).

8 Datatilsynet (2021).

9 Schartum, Jansen og Tranvik (2017) s. 56.

10 Schartum (2018) s. 25.

parts behov, atferd eller preferanser. Dette er informasjon som kan være nyttig for saksbehandleren i vurderingen av hva som er en god og hensiktsmessig fremgangsmåte i en sak. Prediksjoner om individer kalles *profilerings*.¹¹ Det finnes flere eksempler på at forvaltningen har tatt i bruk prediksjonsmodeller basert på maskinlæring i beslutningsstøttesystemer. Eksempelvis har Skatteetaten siden 2016 jobbet med å utvikle prediktive modeller som kan kontrollere skattemeldinger og indikere om de bør kontrolleres for feil eller juks.¹² Også NAV har pågående prosjekter der det utvikles prediksjonsmodeller som kan brukes i beslutningsstøtte, blant annet en modell som kan predikere sannsynlighet for at en sykemeldt vil ha nytte av tidlig dialogmøte med arbeidsgiver og lege.¹³

Som tidligere beskrevet utvikles prediksjonsmodeller ved at en maskinlæringsalgoritme lærer fra et datamateriale og danner en modell med utgangspunkt i statistiske korrelasjoner den finner i dette. Det er som regel behov for store mengder data for at maskinlæringsalgoritmen skal være i stand til å finne mønstre som kan benyttes som utgangspunkt for prediksjoner.¹⁴ Utvikling av prediksjonsmodeller ved anvendelse av maskinlæring er derfor helt avhengig av god tilgang til data. Med den gradvise utbyggingen av velferdsstaten fra 50-tallet til i dag, har Norge en omfattende offentlig forvaltning som gir tjenester til alle deler av samfunnet.¹⁵ Som en følge av dette sitter forvaltningen på store mengder opplysninger som potensielt kan danne grunnlag for prediksjonsmodeller. Forvaltningen har derfor gode forutsetninger for å kunne utvikle og anvende løsninger basert på maskinlæring til ulike formål.

For at den digitale transformasjonen mot økt bruk av maskinlæring skal være mulig, må forvaltningen ha lovlig tilgang til det datamaterialet de besitter. Dette aktualiserer spørsmålet om hvilke rettslige grenser forvaltningen må ta hensyn til når de ønsker å anvende data til å utvikle prediksjonsmodeller som kan anvendes i beslutningsstøttesystemer.

11 Profilerings er særskilt regulert i personvernforordningen, jf. art. 4(4). Se nærmere om profilerings nedenfor i avsnitt 2.2.1.

12 Løland, Berset og Haff (2017) s. 345.

13 Datatilsynet (2021).

14 STOA (2020) s. 16.

15 Christensen og Berg (2019).

1.2 Problemstilling

I denne rapporten diskuterer jeg i hvilken grad og på hvilken måte personvernrettslige og forvaltningsrettslige bestemmelser medfører rettslige skranker for forvaltningens bruk av data til utvikling av prediksjonsmodeller. Som nevnt ovenfor, vil man ved selve utviklingen av prediksjonsmodellen trenge et datamateriale som kan danne utgangspunkt for de korrelasjonene den ferdige modellen vil inneholde. I forbindelse med sammenstillingen av et slik treningsmateriale, oppstår spørsmålet om hvilke rettslige grenser forvaltningen må overholde for at dataene lovlig skal kunne anvendes til utviklingen. Personvernforordningen (forkortet PVF) er en sentral rettslig ramme, da det er stor sannsynlighet for at mye av datamaterialet forvaltningen vil anvende vil inneholde personopplysninger. Personvernforordningen gjelder ved helt eller delvis automatisert behandling av personopplysninger¹⁶ og har vært gjennomført i norsk rett siden 2018.¹⁷ Forvaltningsloven (forkortet fvl.) regulerer virksomheten til forvaltningsorganer, jf. fvl. § 1. Utover taushetspliktreglene, inneholder ikke forvaltningsloven bestemmelser som får direkte anvendelse på forvaltningens bruk av data til utvikling av prediksjonsmodeller. Forvaltningsloven kan imidlertid suppleres av ulovfestede forvaltningsrettslige prinsipper ved vurderingen av forvaltningens adgang til å benytte opplysninger ved trening av maskinlæring.¹⁸

Jeg har valgt å ta utgangspunkt i de personvern- og forvaltningsrettslige reglene som har størst betydning for forvaltningens tilgang til data. Dette gjelder reguleringer som har direkte betydning for lovlig adgang til opplysninger, samt reguleringer som stiller krav til opplysningens kvalitet og mengde. En konsekvens av dette er at jeg ikke vurderer betydningen av personvernforordningens øvrige krav for lovlig behandling av personopplysninger og heller ikke andre mer perifere forvaltningsrettslige regler.¹⁹

Formålsbegrensningsprinsippet i PVF art. 5(1)(b) er relevant for lovlig tilgang til opplysninger. Det følger av bestemmelsen at personopplysninger ikke kan viderebehandles på en måte som er «uforenlig» med de opprinnelige formålene opplysningene ble samlet inn for. Formålet med bestemmelsen er å sørge for en effektiv beskyttelse av den registrertes rettigheter ved at behandlingsansvarlige ikke kan anvende opplysninger til annet enn det de er samlet inn for.²⁰ Data forvaltningen besitter, er som regel samlet inn i forbindelse med utredning ved

16 PVF art. 2(1).

17 Personopplysningsloven § 1.

18 Schartum (2020) s. 172.

19 Eksempelvis beskyttelsesinstruksen og folkeregisterloven kap. 10.

20 Datatilsynet (2018) s. 16.

enkeltsaksbehandling knyttet til forvaltningens ulike ansvarsområder. Spørsmålet som oppstår, er i hvilken utstrekning bruk av data til utvikling av prediksjonsmodeller vil være «uforenelig» med disse innsamlingsformålene.

En annen mulig begrensning finnes i de alminnelige taushetspliktreglene i fvl. §§ 13 flg. Forvaltningsloven § 13 fastslår at enhver som utfører arbeid for forvaltningen plikter å hindre at «andre» får adgang eller kjennskap til det han i forbindelse med arbeidet får vite om «noens personlige forhold». Dersom forvaltningen vil anvende opplysninger som gjelder «noens personlige forhold» ved utvikling av en prediksjonsmodell blir spørsmålet om dette vil medføre en krenkelse av kravet om å hindre at «andre» får adgang eller kjennskap til informasjonen.

En tredje begrensning er knyttet til rettslige krav om kvalitet og minimering. Det følger av dataminimeringsprinsippet i PVF art. 5(1)(c) at behandlingen av opplysninger skal være «adekvate, relevante og begrenset til det som er nødvendig for formålene de behandles for.» Prinsippet stiller krav til at det skal behandles så lite opplysninger som mulig, men samtidig at det skal behandles et tilstrekkelig antall opplysninger for å oppnå formålet med behandlingen. Prinsippet kan medføre en utfordring ved maskinlæring fordi det er vanskelig å si på forhånd hvor mye og hvilke typer opplysninger man vil trenge for å utvikle en prediksjonsmodell.²¹ Dette må ses i sammenheng med at man ofte trenger store mengder opplysninger ved anvendelse av maskinlæring.²²

Der hvor dataminimeringsprinsippet krever at bruk av personopplysninger skal begrenses så mye som mulig, kan riktighetsprinsippet i PVF art. 5(1)(d) medføre begrensninger i form av krav om at opplysningene må være «korrekte og om nødvendig oppdaterte». I offentlig forvaltning må riktighetsprinsippet suppleres med kravet om at forvaltningens saksbehandling skal være forsvarlig. Prinsippet kan supplere forvaltningsloven ved at forvaltningens saksbehandling kan anses uforsvarlig, selv om ikke forvaltningslovens bestemmelser er brutt.²³ Spesielt den delen av forsvarlighetskravet som gjelder saksutredningen er sentral for rapportens tema. Kravet om utredning er lovfestet i fvl. §§ 17 og 37, men er ansett for å gjelde som et generelt prinsipp.²⁴ Det følger av utredningsprinsippet at forvaltningen har plikt til å sørge for at grunnlaget for avgjørelser har tilstrekkelig kvalitet. Når riktighetsprinsippet og utredningsprinsippet ses i sammenheng, oppstår spørsmålet om hvilke kvalitetskrav som må stilles til datamaterialet som skal benyttes ved trening av maskinlæringsalgoritmer. I tillegg oppstår

21 Datatilsynet (2018) s. 17.

22 STOA (2020) s. 16.

23 Woxholt (2011) s. 45-46

24 Woxholt (2011) s. 354.

spørsmålet om hvordan disse kvalitetskravene skal avveies mot plikten til å begrense behandlingen av personopplysninger som følger av dataminimeringsprinsippet.

Som fremstillingen viser, vil bruk av data til utvikling av prediksjonsmodeller være underlagt både personvernrettslige og forvaltningsrettslige reguleringer. Et formål med rapporten er å få frem i hvilken utstrekning det oppstår et samspill mellom regelsettene når de rettslige rammene for bruk av data skal klarlegges.

I denne rapporten tar jeg derfor for meg følgende problemstillinger:

- I hvilken utstrekning og på hvilke måter vil overnevnte forvaltningsrettslige og personvernrettslige bestemmelser begrense forvaltningens adgang til bruk av data til trening av maskinlæringsalgoritmer som skal brukes i beslutningsstøtteverktøy?
- På hvilke måter skjer det et samspill mellom personvernretten og forvaltningsretten når disse begrensningene skal klarlegges?

1.3 Rettskilder og metode

For å vurdere rettslige begrensninger for forvaltningens bruk av data til utvikling av prediksjonsmodeller gjør jeg en rettslig analyse av de nevnte personvernrettslige og forvaltningsrettslige reguleringene. For spørsmålet om på hvilke måter det skjer et samspill mellom personvernretten og forvaltningsretten velger jeg å belyse dette ved å trekke linjer mellom rettsreglene der dette kan få frem relevante sammenhenger.

På personvernrettens område er personvernforordningen den primære rettslige reguleringen. Som angitt ovenfor, tar rapportens problemstillinger utgangspunkt i relevante personvernprinsipper i PVF art. 5(1): formålsbegrensningsprinsippet i bokstav b), dataminimeringsprinsippet i bokstav c) og riktighetsprinsippet i bokstav d). Også andre personvernrettslige bestemmelser vil bli trukket inn i den grad de kan belyse rapportens tema. Personvernprinsippene er formulert som prinsipper og ikke som absolutte rettsregler. I tillegg er vilkårene skjønnsmessig og lite presist formulert. Dette medfører at reglene har mindre gjennomslagskraft enn alminnelige positivrettslige regler, og således ikke fungerer som absolutte skranker.²⁵ I hvilken utstrekning prinsippene må gis gjen-

25 Schartum (2020) s. 87.

nomslag beror på avveininger mellom de interesser prinsippene gir uttrykk for og eventuelle motstridende interesser.²⁶

Personvernforordningens korte virketid medfører metodiske utfordringer når betydningen av personvernprinsippene i den enkelte sak skal klarlegges. Jeg har ikke funnet praksis fra EU-domstolen som kan belyse min problemstilling direkte. Jeg har videre valgt å ikke benytte praksis fra det opphevede personverndirektivet (96/46/EF). Selv om mye av innholdet i personverndirektivet er videreført i personvernforordningen,²⁷ har personvernforordningen flere nye bestemmelser og gjennomgående en annen ordlyd. I tillegg kommer at personvernets status og prioritet er i endring i en stadig mer digitalisert verden.²⁸ Jeg anser derfor den tolkningsmessige sammenhengen for å være endret på en slik måte at praksis fra det tidligere personverndirektivet ikke kan tillegges vekt uten en nærmere begrunnelse. Jeg har heller ikke funnet praksis fra det opphevede personverndirektivet som gir direkte bidrag til behandling av rapportens problemstilling, og har derfor ikke anvendt rettspraksis fra EU-domstolen i rapporten.

Med manglende avklaringer i rettspraksis, må anvendelse av personvernforordningen bygge på andre rettskilder. I henhold til EU-rettslig metode er formål og kontekst sentrale tolkningsmomenter, og forordningens fortale er derfor relevant ved tolkningen.²⁹ Fortalen er allikevel ikke rettslig bindende, og har derfor i utgangspunktet begrenset vekt som rettskilde.³⁰ Videre anvender jeg rådgivende uttalelser fra Artikkel 29-gruppen (heretter A29-gruppen), der disse kan gi tolkningsbidrag. A29-gruppen var EUs rådgivende organ under det tidligere personverndirektivet, men har siden blitt avløst av Det europeiske personvernrådet. Det europeiske personvernrådet har gitt sin tilslutning til flere av A29-gruppens tolkningsuttalelser,³¹ og disse anses for å ha forholdsvis stor rettskildevekt ved tolkning av forordningen.³² Jeg benytter meg også av juridisk litteratur knyttet til personvernforordningen i forbindelse med arbeidet med rapporten. I henhold til EU-rettslig metode skal EU-retten praktiseres autonomt og ensartet i EU/EØS-landene,³³ og jeg anvender derfor både norsk og interna-

26 Schartum (2020) s. 87.

27 Schartum (2020) s. 25.

28 Datatilsynet (2018) s. 4.

29 Sejersted (2011) s. 57.

30 Schartum (2020) s. 23.

31 Blant annet retningslinjer om automatiske avgjørelser og profilering (2018), se The European Data Protection Board (2018) s. 1-2.

32 Schartum (2020) s. 24.

33 Fredriksen og Mathisen (2019) s. 403.

sjonal litteratur. Dette gir en mulighet til å identifisere enigheter og uenigheter i forståelsen av forordningens bestemmelser.

En metodisk utfordring knyttet til det forvaltningsrettslige rettskildebildet er at det i stor grad mangler reguleringer knyttet til systemutvikling.³⁴ Med unntak av forvaltningslovens regler om taushetsplikt, kommer derfor ikke forvaltningsloven direkte til anvendelse på rapportens problemstillinger. Heller ikke *NOU 2019:5 Ny forvaltningslov* foreslår å regulere forvaltningens systemutvikling, og kan således ikke gi veiledning.³⁵ Flere av forvaltningslovens regler anses allikevel for å gi uttrykk for generelle forvaltningsrettslige prinsipper.³⁶ Det er noe uenighet i litteraturen om hvor stor rolle de ulovfestede forvaltningsrettslige prinsippene spiller ved siden av forvaltningsloven,³⁷ men det anses ikke tvilsomt at forvaltningsloven suppleres av forvaltningsrettslige prinsipper.³⁸ Det er disse prinsippene som vil danne utgangspunktet ved vurderingen av rettslige rammer for forvaltningens bruk av data til trening av maskinlæring. De forvaltningsrettslige prinsippene er ingen fast størrelse, men heller en gruppe av rettsregler som anses som så grunnleggende i forvaltningsretten at de betegnes prinsipper.³⁹ Prinsippet om at forvaltningens saksbehandling skal være forsvarlig er relevant for rapportens problemstillinger. Forsvarlighetsprinsippet medfører et krav om at forvaltningens avgjørelser skal treffes på grunnlag av en forsvarlig behandling.⁴⁰ Forsvarlighetskravet ligger til grunn for forvaltningslovens saksbehandlingsregler, og disse kan således gi veiledning når forsvarlighet skal vurderes på ulovfestet grunnlag. Særlig prinsippet om at forvaltningens utredning skal være forsvarlig har stor betydning for vurderingen av rettslige rammer for utvikling av prediksjonsmodeller som skal anvendes i beslutningsstøttesystemer.⁴¹

Forsvarlighetsprinsippet blir i det følgende brukt både som selvstendig rettskilde og som et tolkningsmoment i rapporten. Med det første mener jeg at jeg vurderer i hvilken utstrekning kravet om at forvaltningens saksbehandling skal være forsvarlig vil medføre begrensninger for forvaltningens bruk av data til utvikling av prediksjonsmodeller. At forsvarlighetsprinsippet blir brukt som et tolkningsmoment, medfører at jeg anvender prinsippet ved tolkningen av rettslige krav som følger av andre bestemmelser.

34 Schartum (2018) s. 171.

35 NOU 2019:5 s. 159.

36 Schartum (2018) s. 172.

37 Smith (2018) s. 204 og Boe (2018) s. 137.

38 Se blant annet Woxholt (2011) s. 45-46, Boe (2018) s. 136, Graver (2019) s. 75 flg.

39 Innjord (2014) s. 575.

40 Graver (2019) s. 101.

41 Schartum (2018) s. 45.

Anvendelse av forvaltningsrettslige prinsipper både som rettskilde og som tolkningsmoment medfører metodiske utfordringer. Innholdet i kravet til forsvarlig saksbehandling formuleres ulikt av ulike forfattere,⁴² og det er derfor vanskelig å legge til grunn en fast definisjon. Videre beror prinsippene på skjønnsmessige vurderinger, og innebærer derfor ingen absolutte krav.⁴³ Uklarhet rundt innholdet i forsvarlighetskravet er bakgrunnen for at Forvaltningslovutvalget avviste å lovfeste et krav til forsvarlighet i utredningen om ny forvaltningslov.⁴⁴ Som beskrevet ovenfor, medfører også personvernprinsippene skjønnsmessige vurderinger, men i motsetning til prinsippet om forsvarlig saksbehandling, er personvernprinsippene kodifiserte. Personvernprinsippene har derfor en tydeligere kjerne enn prinsippet om forsvarlighet og er således mer tilgjengelige.

1.4 Fremstillingen videre

For å vurdere rettslige begrensninger for bruk av data til utvikling av maskinlæringsmodeller (ML-modeller), er det nødvendig å forstå hvilken rolle data spiller ved utviklingen. Dette forutsetter kunnskap om sammenhengen mellom maskinlæring og data ved utvikling av prediksjonsmodeller, samt hvilken betydning dataene vil ha for det endelige beslutningsstøtteverktøyet. Videre er det nødvendig å forstå hvilke utfordringer som er knyttet til bruk av prediksjonsmodeller i beslutningsstøtte.

Jeg gir derfor i kapittel 2 et overblikk over maskinlæringsteknologien og hvordan den benyttes ved utvikling av ML-modeller. Deretter forklarer jeg de ulike betydningene av data i rapporten. Dette gjelder for det første den faktiske sammenhengen mellom dataene som anvendes og prediksjonsmodellens egenskaper. I tillegg forklarer jeg den rettslige sammenhengen mellom data og reguleringene som er tema for rapporten. Avslutningsvis i kapittel 2 skisserer jeg noen utfordringer som er knyttet til bruk av prediksjonsmodeller i beslutningsstøtte. Dette gjør jeg ved å forklare implikasjonene av *profilering*.

I rapportens hoveddel gjør jeg en rettslig analyse av reguleringene som er tema for rapporten. Jeg behandler først de rettslige reguleringene som regulerer *lovlig tilgang* til opplysninger, altså formålsbegrensningsprinsippet i art. 5(1)(b) og taushetspliktreglene i fvl. §§ 13 flg. Deretter analyserer jeg de rettslige begrensningene som kan stille krav til opplysningenes mengde og kvalitet, dvs. data-

42 Se blant annet Schartum (2018) s. 45, Boe (2018) s.121-132, Graver (2019) s. 102.

43 Schartum (2018) s. 45.

44 NOU 2019:5 s. 158.

minimerings- og personvernprinsippet og det forvaltningsrettslige forsvarlighetsprinsippet.

Med utgangspunkt i funn jeg har gjort i den rettslige analysen, forsøker jeg i avsnitt 3.5 å belyse samspillet mellom personvernretten og forvaltningsretten når begrensninger knyttet til bruk av opplysninger skal klarlegges. Avslutningsvis knytter jeg i kapittel 4 noen bemerkninger til forholdet mellom det eksisterende rettslige rammeverket og kunstig intelligens *de lege ferenda*.

2 Nærmere om maskinlæring og betydningen av data

2.1 Om maskinlæring og prediksjonsmodeller

Innledningsvis har jeg beskrevet maskinlæring som en samling datateknikker der en algoritme lærer fra et datamateriale, uten at denne læringen spesifikt programmeres av et menneske. Ved utvikling av prediksjonsmodeller er særlig maskinlæringsteknikker for *veiledet læring* en aktuell metode.⁴⁵ Ved veiledet læring anvendes et kategorisert datasett til å trene opp modellen.⁴⁶ At datasettet er kategorisert betyr at de ulike typene av data i treningssettet vil være koblet til en «merkelapp».⁴⁷ Læringen skjer ved at algoritmen anvender statistiske metoder for å analysere sammenhenger mellom dataene og de ulike merkelappene.⁴⁸

En utfordring ved maskinlæring er at det kan være vanskelig å vite på forhånd hvilke opplysninger som er nødvendige for å utvikle en bestemt modell.⁴⁹ I tillegg vil større mengder treningsdata i utviklingen ofte kunne forbedre presisjonen i algoritmen.⁵⁰ Utvikling av en prediksjonsmodell gjøres derfor gradvis frem til modellen har lært de sammenhengene som er ønsket. Når modellen er blitt tilstrekkelig utviklet vil den være i stand til å koble ukjente data av samme type til de eksisterende merkelappene, og det er dette som gjør at modellen kan brukes til å gjøre prediksjoner.⁵¹

Et eksempel kan være at man vil utvikle en modell som skal predikere sannsynlig varighet av en sykemelding. Man kan da tenke seg at forvaltningen kan anvende opplysninger fra et utvalg av tidligere saker, eksempelvis medisinsk diagnose, kjønn og alder (altså «data») på den ene siden, og hvilken varighet sykemeldingen hadde på den andre siden (altså «merkelapp»).⁵² Ved å anvende statistiske metoder vil algoritmen kunne generere en modell som reflekterer sammenhengene mellom de ulike opplysningene og varighet av en sykemelding. Prediksjonsmodellen representerer innsiktene fra læringen,⁵³ men vil ikke selv

45 Henriksen (2019).

46 Datatilsynet (2018) s. 7.

47 Bendiksen og Normann Hansen (2019) s. 19.

48 STOA (2020) s. 11.

49 Datatilsynet (2018) s. 17.

50 Teknologirådet (2018) s. 19.

51 STOA (2020) s. 15

52 Eksempelet er ment å illustrere bruken av data ved veiledet læring, og er ikke ment å gi et faktisk eksempel på hvordan det kan utvikles prediksjonsmodeller til bruk i beslutningsstøtte.

53 Datatilsynet (2018) s. 9.

inneholde de opplysningene som er brukt til å lære opp systemet.⁵⁴ Man vil derfor ikke kunne avsløre opplysninger om personer ved å studere modellen.

2.2 Betydningen av data

2.2.1 Den faktiske betydningen av data

Gjennomgangen over viser betydningen av data når man skal utvikle prediksjonsmodeller gjennom bruk av maskinlæring. Den ferdige prediksjonsmodellen er en representasjon av det datamaterialet som er brukt til å trene den.⁵⁵ Dette er bakgrunnen for at et vanlig uttrykk innenfor maskinlæring er «søppel inn, søppel ut».⁵⁶ Datamaterialet har betydning for modellens prediksjonsevner og for om hvorvidt den vil ha uønskede eller til og med skadelige egenskaper. Prediksjonsmodellen vil blant annet påvirkes av datamaterialets størrelse, utvalg av typer data og hvordan disse er «merket», samt hvordan ulike grupper er representert i datamaterialet.⁵⁷

Sammenhengen mellom datamaterialet og den endelige prediksjonsmodellens egenskaper, samt den rettslige betydningen av dette vil bli nærmere gjennomgått nedenfor i avsnitt 3.4.

2.2.2 Den rettslige betydningen av data

Den rettslige betydningen av data referer til data som «rettsfakta».⁵⁸ I elektronisk databehandling er «data» betegnelsen på «symboler organisert i bestemte mønstre som representerer fakta, observasjoner og/eller ideer.»⁵⁹ Dataene må oppfylle bestemte krav for at de rettslige reguleringene som er tema for rapportens skal komme til anvendelse. Når data representerer personopplysninger utgjør dette en kobling til personvernforordningen, jf. PVF art. 2(1). Etter artikkel 4(1) er «personopplysninger» forstått som «enhver opplysning om en identifisert eller identifiserbar fysisk person». «Identifiserbar» viser til at det er *mulig* å identifisere en person direkte eller indirekte på bakgrunn av opplysningene som behandles. Hvorvidt det er mulig å identifisere en bestemt fysisk person beror på en konkret vurdering i det enkelte tilfelle.⁶⁰

54 Datatilsynet (2018) s. 9.

55 Datatilsynet (2018) s. 5.

56 Datatilsynet (2018) s. 10.

57 Se nærmere i avsnitt 4.4.4.

58 Aarli og Mæhle (2018) s. 235.

59 Schartum, Jansen og Tranvik (2017) s. 188

60 Skullerud (2019) s. 152.

Innholdet i dataene henger også sammen med anvendelsen av PVF art. 9(1). I den utstrekning forvaltningen vil benytte opplysninger som faller inn under særlige kategorier personopplysninger, for eksempel helseopplysninger, må denne behandlingen oppfylle forordningens krav. Dette reiser blant annet krav om særskilt behandlingsgrunnlag etter art. 9(2), samt at behandlingsopplegget i alminnelighet vil anses for å medføre økt risiko for den registrerte.

Også de forvaltningsrettslige reglene om taushetsplikt henger sammen med egenskaper ved dataene som brukes. Det følger av fvl. § 13 at taushetsplikten medfører plikt til å hindre at andre får adgang eller kjennskap til det man i forbindelse med arbeidet får vite om «noens personlige forhold». «Noens personlige forhold» er forstått som «opplysninger det er vanlig å ville holde for seg selv».⁶¹ Hvorvidt opplysninger forvaltningen vil bruke til trening av maskinlæring skal anses for gjelde «noens personlige forhold» beror på en skjønsmessig vurdering som må avgjøres konkret.⁶²

Når de rettslige reguleringene er knyttet til egenskaper ved dataene vil de kunne komme til anvendelse samtidig. Helseopplysninger vil eksempelvis både omfattes PVF art. 9(1) og vil regnes som «personlig forhold» etter fvl. § 13.⁶³ Når taushetspliktreglene beror på en skjønsmessig vurdering, medfører det at taushetspliktreglene kan ha et videre anvendelsesområde enn art. 9. Dette er tilfellet dersom opplysninger som ikke omfattes av kategoriene i art. 9(1) anses for å omhandle «personlige forhold». Taushetspliktige opplysninger vil allikevel være underlagt personvernforordningen da de vil være opplysninger om en «identifisert person», jf. definisjonen i art. 4(1).

Forvaltningsorganer besitter ofte data samlet inn i forbindelse med utredning knyttet til enkeltsaksbehandling. Disse dataene vil som regel være personopplysninger, hvor en del i tillegg vil være underlagt taushetsplikt. Ikke alle maskinlæringsalgoritmer vil måtte utvikles ved å bruke personopplysninger eller taushetspliktige opplysninger. For det første kan det anvendes syntetiske testdata⁶⁴ eller data som ikke inneholder informasjon om personer overhodet. Videre kan opplysningene anonymiseres slik at de ikke lenger kan kobles til en identifisert person. I så fall vil verken personvernforordningen eller de forvaltningsrettslige taushetspliktreglene gjelde.⁶⁵ Anonymisering forutsetter at det ikke er mulig å

61 Woxholt (2011) s. 265.

62 Graver (2019) s. 369.

63 Ot.prp.nr.3 (1976–1977) s. 14.

64 Syntetiske testdata er fiktive personopplysninger som er laget for testformål, se Ringnes (2019).

65 Jf. Fortalepunkt 26 og fvl. § 13a nr. 2.

identifisere et individ direkte eller indirekte gjennom opplysningene.⁶⁶ Dette kan være en utfordring ved anvendelse av maskinlæring. Store mengder opplysninger medfører en fare for at opplysninger kan kombineres slik at individer kan re-identifiseres.⁶⁷ Spørsmålet om hvilke rettslige begrensninger som gjelder for bruk av slike opplysninger er derfor aktuelt. I tillegg er ikke alle rettsregler som gjelder forvaltningens bruk av data tilknyttet dataene. Eksempelvis vil kravet om at forvaltningens saksbehandling skal være forsvarlig gjelde generelt.

2.3 Særlig om profilering

Bruk av personopplysninger for å evaluere eller gjøre antagelser om en person kalles *profilering*. Profilering er definert i PVF art. 4(4) som enhver automatisert behandling av personopplysninger der personlige aspekter vurderes «særlig for å (..) forutsi aspekter som gjelder nevnte fysiske person(..)». Dette medfører at bruk av en prediksjonsmodell i beslutningsstøtte for å gjøre prediksjoner om en person vil falle innenfor personvernforordningens definisjon av profilering.

Profilering regnes for å være inngripende, og kan potensielt krenke en persons rettigheter og friheter. Dette gjelder særlig når profilen benyttes som utgangspunkt for en beslutning om en person. Dersom prediksjonsmodellen gjør en uriktig prediksjon, vil det kunne lede til uriktig beslutningsgrunnlag. Det er også en fare for at profilen kan bidra til diskriminering eller sosial eksklusjon dersom prediksjonsmodellen ikke er rettferdig.⁶⁸ I tillegg medfører profilering at det genereres nye personopplysninger om personen som utsettes for profileringen. Prognosen vil utgjøre en ny opplysning om personen som vanskelig kan verifiseres eller avkreftes, og kan derfor beskrives som en «subjektiv» form for personopplysning.⁶⁹

Faren for negative konsekvenser for den som utsettes for profilering er bakgrunnen for at profilering er særskilt regulert i personvernforordningen.⁷⁰ Reguleringen av profilering knytter seg allikevel bare til situasjonen der profilen benyttes til å gjøre prediksjoner om den registrerte,⁷¹ altså situasjonen *etter* at prediksjonsmodellen er tatt i bruk. Bruk av personopplysninger til å utvikle predik-

66 Skullerud (2019) s. 152.

67 STOA (2020) s. 36.

68 A29-gruppen (2018) s. 5-6.

69 Motzfeldt og Abkenar (Red.) (2019) s. 152.

70 Profilering utløser blant annet informasjonsplikter etter PVF art. 13(2)(f) og 14(2)(g), rett til begrenset behandling etter art. 21(1), samt at vilkår etter PVF art. 22 må være oppfylt.

71 Dette følger av definisjonen av profilering i art. 4(4) som gjelder bruken av personopplysninger til å vurdere personlige aspekter knyttet til «nevnte fysiske person».

sjonsmodellen som skal anvendes ved profileringen faller således utenfor de særskilte reguleringene. Selv om utviklingen av profileringsverktøy ikke er direkte regulert av forordningen, vil det forhold at dataene skal benyttes til å utvikle prediksjonsmodeller ha betydning for den rettslige analysen av forordningens bestemmelser.

I den kommende rettslige analysen vil risikoen knyttet til profilering være ett moment som må vurderes. Vurderingen vil i tillegg måtte se hen til at data spiller en avgjørende rolle ved utvikling av prediksjonsmodeller, samt at det er en nær sammenheng mellom dataene som anvendes ved treningen og det endelige beslutningsstøtteverktøyet sine egenskaper.

3 Rettslige begrensninger for bruk av data

3.1 Innledning

Behandling av personopplysninger skal ha rettslig grunnlag i art. 6 og 9. En av de viktigste årsakene til at forvaltningen behandler personopplysninger er for å oppfylle utredningskravet i fvl. § 17 i forbindelse med enkeltsaksbehandling. Videre vil forvaltningen ofte behandle opplysninger for oppfølging og kontroll, som avledede formål fra enkeltsaksbehandlingen. Det sentrale grunnlaget for at forvaltningen kan behandle alminnelige personopplysninger for disse formålene finnes i art. 6(1)(e). Bestemmelsen gjelder behandlinger som er nødvendige «for å utføre en oppgave i allmennhetens interesse» eller «utøve offentlig myndighet som den behandlingsansvarlige er pålagt». Myndighetsutøvelse kan forstås som bebyrdende eller positive avgjørelser og omfatter forvaltningens enkeltsaksbehandling.⁷² Behandling av opplysninger for oppfølging og kontroll, vil trolig enten anses for å være nødvendig for myndighetsutøvelse eller for utførelse av oppgaver i allmennhetens interesse. Det er derfor grunn til å anta at art. 6(1)(e) vil være dekkende for behandling av personopplysninger for de overnevnte formålene innenfor det enkelte forvaltningsorgans ansvarsområde. I den utstrekning forvaltningen vil behandle særlige kategorier personopplysninger, kreves i tillegg grunnlag i et av behandlingsoalternativene i art. 9(2). Særlig alternativene i bokstav b)⁷³ og g)⁷⁴ vil være aktuelle grunnlag for forvaltningens enkeltsaksbehandling. Av hensyn til omfanget av fremstillingen drøfter jeg ut fra at forvaltningen behandler opplysninger med grunnlag i art. 6(1)(e). Jeg trekker kun inn behandling av særlige kategorier personopplysninger etter art. 9(1) i den utstrekning det kan illustrere sentrale momenter knyttet til adgangen til å benytte opplysninger til å utvikle prediksjonsmodeller.

Behandling av opplysninger med utgangspunkt i art. 6(1)(e) skiller seg ut ved at det etter art. 6(3) kreves at slik behandling må ha grunnlag i «nasjonal rett». I høringsnotatet om gjennomføring av forordningen er det skrevet at nasjonal rett kan forstås som lov- og forskriftsbestemmelser, men at trolig også vedtak fattet i medhold av lov eller forskrift er omfattet.⁷⁵ Etter norsk rettstradisjon er

72 Schartum (2020) s. 130-131 og Skullerud (2019) s. 182.

73 Bokstav b) gjelder behandlinger som er nødvendige for å (...) oppfylle forpliktelser eller rettigheter på området arbeidsrett, trykderett og sosialrett.

74 Bokstav g) gjelder behandlinger som er nødvendige av hensyn til viktige allmenne interesser.

75 Prop.56 LS (2017–2018) s. 33.

forvaltningens myndighetsutøvelse regulert i lov eller forskrift.⁷⁶ Trolig vil også forvaltningens utføring av «oppgaver i allmennhetens interesse» ha grunnlag i lov eller forskrift, i hvert fall hvis oppgaven forutsetter behandling av større mengder personopplysninger. Situasjonen etter dette er at behandlinger med grunnlag i art. 6(1)(e) skal ha grunnlag i lov eller forskrift. Dette kan trolig ikke utledes direkte av art. 6(3), men følger av norsk rettstradisjon.⁷⁷ Utgangspunktet står særlig sterkt på legalitetsprinsippets område, slik at jo mer inngripende en behandling av personopplysninger kan være, jo sterkere står kravet om et klart hjemmelsgrunnlag.⁷⁸

Tema i det følgende er på hvilke måter og i hvilken utstrekning forvaltningsrettslige og personvernrettslige bestemmelser begrenser forvaltningens bruk av opplysninger til trening av maskinlæringsalgoritmer som skal benyttes i beslutningsstøtte. Ved denne vurderingen har behandlingsformålet stor betydning. Behandlingsformålet vil for det første ha en nær sammenheng med behandlingsgrunnlaget som hjemler behandlingen.⁷⁹ Jeg diskuterer derfor i rapportens første del premissene for at utvikling av prediksjonsmodeller skal ha grunnlag i art. 6(1)(e). For det tilfelle at behandlingen ikke kan hjemles direkte i art. 6(1)(e), blir spørsmålet om opplysningene kan viderebehandles i medhold av art. 5(1)(b), jf. art. 6(4). Også her vil behandlingsformålet være avgjørende for vurderingen. Spørsmålet om behandlingsgrunnlag og viderebehandling blir behandlet i avsnitt 3.2.

Forutsatt at opplysningene kan behandles lovlig i medhold av art. 6(1)(e), alternativt lovlig viderebehandles etter art. 5(1)(b), blir spørsmålet i hvilken utstrekning taushetspliktreglene i forvaltningsloven vil kunne utgjøre en skranke. Dette er tema i avsnitt 3.3.

I den utstrekning verken krav om behandlingsgrunnlag eller taushetsplikt stenger for bruk av opplysningene, må det tas stilling til hvilken mengde personopplysninger som kan anvendes, samt hvilke kvaliteter disse opplysningene må ha. Dette aktualiserer personvernprinsippene, da primært dataminimerings- og riktighetsprinsippet. Også her vil behandlingsformålet ha stor betydning ved vurderingen av hvilke opplysninger som er «nødvendige» for at formålet skal nås. Behandlingsformålet har også betydning ved vurderingen av hvilke krav som følger av prinsippet om forsvarlig saksbehandling. I avsnitt 3.4 drøfter jeg hvilke krav som stilles til datamaterialet etter dataminimerings- og riktighets-

76 Smith (2018) s. 37.

77 Skullerud (2019) s. 187.

78 Skullerud (2019) s. 187.

79 Schartum (2020) s. 126.

prinsippene tolket i lys av det forvaltningsrettslige kravet til forsvarlig saksbehandling, herunder utredningsprinsippet.

Avslutningsvis gjør jeg en samlet vurdering av i hvilken utstrekning forvaltningsretten og personvernretten medfører begrensninger for bruk av opplysninger til utvikling av maskinlæring. Fremstillingen fokuserer på hvordan det skjer et samspill mellom de to regelsettene.

3.2 Prinsippet om formålsbegrensning, jf. art. 5(1)(b)

3.2.1 Innledning

Formålsbegrensingsprinsippet i PVF art. 5(1)(b) stiller krav knyttet til formålene personopplysningene behandles for. Prinsippet består av tre hoveddeler: For det første skal opplysninger samles inn for «spesifikke, uttrykkelig angitte og berettigede formål». Ordlyden viser at det ikke er noen begrensning på antall behandlingsformål den behandlingsansvarlige kan fastsette, og det kan også fastsettes flere formål for den enkelte behandling.⁸⁰ For det andre skal opplysningene «ikke viderebehandles på en måte som er uforenelig med disse formålene». Artikkel 6(4) angir en ikke-uttømmende liste over momenter som kan vektlegges ved vurderingen av om et nytt formål er forenelig med det opprinnelige. Bestemmelsen angir videre opplysninger lovlig kan behandles for bestemte sekundærformål, forutsatt at slik behandling av personopplysninger oppfyller garantier som kreves etter art. 89 nr. 1. De lovlige sekundærformålene er «arkivformål i allmennhetens interesse», «formål knyttet til vitenskapelig eller historisk forskning» eller «statistiske formål».

Prinsippet om formålsbegrensning spiller en avgjørende rolle for personopplysningsvernet, og sammenlignet med de øvrige personvernprinsippene har formålsbegrensingsprinsippet en tydelig kjerne.⁸¹ Bestemmelsen er både formulert som et prinsipp og et krav for lovlig behandling.⁸² Den behandlingsansvarlige er pålagt å konkretisere hvilke formål opplysningene skal behandles for før innsamlingen, og det er forbudt å bruke opplysningene til andre *uforenelige* formål etter at de er samlet inn. Sammen med kravet om behandlingsgrunnlag, er formålsbegrensingsprinsippet derfor en viktig garanti for effektiv beskyttelse av den registrertes personvern og en forutsetning for den registrertes autonomi og kontroll over egne personopplysninger.⁸³ Vilkårene etter formålsbegrensingsprinsippet beror allikevel på skjønsmessige vurderinger.

80 Schartum (2020) s. 100.

81 Jf. ovenfor i avsnitt 1.3.

82 Schartum (2020) s. 91.

83 Datatilsynet (2018) s. 16.

Personvernforordningen inneholder ingen eksplisitt formulering om at det er lovlig å behandle personopplysninger for et formål de er samlet inn for, men dette utgangspunktet følger forutsetningsvis av forordningen. Derav kan opplysningene lovlig behandles for det formål de er samlet inn for i samsvar med behandlingsgrunnlaget som hjemler behandlingen. Da blir spørsmålet hva premissene er for at maskinlæring skal anses for å inngå under primærformålene, og derfor lovlig kunne behandles i medhold av art. 6(1)(e). Avgjørende for denne vurderingen er kravene om at formål skal være «spesifikke, uttrykkelig angitte og berettigede». Dette er tema i avsnitt 3.2.2. For det tilfelle at utvikling av prediksjonsmodeller ikke kan hjemles direkte i art. 6(1)(e), blir spørsmålet om et slikt formål er en forenelig viderebehandling. Vurderingen tar utgangspunkt i kriteriene angitt i art. 6(4). Avslutningsvis vurderer jeg om utvikling av prediksjonsmodeller kan anses for å inngå under et av de lovlige sekundærformålene, hvor alternativet «statistiske formål» er det sentrale grunnlaget.

Artikkel 6(4) viser at viderebehandling av opplysninger for nye formål kan gjøres dersom den registrerte samtykker eller dersom medlemslandet har vedtatt grunnlag for viderebehandling for bestemte formål i nasjonal lovgivning i medhold av kriteriene i art. 23. Ingen av grunnlagene er praktiske for forvaltningens bruk av data til trening av maskinlæringsalgoritmer, og blir ikke videre drøftet i rapporten.

3.2.2 Lovlig behandling under primærformålene, jf. art. 6(1)(e), jf. art. 5(1)(b)

3.2.2.1 Kravet om «spesifikke, uttrykkelig angitte og berettigede» formål i lys av art. 6(1)(e), jf. 6(3)

I det følgende klargjør jeg premissene for at maskinlæring skal inngå under primærformålene, og derfor lovlig kunne behandles i medhold av behandlingsgrunnlaget i art. 6(1)(e). Vurderingen beror på kravene om at personopplysninger «skal samles inn for spesifikke, uttrykkelig angitte og berettigede formål» i art. 5(1)(b). I utgangspunktet stiller ikke personvernforordningen krav til hvordan behandlingsformål rent faktisk skal angis⁸⁴ utover krav om at den registrerte skal gis tilstrekkelig informasjon etter PVF art. 13(1)(c) eller 14(1)(c) og at den behandlingsansvarlige kan dokumentere at behandlingsformålene er angitt.⁸⁵ Som fremhevet i avsnitt 3.1 ovenfor, skal behandling av opplysninger med grunnlag i art. 6(1)(e) i tillegg ha grunnlag i «nasjonal rett» etter art. 6(3). Jeg minner om at det i henhold til norsk rettstradisjon er sannsynlig at oppgaver som utføres i allmennhetens interesser og utøvelse av offentlig myndighet vil ha grunnlag i lovgivning eller forskrift, og at dette kravet er absolutt på legalitetsprinsippets område. Etter art. 6(3) er det *formålet* med behandlingen som «skal

84 Öman (2019) s. 115.

85 Jf. dokumentasjonskravene i art. 5(2) og 24(1).

være fastsatt i nevnte rettslige grunnlag». Fortalepunkt 45 modifierer tilsynelatende regelen til at formålene «bør» fastsettes i medlemsstatens nasjonale rett, men dette kan trolig ikke vektlegges da fortalen etter EU-rettslig metode ikke kan tilsidesette en klar ordlyd i forordningen.⁸⁶ Dersom kravet om at formålene skal angis i nasjonal rett ikke er absolutt, må det i så fall fastslås av EU-domstolen.

Utgangspunktet etter dette er at behandlingsformål med grunnlag i art. 6(1)(e) skal være fastsatt i nasjonal rett, da fortrinnsvis i lov eller forskrift. Kravene om at opplysninger skal samles inn for «spesifikke, uttrykkelig angitte og berettigede formål» etter art. 5(1)(b) bør derfor vurderes ut fra den kontekst at behandlingsformålene følger av særlovgivningen som regulerer det enkelte forvaltningsorgans ansvarsområde.

Kravet om at formålene må være «spesifikke» betyr at det ikke kan behandles personopplysninger for udefinerte eller ubegrensede formål.⁸⁷ Formålsangivelsen danner utgangspunktet for kravene etter de øvrige personvernprinsippene, herunder dataminimerings- og riktighetsprinsippet. Formålene må altså være angitt så nøyaktig at det blir mulig å vurdere hvilke opplysninger som er nødvendige for å nå formålet.⁸⁸ Eksempelvis vil personopplysningsbehandling for «gjennomføring av forvaltningens lovpålagte oppgaver» ikke oppfylle spesifikasjonskravet, mens «enkeltsaksbehandling i saker om dagpenger ved arbeidsledighet» vil være tilstrekkelig. Videre er det grunn til å lese kravet om at formål skal være spesifikke i sammenheng med legalitetsprinsippet. Dette medfører at kravet til presisjon er høyere dersom formålet hjemler personopplysningsbehandling som kan være inngripende overfor enkeltpersoner.⁸⁹

Artikkel 5(1)(b) fastslår at formål skal være «uttrykkelig angitt», mens art. 6(3) uttrykker at «formålet med behandlingen» skal være fastsatt i det nasjonale rettsgrunnlaget. Artikkel 6(3) medfører derfor ikke noen utvidelse av den regelen som allerede følger av art. 5(1)(b) utover et krav om at formålet må fastsettes i det nasjonale rettsgrunnlaget. «Uttrykkelig» betyr at formålene ikke kan være underforstått, men faktisk må uttrykkes klart.⁹⁰ Kravet vil være oppfylt dersom behandlingsformålet fremgår av loven. Det gjelder ikke et krav om at det uttrykkelig må fastsettes at personopplysninger kan behandles for formålet,⁹¹ og det er

86 Schartum (2020) s. 23.

87 de Terwangne (2020) s. 315.

88 Öman (2019) s. 116.

89 Se Skullerud (2019) s. 187 om lovkravet etter art. 6(3), men samme synspunkt gjør seg gjeldende for formuleringen av formål i lov.

90 Skullerud (2019) s. 173.

91 Skullerud (2019) s. 187.

heller ikke nødvendig å detaljere hvordan formålet skal oppnås.⁹² Derfor vil eksempelvis «enkeltsaksbehandling i saker om dagpenger» være tilstrekkelig, uten at det må uttrykkes eksplisitt at personopplysninger kan behandles for dette formålet. De enkelte behandlinger som er nødvendige for enkeltsaksbehandlingen må heller ikke angis.

For det tredje kreves det at behandlingsformålene må være «berettigede». Vilkåret skiller seg fra de øvrige ved at det stilles et materielt krav til formålets innhold, fremfor dets kvalitative utforming. Det er ulike syn på hvordan kravet om «berettiget» («legitimite» i den engelske språkversjonen) skal forstås. «Berettiget» kan forstås som lovlighet, og således kreve at formålene er i tråd med forordningen, men også annet regelverk, eksempelvis menneskerettighetene.⁹³ Det er lite trolig at dette er den fulle betydningen, sett i sammenheng med kravet om at personopplysninger skal behandles på en «lovlig» måte etter art. 5(1)(a). Kravet kan også referere til at behandlingsformålet må være en naturlig del av den behandlingsansvarliges virksomhet.⁹⁴ I så fall kan forvaltningen bare behandle opplysninger til formål som ligger innenfor forvaltningens alminnelige virksomhet. Trolig må «berettiget» anses for å innebære en kombinasjon av disse kravene. Det må formodes at behandlinger med grunnlag i art. 6(1)(e) vil være lovlige og ligge innenfor forvaltningens alminnelige virksomhet.

Om art. 5(1)(b) ses i sammenheng med art. 6(3) kan det konkluderes med at forvaltningens behandlingsformål etter art. 6(1)(e) skal angis i lovgivningen på en så presis måte at det blir mulig å vurdere hvilke personopplysninger som lovlige kan behandles, samt hvilke behandlingsaktiviteter som naturlig henger sammen med formålet. I den utstrekning formålet nødvendiggjør behandling av personopplysninger som kan være inngripende for enkeltpersoner, vil kravene til presisjon være strengere.

I det følgende vurderer jeg betingelsene for at utvikling av prediksjonsmodeller skal anses å ha grunnlag i art. 6(1)(e). Vurderingen av rettslig grunnlag for bruk av opplysninger til trening av maskinlæringsalgoritmer må selvfølgelig vurderes konkret med utgangspunkt i den enkelte særlov. Formålet med fremstillingen er således å påpeke generelle premisser for at behandlingsgrunnlaget skal gi dekning for en slik behandling i lys av kravene i art. 5(1)(b).

92 Jarbekk (red.) (2019) s. 162.

93 Skullerud (2019) s. 173.

94 Schartum (2020) s. 101, Jarbekk (2019) s. 143 og Skullerud (2019) s. 173-174.

3.2.2.2 Utvikling av prediksjonsmodeller med grunnlag i art. 6(1)(e)

Vurderingen om når bruk av opplysninger til utvikling av prediksjonsmodeller vil ha grunnlag i art. 6(1)(e) må gjøres med utgangspunkt i kravet om at formål skal være «spesifikke, uttrykkelig angitte og berettigede». I lys av konklusjonen ovenfor, vil kravene være oppfylt dersom særloven eksplisitt uttrykker at opplysninger kan behandles for å utvikle systemer basert på maskinlæring som skal benyttes til enkeltsaksbehandling, oppfølging eller kontroll. I et slikt tilfelle vil det være mulig å forutse hva opplysningene vil bli benyttet til, samt hvilke behandlinger som vil inngå i behandlingsopplegget. Trolig vil det også være tilstrekkelig at særloven uttrykker at opplysninger kan behandles for forvaltningens systemutvikling, allikevel slik at formuleringen i mindre grad sikrer forutberegnelighet for den registrerte. Dersom bruken av opplysninger til utvikling av prediksjonsmodeller kan være inngripende, for eksempel fordi det skal utvikles beslutningsstøtteverktøy til kontrollformål, vil formuleringen trolig ikke være tilstrekkelig.⁹⁵

Kravene i art. 5(1)(b) tilsier at det ikke vil være tilstrekkelig at formålsangivelsen er begrenset til enkeltsaksbehandling, oppfølging og kontroll innenfor det aktuelle forvaltningsorgans ansvarsområde. Dersom eksempelvis loven hjemler bruk av opplysninger for enkeltsaksbehandling i saker om dagpenger, vil trolig ikke samme opplysninger kunne benyttes til å utvikle en prediksjonsmodell som skal benyttes ved saksbehandling knyttet til dagpengesøknader. I et slikt tilfelle vil det overordnede formålet med behandlingen være det samme, nemlig gjennomføring av enkeltsaksbehandling i saker om dagpenger. Jeg finner allikevel at utvikling av prediksjonsmodeller må anses for å ligge så langt unna de behandlinger som vanligvis gjøres ved enkeltsaksbehandling, at det må anses for å være et annet formål. I et slikt tilfelle vil kravet om forutberegnelighet ikke være oppfylt. Videre ville en slik løsning medføre at kravene i art. 5(1)(b) om at formål skal være «spesifikke, uttrykkelig angitte og berettigede» miste betydning. For at den registrerte skal kunne ivareta egne rettigheter etter personvernforordningen må personen forstå hvilke formål personopplysningene vil bli brukt til.

Konklusjonen etter dette er at kravene om at formål skal være spesifikke, uttrykkelige og berettigede medfører at formål om behandling av opplysninger til utvikling av maskinlæring må formuleres klart i særloven for å sikres behandlingsgrunnlag etter art. 6(1)(e). Det vil trolig aksepteres at opplysningene samles inn for mer generelle formål, for eksempel «systemutviklingsformål», men dette vil antakelig forutsette at behandlingen ikke er inngripende overfor enkeltpersoner.

95 Skullerud (2019) s. 187.

Behandling av opplysninger for utvikling av prediksjonsmodeller med utgangspunkt i art. 6(1)(e) kan ikke skje dersom det aktuelle forvaltingsorganet ikke kan vise til et tilstrekkelig klart behandlingsformål i lovgivningen. Personvernforordningens system er at det ikke er mulig å erstatte et behandlingsgrunnlag med et alternativt behandlingsgrunnlag etter at opplysningene er samlet inn. Behandlingsgrunnlag skal angis *før* innsamlingen, og kan ikke byttes ut dersom den behandlingsansvarlige vil behandle opplysningene for formål som ikke har grunnlag i det opprinnelige behandlingsgrunnlaget.⁹⁶ Dersom forvaltningen vil anvende opplysningene ved trening av maskinlæringsalgoritmer, må dette formålet anses for å være forenelig med de opprinnelige innsamlingsformålene i medhold av art. 5(1)(b) jf. art. 6(4).

Før jeg vurderer forenelighet av prediksjonsmodeller, vil jeg drøfte hvorvidt det bør stilles krav om at nye formål skal være «spesifikke, uttrykkelig angitt og berettigede», slik det tilsvarende gjøres når opplysningene samles inn.

3.2.2.3 Analogisk anvendelse av krav om at formål skal være «spesifikke, uttrykkelig angitte og berettigede» for nye behandlingsformål

Artikkel 5(1)(b) angir opplysninger skal *samles inn* for «spesifikke, uttrykkelig angitte og berettigede» formål. Spørsmålet er om ordlyden skal tas på ordet, eller om det må gjelde et tilsvarende krav når opplysningene behandles for nye formål *etter at de er samlet inn*.

Jeg vurderer at det er gode grunner for å anvende kravet om at formål skal være «spesifikke, uttrykkelig angitte og berettigede» analogisk på nye behandlingsformål. Det er lite naturlig at det skal stilles lempeligere krav til sekundærformål enn til primærformål. Alternativt kan et slikt krav anses å være utledet av kravet om at nye behandlingsformål må være forenelige med primærformålet. I så fall inkluderer forenelighetskriteriet et krav om at nye formål må være «spesifikke, uttrykkelig angitte og berettigede». At nye formål er spesifikke og uttrykkelige vil i tillegg være en forutsetning for at vurderingen av forenelighet skal kunne gjøres. Det har mindre betydning hvorvidt et slikt krav bygger på en analogisk anvendelse av første del av art. 5(1)(b) eller om det ligger forutsetningsvis i forenelighetsvurderingen. Avgjørende er at det har gode grunner for seg at også nye behandlingsformål må være berettigede, formuleres tilstrekkelig presist og angis uttrykkelig.

Et spørsmål blir da hvordan formål om utvikling av prediksjonsmodeller som skal anvendes i beslutningsstøtteverktøy må ses i lys av disse kravene. I vurderingen av om formålet er «berettiget» er det både et spørsmål om lovlighet og om

⁹⁶ Skullerud (2019) s. 185 og Jarbekk (2019) s. 149-150.

det er innenfor forvaltningens naturlige virksomhet å utvikle slike verktøy. Det er lovlig å utvikle profileringsverktøy som skal benyttes som beslutningsstøtte. Det følger også forutsetningsvis av at personvernforordningen regulerer bruk av data til «profiling». Hvorvidt den konkrete prediksjonsmodellen er lovlig, vil bero på dens utforming og hvilket formål den benyttes til. Et beslutningsstøtteverktøy som predikerer risiko for trygdesvindler med utgangspunkt i etnisitet, vil eksempelvis trolig medføre ulovlig diskriminering i henhold til likestillings- og diskrimineringsloven.⁹⁷ Spørsmålet om det er innenfor forvaltningens naturlige virksomhet å utvikle prediksjonsmodeller som skal benyttes i beslutningsstøtte vil også bero på modellens nærmere utforming og formål. Det må anses naturlig at forvaltningen benytter ny teknologi som muliggjør bedre oppfyllelse av forvaltningens lovpålagte oppgaver. Motsatt vil prediksjonsmodeller for kommersielle formål ikke være berettiget, med mindre utvikling av slike modeller er saklig begrunnet i forvaltningsorganets virksomhet.⁹⁸

Ved lovlig viderebehandling vil det opprinnelige behandlingsgrunnlaget være tilstrekkelig behandlingsgrunnlag.⁹⁹ En konsekvens av dette er at kravet om at formål skal angis uttrykkelig ikke medfører at det nye behandlingsformålet må framgå av lovgivningen. Det følger allikevel av art. 13(3) og 14(4) at den behandlingsansvarlige skal gi den registrerte «informasjon om nevnte andre formål» før opplysninger kan behandles for nye formål. Videre må formålene uttrykkelig angis for å oppfylle dokumentasjonskravet i art. 5(2) og art. 24(1). Det vil altså uansett gjelde et krav om at forvaltningen må formulere og formidle formålet om å utvikle prediksjonsmodeller som skal benyttes i beslutningsstøtte.

Kravet om at behandlingsformål må formuleres og formidles aktualiserer vilkåret om at formål skal være «spesifikke». Spesifikasjonskravet er trolig det mest utfordrende kravet når det kommer til anvendelse av maskinlæring. Som angitt i avsnitt 2.2 kan det være utfordrende å definere formål ved utvikling av maskinlæring. Ofte vet man ikke hvordan maskinlæringsalgoritmen vil utvikle seg med utgangspunkt i de treningsdataene den gis,¹⁰⁰ og modellen kan fortsette å utvikle seg etter hvert som den trenes.¹⁰¹ Et eksempel kan være at man har et formål om å utvikle en modell som skal predikere sykefraværslengde, men ender opp med en modell som indikerer sannsynlighet for uførhet. Ved maskinlæring må i så fall spesifikasjonskravet medføre et minstekrav om at forvaltningen så langt som mulig søker å definere formålet med beslutningsstøtteverktøyet på en presis og nøyaktig måte.

97 Jf. Likestillings- og diskrimineringsloven § 6, jf. §§ 7-11.

98 Schartum (2018) s. 101

99 Se avsnitt 3.2.3.1.

100 Datatilsynet (2018) s. 17.

101 Teknologirådet (2018) s. 19.

3.2.3 Forbud mot viderebehandling av data til «uforenelige» formål, jf. art. 5(1)(b), jf. art. 6(4)

3.2.3.1 Innledning

Artikkel 5(1)(b) viser at opplysninger kan behandles for nye formål etter at de er samlet inn, forutsatt at de nye formålene er forenelige med de opprinnelige. Fortalepunkt 50 angir at «i et slikt tilfelle kreves det ikke et annet rettslig grunnlag enn det som ligger til grunn for innsamlingen av personopplysninger». Rekkevidden av uttalelsen er omdiskutert,¹⁰² men har gjennomgående blitt lagt til grunn som gjeldende rett, blant annet av Justis- og beredskapsdepartementet i høringsnotatet om gjennomføring av forordningen.¹⁰³ I lys av diskusjonen om forutberegnelighet kan det stilles spørsmål ved en løsning der primærformålet vil omfatte alle behandlinger som ikke er for uforenelige formål. Ordningen må bygge på en antagelse om at forutberegneligheten ivaretas gjennom kravet om at den nye behandlingen må være forenelig.¹⁰⁴ Siden ordlyden i fortalepunkt 50 er klar og forståelsen har blitt lagt til grunn i Justis- og beredskapsdepartementets høringsnotat og i juridisk litteratur, vil jeg drøfte ut fra at det opprinnelige behandlingsgrunnlaget vil være tilstrekkelig for viderebehandlingen. Hvorvidt krav om selvstendig behandlingsgrunnlag etter art. 6 eller 9 også gjelder for viderebehandlinger, må avklares av EU-domstolen.

Artikkel 5(1)(b) angir at opplysninger ikke kan viderebehandles på en måte som er «uforenelig» med det opprinnelige formålet med innsamlingen. Kravet er negativt formulert, men i fortalepunkt 50 og art. 6(4) anvendes den positive formuleringen «forenelig». Den positive formuleringen anvendes også av A29-gruppen i deres retningslinjer om formålsbegrensning (riktignok fra 2013),¹⁰⁵ og samme forståelse er fremmet i norsk juridisk litteratur.¹⁰⁶ Motsatt er det fremhevet i dansk og svensk juridisk litteratur at vurderingen etter art. 5(1)(b) beror på en negativ vurdering av om formålene er «uforenelige» med hverandre.¹⁰⁷ Formuleringen som legges til grunn vil ha betydning for hvilken posisjon man inntar ved vurderingen: enten en vurdering der man vektlegger kriterier som taler for at formålene er forenelige eller en negativ vurdering som vektlegger momenter som tilsier at formålene ikke er forenelige. Innfallsvinkel vil ikke få stor betydning for konklusjonen, da man uansett vil måtte kartlegge momenter som taler både for at formålene er forenlige og uforenlige. Om det nye formålet

102 Kotschy (2021) s. 76.

103 Jf. Prop.56 LS (2017-2018) s. 37. For samme forståelse, se Jarbekk (2019) s. 164 og Öman (2019) s.216. Motsatt se Kotschy (2021) s. 76 og STOA (2020) s. 52. Mer forbeholdent, se Skullerud (2019) s. 194.

104 Skullerud (2019) s. 194.

105 A29-gruppen (2013) s. 20 flg.

106 Schartum (2020) s. 102 og Jarbekk (2019) s. 143.

107 Se Öman (2019) s. 119 og Motzfeldt og Abkenar (Red.) (2019) s. 247.

må vurderes mot flere innsamlingsformål vil det eksempelvis ikke utgjøre noen forskjell om det kreves at det nye formålet er forenelig med alle opprinnelige innsamlingsformål eller om det nye formålet ikke kan være uforenelig med noen av innsamlingsformålene. Da det ikke har stor praktisk betydning om man holder seg til den positive eller den negative formuleringen, vurderer jeg det som lite problematisk å omformulere kravet til å gjelde et spørsmål om en viderebehandling er «forenelig» med det opprinnelige formålet.

Kravet om at en viderebehandling må være «forenelig» beror på en skjønnsmessig vurdering, og må vurderes konkret i den enkelte sak. Som beskrevet tidligere, er det i PVF art. 6(4) angitt relevante momenter ved vurderingen av om en ny behandling er forenelig med det opprinnelige formålet for innsamlingen. Kriteriene er ikke uttømmende (jf. formuleringen «blant annet»), og kan derfor suppleres av andre momenter i en helhetsvurdering. Kriteriene er en kodifisering av momenter angitt av A29-gruppen i deres retningslinjer om formålsbegrensning,¹⁰⁸ og retningslinjene kan derfor gi bidrag ved anvendelsen av kriteriene i art. 6(4). Dette må gjelde selv om akkurat denne uttalelsen ikke er blant de retningslinjene Det europeiske personvernrådet har gitt sin tilslutning til.¹⁰⁹

3.2.3.2 Trening av maskinlæringsalgoritmer som et «forenelig» viderebehandlingsformål

Tema i det følgende er i hvilken utstrekning bruk av opplysninger til utvikling av prediksjonsmodeller som skal benyttes i beslutningsstøtteverktøy vil være «forenelig» med de opprinnelige formålene med innsamlingen, sett i lys av kriteriene i art. 6(4). Jeg minner om at utgangspunktet er at opplysningene er samlet inn i forbindelse med utredning etter fvl. § 17. Det primære behandlingsformålet er derfor enkeltsaksbehandling, men også slik at opplysningene kan benyttes til etterfølgende oppfølging og kontroll. Vurderingen av forenelighet må gjøres med utgangspunkt i den konkrete prediksjonsmodell og særloven som begrunner innsamling av opplysningene. Formålet med fremstillingen er derfor å få frem sentrale utgangspunkter for vurderingen av om bruk av data til trening av en maskinlæringsalgoritme er en forenelig viderebehandling.

Etter art. 6(4) bokstav a) vil *forbindelsen mellom formålene* som personopplysningene opprinnelig ble samlet inn for og formålene med den tiltenkte viderebehandlingen ha betydning. Ved vurderingen vil det være relevant å se hen til nærheten og sammenhengen mellom formålene.¹¹⁰ Det kreves ikke en slik nærhet og sammenheng at det nye formålet må være et underformål av det opp-

108 Kotschy (2020) s. 327, se A29-gruppen (2013) s. 23-27.

109 The European Data Protection Board (2018) s. 1-2.

110 A29-gruppen (2013) s. 23-24.

rinnelige formålet,¹¹¹ men jo større avstand det er mellom disse, jo mer kreves for at de skal anses for å være forenelige. Et spørsmål blir da hvilken nærhet og sammenheng det er mellom enkeltsaksbehandling, herunder oppfølging og kontroll og trening av maskinlæring som skal benyttes i beslutningsstøtte ved slike oppgaver. Vurderer man formålene isolert trekker dette i retning av manglende nærhet og sammenheng. Det umiddelbare formålet om å utvikle en prediksjonsmodell er noe annet enn bruk av opplysninger for enkeltsaksbehandling, oppfølging eller kontroll. Det er allikevel klart at det endelige formålet med behandlingen er å utvikle et verktøy som skal forbedre utføringen av slike oppgaver. Det er altså en forbindelse mellom innsamlingsformålet og den videre bruken av opplysningene, selv om denne ikke kan sies å være nær. Siden det ikke kreves en slik nærhet at det nye formålet må anses for å være et underformål og at utvikling av beslutningsstøtteverktøy henger direkte sammen med forvaltningens enkeltsaksbehandling, finner jeg at det er en nærhet og sammenheng som taler for at slike formål vil kunne være forenelige. Det samme må gjelde dersom forvaltningen vil utvikle en prediksjonsmodell som skal benyttes til oppfølging eller kontrollformål.

Etter bokstav b) skal vurderingen se hen til *i hvilken sammenheng opplysningene er samlet inn, særlig med hensyn til forholdet mellom de registrerte og den behandlingsansvarlige*. Et første moment er at relasjonen mellom forvaltningen og individet er en asymmetrisk relasjon. Individet vil ofte ikke ha et valg om vedkommende vil ha kontakt med forvaltningen. I tillegg er relasjonen preget av skjevhet i ressurser og mulighet til å ivareta egne interesser. Individet har derfor et beskyttelsesbehov, noe som taler imot at formålene kan anses for å være forenelige.¹¹² Fortalepunkt 50 angir at vurderingen må ta hensyn til den registrertes rimelige forventninger på grunnlag av forholdet til den behandlingsansvarlige. Et spørsmål blir da om det er innenfor borgerens rimelige forventninger at forvaltningen benytter opplysninger som er samlet inn for enkeltsaksbehandling til trening av prediksjonsmodeller. Det må antas at den enkelte vil forvente at forvaltningen benytter opplysningene til å utføre oppgaver de er pålagt etter særlovgivningen, slik som saksbehandling, oppfølging og kontrollvirksomhet. Dette inkluderer trolig også en forventning om at forvaltningen benytter opplysningene til å organisere utførelsen av slike oppgaver på en mest mulig hensiktsmessig måte, herunder ved å utvikle digitale verktøy. Siden prediksjonsmodeller i beslutningsstøtte fortsatt er ny teknologi er det lite trolig at den enkelte vil ha en forventning om at forvaltningsorganet vil benytte data til dette formålet. Det har derfor formodningen mot seg at det vil være innenfor de registrertes

111 Kuner (2020) s. 341.

112 A29-gruppen (2013) s. 24.

rimelige forventinger at opplysninger benyttes til et slikt formål. Dette taler for at utvikling av prediksjonsmodeller ikke skal anses for å være forenelige formål.

Kriteriene i bokstav c), d) og e) vektlegger faren for negative konsekvenser for den registrerte og hvorvidt disse begrenses gjennom kompenserende tiltak. Bokstav c) angir at vurderingen skal se hen til *arten av de personopplysningene* som vil bli benyttet i viderebehandlingen, især om det er snakk om særlige kategorier personopplysninger. Kriteriet i bokstav c) må ses i sammenheng med kriteriene i bokstav d) og e). Bokstav e) angir at eksistensen av *nødvendige garantier* vil være relevante vurderingen. Betydningen av at det behandles særlige kategorier personopplysninger reduseres dersom den behandlingsansvarlige iverksetter kompenserende tiltak som i tilstrekkelig grad minsker risikoen for negative konsekvenser som følger av behandlingen. Bestemmelsen selv nevner kryptering og pseudonymisering¹¹³ som mulige kompenserende tiltak. Også kriteriet om å ta hensyn til de *mulige konsekvensene av den tiltenkte viderebehandling for de registrerte* i bokstav d) må ses i sammenheng med det overnevnte. Ved vurderingen av hvilke konsekvenser behandlingen vil ha for den registrerte vil eventuell behandling av særlige kategorier personopplysninger ha betydning, og i tilfelle om det finnes garantier som veier opp for den økte risikoen.

Det er trolig at en del prediksjonsmodeller vil trenes på særlige kategorier av personopplysninger etter art. 9(1), for eksempel helseopplysninger. Det er også sannsynlig at en del prediksjonsmodeller vil benytte opplysninger som gjelder «noens personlige forhold», og således er underlagt taushetsplikt etter fvl. § 13. Selv der taushetsbelagte opplysninger ikke omfattes av kategoriene som angis i art. 9(1), vil de trolig måtte vektlegges etter bokstav c). Det skyldes at også slike opplysninger leder til at behandlingen innebærer høyere risiko for den registrerte. I den utstrekning det benyttes særlige kategorier personopplysninger etter art. 9(1) eller taushetspliktige opplysninger etter fvl. § 13 ved utvikling av en prediksjonsmodell vil det være et moment som trekker i retning av at formålene ikke skal anses for å være forenelige. Betydningen av at det behandles slike opplysninger vil allikevel kunne reduseres dersom forvaltningen iverksetter kompenserende tiltak, jf. bokstav e).

Etter bokstav d) vil formål lettere aksepteres som forenelige dersom konsekvensene av behandlingen for den registrerte er små eller uvesentlige.¹¹⁴ Det kan det stilles spørsmål ved hvorvidt bruk av opplysninger til trening av en maskin-

113 Pseudonymisering er behandling av personopplysninger på en slik måte at de ikke lenger kan knyttes til en bestemt registrert uten bruk av tilleggsopplysninger, jf. PVF art. 4(5).

114 Skullerud (2019) s. 192.

læringsalgoritme vil ha konsekvenser for de registrerte. En mulig konsekvens for den registrerte er informasjonssikkerhetsrisikoen som følger med når opplysninger behandles digitalt.¹¹⁵ Menneskelige eller tekniske feil under utviklingen kan lede til at opplysningene endres, går tapt eller til at de avsløres for uvedkommende. Jeg vurderer at utover risikoen knyttet til informasjonssikkerhet, vil den enkelte i liten grad påvirkes direkte av behandlingen. At opplysninger brukes til trening av et profileringsverktøy medfører ikke den samme faren for negative konsekvenser som det å bli utsatt for selve profileringen.¹¹⁶ Det er allikevel en hypotetisk mulighet for at den ferdigutviklede modellen vil ha direkte konsekvenser for den registrerte. Om den registrerte kommer i kontakt med forvaltningen kan vedkommende selv risikere å bli profilert. Profileringen kan ha et «positivt» formål og bidra til at personen får bedre oppfølging fra forvaltningen, men formålet kan også være «negativt» og ha betydning for om hvorvidt man får innvilget en ytelse eller hvorvidt man plukkes ut for kontroll. Spørsmålet er om den teoretiske muligheten for at den registrerte senere kan profileres av prediksjonsmodellen skal inkluderes i vurderingen av konsekvenser etter bokstav d). Min oppfatning er at dette er hensyn som ikke bør vektlegges. Personvernforordningen skal beskytte personopplysningsvernet til den registrerte i relasjon til den enkelte behandling.¹¹⁷ I dette tilfellet gjelder behandlingen trening av en maskinlæringsalgoritme, og utover informasjonssikkerhetsrisikoen, vil den registrerte ikke påvirkes direkte av denne behandlingen.

På bakgrunn av at de registrerte i liten grad vil påvirkes av viderebehandlingen, finner jeg at bokstav c), d) og e) taler for utvikling av prediksjonsmodeller kan anses for å være forenelige formål. Dette forutsetter at det iverksettes tilstrekkelige tiltak som kan kompensere for risikoen for negative konsekvenser, særlig dersom forvaltningen benytter opplysninger etter art. 9(1) eller fvl. § 13. Iverksetting av slike tiltak vil allerede være pålagt den behandlingsansvarlige etter art. 32(1) og art. 24(1). Blant mulige tiltak er pseudonymisering, mulighet for å begrense kretsen av personer med tilgang til opplysningene, god ivaretagelse av informasjonssikkerhet, samt taushetsplikt for alle involverte.

3.2.3.3 Samlet vurdering

Formålet med viderebehandlingsforbudet i art 5(1)(b) er å sørge for at den enkelte skal kunne overskue konsekvensene av en personopplysningsbehandling, slik at den registrerte kan ivareta egne interesser etter forordningen. Allikevel har personvernforordningen gitt en viss fleksibilitet for den behandlingsansvarlige gjennom muligheten for å behandle opplysninger for forenelige formål.

¹¹⁵ STOA (2020) s. 46.

¹¹⁶ Se avsnitt 2.3. Se også STOA (2020) s. 46-47.

¹¹⁷ Personvernprinsippene i art. 5(1) vurderes opp mot den enkelte «behandlingen» av personopplysninger.

En lempelig tolkning av rammene for viderebehandlingsforbudet vil gi større frihet for behandlingsansvarlige i situasjoner der behandlingsformål ikke er blitt uttrykt ved innsamlingen av opplysningene. Motsatt vil en streng linje kunne medføre at personvernforordningen blir en hindring for de som behandler personopplysninger innenfor forordningens virkeområde. Det er altså nødvendig å finne en balanse mellom *vern* og *fleksibilitet* i formålsbegrensningsprinsippet. Dette dilemmaet er bakgrunnen for påstanden om at personvernforordningen og kunstig intelligens er uforenelige.¹¹⁸ Enkelte forfattere har således tatt til orde for en dynamisk tolkning av personvernforordningen og formålsbegrensningsprinsippet, der håndtering av risiko åpner for lempeligere anvendelse av bestemmelser som ellers vil begrense muligheten for å utvikle kunstig intelligens.¹¹⁹

En mulig tilnærming til vurderingen kan være se hen til anvendelsesområdet for det konkrete beslutningsstøtteverktøyet. Det kan for eksempel argumenteres for at opplysninger bør kunne benyttes til å utvikle maskinlæringsalgoritmer som kun vil være gunstige for borgeren. Med utgangspunkt i en slik betraktning bør eksempelvis utvikling av beslutningsstøtteverktøy som skal avdekke underforbruk av en trygdeytelse anses for å være en forenelig viderebehandling. Motsatt bør utvikling av prediksjonsmodeller som vil kunne ha negative konsekvenser for individet lede til en mer restriktiv anvendelse av formålsbegrensningsprinsippet. Dette kan være prediksjonsmodeller som skal benyttes til kontrollformål eller i enkeltsaksbehandling.

En restriktiv anvendelse av forenelighetskravet er begrunnet i at det trolig ligger utenfor den enkeltes rimelige forventninger at opplysninger skal benyttes til utvikling av maskinlæring. Dette må ses i sammenheng med at individet ofte ikke velger selv å ha kontakt med forvaltningen. Den registrertes rimelige forventninger er fremhevet som et grunnleggende moment av mange aktører,¹²⁰ og har derfor stor vekt. Dette må gjelde selv om den registrerte ikke påvirkes direkte av behandlingen i den enkelte sak. Videre finner jeg at en lempelig tolkning i stor grad vil utarme den beskyttelsen viderebehandlingsforbudet gir, og således tilrettelegge for formålsutglidning. Vektlegger man hvorvidt beslutningsstøtteverktøyet vil kunne være inngripende overfor den enkelte ved forenelighetsvurderingen vil dette være i samsvar med legalitetsprinsippet. En restriktiv anvendelse av forenelighetskriteriet vil tilrettelegge for at forvaltningens behandlingsformål heller må inntas i lovgivningen i medhold av art. 6(1)(e), jf. art. 6(3). En slik løsning vil sikre forutberegnelighet for individet.

118 ICO (2017) s. 37.

119 Stalla-Bourdillon and Knight (2018) s. 1 og STOA (2020) s. 76.

120 Se fortalepunkt 50, A29-gruppen (2013) s. 24, og ICO (2017) s. 38.

3.2.4 Unntak fra viderebehandlingsforbud for «statistiske formål», jf. art. 5(1)(b), jf. art. 89 nr. 1

Hvis utvikling av prediksjonsmodeller ikke anses for å være forenelig med de opprinnelige formålene for innsamlingen, er spørsmålet om et slikt formål kan omfattes av alternativet «statistiske formål» etter art. 5(1)(b). Som angitt innledningsvis åpner art. 5(1)(b) for at personopplysninger kan viderebehandles for bestemte sekundærformål forutsatt at den behandlingsansvarlige sørger for at behandlingsopplegget overholder krav i art. 89 nr. 1.

I fortalepunkt 162 uttales at statistiske formål skal forstås som «enhver innsamling og behandling av personopplysninger som er nødvendig i forbindelse med *statistiske undersøkelser* eller for å fremskaffe *statistiske resultater*» (min utheving). Som forklart i avsnitt 2.2 anvendes statistiske metoder ved utvikling av prediksjonsmodeller, og prediksjonsmodellen er en representasjon av statistiske korrelasjoner læringsalgoritmen har funnet i et treningsmateriale. Utvikling av prediksjonsmodeller forutsetter altså statistiske undersøkelser, og prediksjonsmodellen vil således kunne representere statistiske resultater.

Det kan stilles spørsmål ved om «statistiske formål» er ment å omfatte viderebehandling av data til utvikling av prediksjonsmodeller som igjen skal brukes til å gjøre prediksjoner om individer. En naturlig språklig forståelse av «statistiske formål» er at formålet er å fremskaffe statistisk informasjon, som igjen kan danne grunnlag for blant annet analyse, vurderinger eller beslutninger. Utvikling av teknologiske verktøy som skal brukes i beslutningsstøtte kan intuitivt anses for å innebære noe annet. Dette henger sammen med at hovedformålet ikke er «statistikk» i seg selv, men heller å utvikle et verktøy i form av en prediksjonsmodell. Den statistiske undersøkelsen er i så måte bare en del av metoden som anvendes for å utvikle modellen

Det er allikevel hevdet at «statistiske formål» åpner for at data kan viderebehandles ved anvendelse av maskinlæring.¹²¹ Forutsetning er i så fall at «statistiske formål» kun omfatter anvendelsen av data til å utvikle selve prediksjonsmodellen, og ikke den videre bruken av denne i beslutningsstøtte senere. Dette skyldes at det i fortalepunkt 162 uttrykkelig fastslås at bruk av det statistiske resultatet som støtte til «avgjørelser som gjelder en bestemt fysisk person» ikke omfattes av «statistiske formål». Således avgrenses det mot viderebehandling der statistiske resultater brukes til profilering, jf. avsnitt 2.3 ovenfor.

Spørsmålet om hvorvidt «statistiske formål» omfatter bruk av data til utvikling av selve prediksjonsmodellen med den følge at slik bruk av data medfører en

121 STOA (2020) s. 77.

lovlig viderebehandling etter art. 5(1)(b), må avgjøres i praksis. Min oppfatning er at en naturlig språklig forståelse trekker i retning av at bruk av data til utvikling av prediksjonsmodeller er noe annet en «statistiske formål», selv om metodikken som anvendes i stor grad bygger på statistikk.

3.2.5 Oppsummering

For at forvaltningen skal kunne benytte opplysninger ved anvendelse av maskinlæring med grunnlag i art. 6(1)(e) må det fremgå av særlovgivningen at forvaltningen kan benytte opplysninger til dette formålet *før* opplysningene samles inn. I henhold til art. 5(1)(b) må formålene formuleres på en tilstrekkelig klar måte.

Det bør stilles samme krav til nye behandlingsformål, som til opprinnelige formål. Som en følge av dette må forvaltningen sikre at formålet med beslutningsstøtteverktøyet er berettiget, at formålet formuleres og formidles til den registrerte etter art. 13 eller 14 og at forvaltningsorganet så langt det lar seg gjøre søker å konkretisere formålet med modellen på en presis og nøyaktig måte.

I vurderingen av om bruk av opplysninger ved anvendelse av maskinlæring er en forenelig viderebehandling, bør en særlig se hen til hvorvidt beslutningsstøtteverktøyet vil kunne ha negative konsekvenser for individet. Ved risiko for negative konsekvenser, taler hensynet til forutberegnelighet for at viderebehandlingen ikke anses for å være forenelig. Det er allikevel klart at forenelighet må vurderes i den konkrete situasjonen der en prediksjonsmodell skal utvikles. Forenelighetsvurderingen i den enkelte sak vil særlig påvirkes av hvordan synet på maskinlæring utvikles i samfunnet, hvilken informasjon som gis den registrerte ved innsamlingen, hva prediksjonsmodellen skal benyttes til, samt hvilke garantier som iverksettes for å begrense mulige negative konsekvenser av behandlingen.

Avslutningsvis, har jeg funnet at utvikling av maskinlæring trolig ikke vil anses for å være «statiske formål» og derfor ikke omfattes av sekundærformålene i art. 5(1)(b).

Konsekvensen av en streng tolkning av viderebehandlingsforbudet er at formålsbegrensningsprinsippet i art. 5(1)(b) vil begrense muligheten til å anvende innsamlede opplysninger til å utvikle verktøy gjennom anvendelse av maskinlæring. Dersom det ikke finnes grunnlag for viderebehandling i lov eller samtykke etter art. 6(4), vil opplysningene måtte samles inn på nytt dersom de skal behandles for utvikling av ML-modeller.¹²² Etter hvert som maskinlæring blir

122 Prop. 56 LS (2017–2018) s. 37.

mer utbredt i samfunnet, vil trolig behandlingsgrunnlag sikres ved at dataanalyse og maskinlæring angis som behandlingsformål i lovgivningen. Viderebehandlingsforbudet vil derfor få mindre betydning for forvaltningens mulighet for å utvikle ML-modeller. Viderebehandlingsforbudet begrenses også av muligheten til å anonymisere personopplysninger,¹²³ samt muligheten til å anvende syntetiske testdata.¹²⁴

3.3 Forbud mot å dele taushetsbelagte opplysninger med andre, jf. fvl. §§ 13 flg.

3.3.1 Innledning

Forvaltningsloven § 13 angir at «enhver som utfører tjeneste eller arbeid for et forvaltningsorgan, plikter å hindre at andre får adgang eller kjennskap til det han i forbindelse med tjenesten eller arbeidet får vite om noens personlige forhold». Hovedregelen i § 13 begrenses av unntak i §§ 13a flg. Hensynet bak taushetsplikten er å beskytte private interesser som kan bli lidende av at forvaltningen deler personlige opplysninger.¹²⁵ Således er taushetsplikten også begrunnet i hensynet til tillitt til forvaltningen. Det er viktig at private skal kunne samhandle med forvaltningen uten å risikere at personlige opplysninger avsløres. Taushetsplikten spiller derfor en viktig rolle for at forvaltningen skal kunne utføre sin funksjon i samfunnet.

Som fremholdt i avsnitt 2.2.2 omfatter taushetsplikten opplysninger som det er «vanlig å ville holde for seg selv» som er knyttet til en bestemt person. Spørsmålet beror på en skjønnsmessig vurdering, der det avgjørende er hva som normalt vil oppleves som utleverende eller krenkende.¹²⁶ Forarbeidene nevner som eksempler at opplysningene i seg selv kan være følsomme, at situasjonen opplysningene har fremkommet i kan tilsi at de beskyttes eller at en kombinasjon av flere opplysninger kan medføre en fare for utlevering av en person.¹²⁷ I tillegg kan taushetsplikt følge av særlovgivning. Det er vanskelig å si noe generelt om forvaltningens bruk av taushetspliktige opplysninger ved trening av prediksjonsmodeller, men i den utstrekning slik informasjon vil omfatte helseopplysninger vil også taushetsplikten gjelde.¹²⁸ Det samme gjelder opplysninger som viser at man har et klientforhold til et forvaltningsorgan som kan anses som

123 Se allikevel om utfordringer knyttet til anonymisering ved dataanalyse i avsnitt 2.2.2.

124 Ringnes (2019).

125 Ot.prp.nr.3 (1976–1977) s. 13.

126 Graver (2019) s. 369.

127 Ot.prp.nr.3 (1976–1977) s. 13-14.

128 Ot.prp.nr.3 (1976–1977) s. 14.

belastende.¹²⁹ Det er naturlig å anta at hvis forvaltningsområdet som skal omfattes av prediksjonsbasert beslutningsstøtte behandler taushetspliktige opplysninger, vil det benyttes slike opplysninger i utviklingen av modellen, samt at de samme typer opplysninger vil være grunnlag for selve beslutningsstøtten.

Spørsmålet etter dette er hvilken betydning taushetsplikten har for bruk av taushetspliktige opplysninger til trening av maskinlæring. For det første følger det av ordlyden i fvl. § 13 at taushetsplikten er en personlig plikt som påligger individet som har mottatt den taushetspliktige opplysningen. Bestemmelsen angir videre at taushetsplikten medfører en plikt til å «hindre at andre får adgang eller kjennskap» til opplysningen. Plikten til å «hindre» krever både at opplysningene ikke aktivt deles og at personen sørger for at opplysningene er tilstrekkelig beskyttet, jf. fvl. § 13 c (2). Videre viser angivelsen av «andre» at taushetsplikten i utgangspunktet gjelder overfor alle. Taushetsplikten gjelder derfor all formidling av opplysninger, også om det skjer innenfor forvaltningsorganet den aktuelle personen tilhører. Dette gir opphav til flere spørsmål. For det første er det et spørsmål om taushetsplikten også gjelder ved bruk av opplysninger til trening av maskinlæring. Hvis ja, blir spørsmålet om noen av forvaltningslovens unntak fra taushetsplikten kan komme til anvendelse.

3.3.2 Gjelder taushetsplikten ved bruk av opplysninger til trening av maskinlæring?

Som angitt ovenfor vil taushetsplikten i alminnelighet gjelde all formidling av taushetsbelagte opplysninger, også innad i forvaltningsorganet. Utgangspunktet er derfor at i den utstrekning bruk av opplysninger til trening av maskinlæring kan medføre «at andre får adgang eller kjennskap til opplysningene», vil taushetsplikten gjelde. I hvilken utstrekning dette faktisk vil skje, beror på hvordan arbeidet med å utvikle prediksjonsmodeller organiseres og utføres. Arbeidet med maskinlæringen kan muligens organiseres på en slik måte at de som arbeider med utviklingen ikke får adgang eller kjennskap til treningsdataene som benyttes. Videre vil den ferdigutviklede prediksjonsmodellen ikke inneholde opplysningene som er benyttet ved treningen, jf. ovenfor i avsnitt 2.1. Dermed utviklingen av ML-modellen ikke medfører noen fare for at personer får adgang eller kjennskap til opplysningene, vil taushetsplikten således ikke begrense bruken. Om de derimot får en slik adgang eller kjennskap, vil taushetsplikten hindre at taushetspliktige opplysninger benyttes ved utviklingen.

Taushetsplikten gjelder altså ved utvikling av prediksjonsmodeller, og kan begrense forvaltningens tilgang til opplysninger ved trening av maskinlæring. Det kan allikevel være grunnlag for en videre anvendelse av taushetspliktreglene

129 Graver (2019) s. 370.

enn hva som følger direkte av ordlyden, slik at taushetsplikten også gjelder for det tilfelle at ingen annen *person* risikerer å få adgang eller kjennskap til opplysningene. NOU 2019:5 om ny forvaltningslov synes å fremme et slikt synspunkt. Ved vurderingen av taushetspliktens innhold etter gjeldende rett uttales: «Med «andre» menes først og fremst andre personer, men taushetsplikten må antakelig også gjelde overfor et datasystem.»¹³⁰ Synpunktet er knyttet til tilfeller der saksbehandlingen er helt eller delvis automatisert, og hvor et datasystem har tilgang til opplysninger som er lagret i andre datasystemer. Rekkevidden av uttalelsen er uklar, og utredningen drøfter ikke spørsmålet videre. Det er foreløpig ikke vedtatt ny forvaltningslov (per april 2021), og uttalelsen har derfor liten rettskildevekt. Spørsmålet som reises er allikevel interessant, og gir grunnlag for prinsipielle drøftelser om hva taushetsplikten handler om i relasjon til elektronisk behandling av opplysninger.

Utredningen uttaler seg om «datasystemer» i relasjon til helt eller delvis automatisert saksbehandling, men det kan tenkes at begrunnelsen også vil gjøre seg gjeldende for maskinlæringsalgoritmer (som også er datasystemer med en viss grad av autonomi). Således blir spørsmålet om taushetsplikten også begrenser deling av opplysninger til en maskinlæringsalgoritme som skal utvikle en prediksjonsmodell. En slik løsning må følge av at de hensyn som begrunner taushetsplikten gjør seg gjeldende på en slik måte at det begrunner en analogisk anvendelse av taushetspliktreglene. Som angitt ovenfor er hensynet bak taushetspliktreglene å verne private parter som kan bli skadelidende av at opplysninger avsløres, og således også ivareta individets tillit til forvaltningen. Jeg anser det som lite trolig at interessene til en privat part vil bli skadelidende som en følge av at opplysningene benyttes til å trene en prediksjonsmodell. Det kan allikevel stilles spørsmål ved om en slik bruk kan utfordre tilliten til forvaltningen ved at forvaltningen benytter opplysningen til noe annet enn hva den er avgitt for. Hvorvidt bruk av taushetsbelagte opplysninger til utvikling av maskinlæring truer disse interessene på en slik måte at taushetsplikten må komme til anvendelse må trolig vurderes konkret i relasjon til det enkelte maskinlæringsprosjekt. Jeg antar allikevel at det bærende hensynet bak taushetspliktreglene er den skade og ulempe som oppstår ved at andre *personer* får kjennskap til opplysningene, og at begrunnelsen for å anvende taushetspliktreglene analogisk utover disse tilfellene mangler.

Konklusjonen etter dette blir at taushetspliktreglene vil gjelde i den utstrekning bruk av taushetsbelagte opplysninger til trening av maskinlæring vil medføre at andre personer får adgang eller kjennskap til opplysningene. Hvorvidt det er tilfelle, vil bero på hvordan utviklingsarbeidet organiseres og utføres. Om taus-

130 NOU 2019:5 s. 283.

hetsplikten også gjelder deling av opplysninger til datasystemer, er et spørsmål som må avgjøres i praksis, eventuelt ved vedtagelse av ny forvaltningslov. Det ble ikke fremmet forslag om taushetspliktregler som omfatter datasystemer i NOU 2019:5.¹³¹

3.3.3 Unntak fra taushetsplikten

Når det er lagt til grunn at taushetsplikten vil gjelde ved utvikling av prediksjonsmodeller, blir spørsmålet om noen av unntakene fra taushetsplikten kommer til anvendelse. Forvaltningsloven § 13 b nr. 2 angir for det første at opplysningene kan benyttes til det «formål de er gitt eller innhentet for». Unntaket har en parallell i formålsbegrensningsprinsippet som også viser sammenhengen mellom behandling i tråd med innsamlingsformålet og legitim behandling. Betydningen av formål etter § 13 b nr. 2 er allikevel snevrere. Unntaket gjelder behandling av den enkelte sak, da det er disse formålene opplysningene er samlet inn for.¹³² Dette illustreres av at bestemmelsen selv gir eksempler på typiske aktiviteter knyttet til enkeltsaksbehandling, slik som «saksforberedelse, avgjørelse, gjennomføring av avgjørelsen, oppfølging og kontroll». I tråd med diskusjonen i avsnitt 3.2 ovenfor er det klart at utvikling av prediksjonsmodeller er et annet formål enn enkeltsaksbehandling, oppfølging og kontroll, selv om det overordnede formålet kan være det samme. Unntaket er derfor ikke aktuelt ved utvikling av ML-baserte beslutningsstøtteverktøy.

Det følger videre av fvl. § 13 b nr. 3 at opplysningene kan gjøres tilgjengelig «for andre tjenestemenn innen organet eller etaten i den utstrekning som trengs for en hensiktsmessig arbeids- og arkivordning». Begrensningen av tilgjengeliggjøringen til andre «tjenestemenn» trekker i retning av at unntaket gjelder noe annet enn at opplysningene skal tilgjengeliggjøres for interne formål i organet som sådan. En mulig forståelse er at opplysningene må være nødvendige for at andre i forvaltningsorganet skal kunne løse sine arbeidsoppgaver. Dette støttes av at bestemmelsen gir som eksempel at opplysninger kan benyttes til «bruk ved vegledning i andre saker». At bestemmelsen begrenser personkretsen til «tjenestemenn» og i tillegg krever at bruken må være nødvendig for en hensiktsmessig arbeidsordning, trekker i retning av at bestemmelsen ikke hjemler unntak for bruk av taushetsbelagte opplysninger til generelle formål innad i forvaltningsorganet. Jeg finner derfor at unntaket ikke medfører at taushetsbelagte opplysninger kan benyttes til utvikling av maskinlæring.

Videre følger det av § 13 b nr. 4 at taushetsplikten ikke begrenser at opplysningene «brukes for statistisk bearbeiding». Unntaket har en parallell til unntaket

131 NOU 2019:5 s. 26-28.

132 Woxholt (2011) s. 294.

for «statistiske formål» etter PVF. art. 5(1)(b), allikevel slik at formuleringen «statistisk bearbeiding» må anses for å favne videre. Også «statistisk bearbeiding» må tolkes slik at det er det statistiske formålet som er avgjørende, og ikke det forhold at det skjer en statistisk bearbeiding av opplysninger. Utvikling av beslutningsstøtteverktøy blir således noe annet. I tråd med diskusjonen i avsnitt 3.2.4 blir konklusjonen at «statisk bearbeiding» ikke omfatter utvikling av maskinlæring.

Forvaltningen kan anvende taushetspliktige opplysninger til trening av maskinlæring når den som har krav på taushet «samtykker», jf. fvl. § 13 a nr. 1. Regelen har en parallell i unntaket fra viderebehandlingsforbudet ved samtykke etter PVF art. 6(4). Lovene har allikevel en noe ulik begrunnelse bak unntaket for samtykke. Overskriften i forvaltningsloven § 13 a viser at unntakene gjelder situasjonen der «det ikke er behov for beskyttelse». Hensynet bak taushetsplikten er å beskytte den som kan påvirkes negativt av at opplysninger deles med andre, og denne begrunnelsen faller bort når personen samtykker til at opplysningene deles. Selv om personvernforordningen også har beskyttelse av den registrertes interesser som sentralt formål, er samtykket særlig begrunnet i hensynet til den registrertes autonomi og selvbestemmelse.¹³³ Regelsettene har altså en sammenfallende løsning, men med en noe ulik begrunnelse.

Taushetsbelagte opplysninger kan også brukes til å utvikle prediksjonsmodeller hvis opplysningene gis i «statistisk form eller ved at individualiserende kjennetegn utelates på annen måte», jf. § 13 a nr. 2. At individualiserende kjennetegn utelates kan bety at opplysningene pseudonymiseres¹³⁴ eller anonymiseres.¹³⁵ Også her er begrunnelsen at beskyttelsesbehovet opphører når det ikke er mulig å identifisere den opplysningene gjelder. I motsetning til unntaket for samtykke har ikke regelen en lovfestet parallell i personvernforordningen. I tillegg gjelder personvernforordningen selv når opplysningene pseudonymiseres. Begrunnelsen er at det er mulig å identifisere den registrerte, og at opplysningene derfor er personopplysninger etter art. 4(1). Personvernforordningen vil ikke gjelde når opplysninger anonymiseres,¹³⁶ men som påpekt i avsnitt 2.2.2 kan anonymisering være utfordrende ved dataanalyse. Da fvl. § 13 a nr. 2 trolig omfatter både pseudonymisering og anonymisering vil unntaket ha et videre anvendelsesområde enn hva som er tilfelle etter personvernforordningen.

133 Schartum (2020) s. 113.

134 Jf. PVF art. 4(5).

135 Schartum (2015) s. 148.

136 Jf. PVF art. 2(1), jf. art. 4(1). Se også fortalepunkt 26.

3.3.4 Oppsummering

Taushetspliktreglene vil bare kunne begrense bruk av opplysninger til utvikling av maskinlæring i den utstrekning forvaltningen vil benytte taushetsbelagte opplysninger etter fvl. § 13 i treningsmaterialet. Videre må bruken av opplysningene medføre at «andre» kan få adgang eller kjennskap til opplysningene for at taushetsplikten skal inntreffe. Avgjørende vil være hvordan arbeidet med den konkrete prediksjonsmodellen organiseres og utføres.

Jeg har videre konkludert med at det trolig ikke er grunnlag for at taushetspliktreglene skal gjelde også i det tilfellet at ingen «andre» vil få tilgang til opplysningene, men der opplysningene deles med maskinlæringsalgoritmen.

Avslutningsvis har jeg funnet at samtykke fra den som har krav på taushet, samt pseudonymisering eller anonymisering vil gjøre at taushetspliktige opplysninger lovlig kan anvendes ved utvikling av maskinlæring.

3.4 Krav om kvalitet og minimering, jf. PVF art. 5(1)(c) og (d) og det forvaltningsrettslige prinsippet om forsvarlig saksbehandling

3.4.1 Innledning

I personvernforordningen stiller dataminimeringsprinsippet i art. 5(1)(c) og riktighetsprinsippet i art. 5(1)(d) uttrykkelige krav til mengden og kvaliteten på de personopplysninger den behandlingsansvarlige skal anvende for å oppnå et formål. Formålet med behandlingen er referanseramme når kravene etter dataminimeringsprinsippet og riktighetsprinsippet skal klarlegges. Vurderingen etter bokstav c) og d) må altså gjøres med det utgangspunkt at opplysningene skal anvendes til å utvikle en prediksjonsmodell som skal brukes i forvaltningens beslutningsstøtteverktøy.

Som påpekt, inneholder ikke forvaltningsretten direkte krav til hvordan forvaltningen skal utvikle digitale systemer.¹³⁷ Det kan allikevel utledes indirekte krav til utviklingsprosessen fra prinsippet om at forvaltningens saksbehandling skal være forsvarlig, herunder særlig kravet om at forvaltningens saksutredning skal være forsvarlig.¹³⁸ Det må derfor ut fra et forvaltningsrettslig perspektiv stilles krav til datamaterialets kvalitet og mengde i den utstrekning det har betydning for kravet om at forvaltningens saksbehandling skal være forsvarlig.

137 Schartum (2018) s. 55.

138 Schartum (2018) s. 55.

Som angitt i avsnitt 2.2.1 ovenfor har dataene som benyttes ved trening av maskinlæringsalgoritmen stor betydning for den endelige prediksjonsmodellens nøyaktighet og for hvorvidt modellen vil ha uønskede egenskaper. Dette er bakteppet når de skranker som følger av dataminimerings- og riktighetsprinsippet og prinsippet om forsvarlig saksbehandling skal klarlegges.

I det følgende gir jeg først et overblikk over dataminimerings- og riktighetsprinsippene og prinsippet om forsvarlig saksbehandling. Deretter klargjør jeg med utgangspunkt i forsvarlighetsprinsippet hvilke krav som skal stilles til forsvarlighet i ML-baserte beslutningsstøttesystemer. Avslutningsvis, vurderer jeg betydningen av dataminimerings- og riktighetsprinsippene for forvaltningens bruk av data til utvikling av prediksjonsmodeller med utgangspunkt i prinsippet om forsvarlig saksbehandling.

3.4.2 Rettslige utgangspunkter

3.4.2.1 Riktighetsprinsippet, jf. art. 5(1)(d)

Riktighetsprinsippet krever at opplysninger skal være «korrekte og om nødvendig oppdaterte» og at det må treffes «ethvert rimelig tiltak for å sikre at personopplysninger som er uriktige med hensyn til formålene de behandles for, uten opphold slettes eller rettes». Bestemmelsens første del modifiseres av at den behandlingsansvarliges plikt til å rette eller slette opplysningene kun gjelder i den utstrekning som er nødvendige for de formål opplysningene behandles for. Det er gjeldende altså ikke et absolutt krav om at opplysningene skal være korrekte, med mindre dette er nødvendig for behandlingsformålet.¹³⁹

Riktighetsprinsippet inneholder en aktivitetsplikt for den behandlingsansvarlige.¹⁴⁰ Ordlyden viser klart at den behandlingsansvarlige skal sørge for at uriktige opplysninger rettes eller slettes når behandlingsformålet tilsier det, og spiller således den registrertes rett til retting og sletting i PVF art. 16 og 17. Videre viser ordlyden at prinsippet gjelder opplysninger som behandles generelt. Riktighetsprinsippet vil både gjelde for de opplysninger den behandlingsansvarlige anvender til behandlingsformålet og for eventuelle personopplysninger behandlingen genererer. Riktighetsprinsippet krever derfor at den behandlingsansvarlige skal sørge for at beslutningsstøttesystemer ikke produserer uriktige prediksjoner om individer.

Det er behandlingsformålet som er avgjørende for de krav som stilles etter riktighetsprinsippet. Et sentralt spørsmål i avsnitt 3.4.4 er hvilke krav som skal

¹³⁹ Schartum (2020) s. 94.

¹⁴⁰ Öman (2019) s. 133.

stilles til opplysningenes riktighet når behandlingsformålet er å utvikle prediksjonsmodeller som skal benyttes i forvaltningens beslutningsstøtteverktøy.

3.4.2.2 Dataminimeringsprinsippet, art. 5(1)(c)

Det følger av dataminimeringsprinsippet at personopplysninger skal være «adekvate, relevante og begrenset til det som er nødvendig for formålene de behandles for». Dataminimeringsprinsippet må vurderes både mot antall opplysningstyper og mengden av opplysninger¹⁴¹ og mot antall registrerte det behandles opplysninger fra.¹⁴² Kravet om at opplysninger skal være «relevante» medfører at bruken av opplysningen skal kunne begrunnes i formålet. Det er ikke adgang til å behandle opplysninger som ikke har en betydning for realisering av behandlingsformålet.¹⁴³ Kravet om at opplysningene skal være «begrenset til det som er nødvendig for formålet» medfører at den behandlingsansvarlige skal søke å begrense bruken av opplysninger i den utstrekning det lar seg gjøre. I fortalepunkt 39 er det presisert at opplysninger bør behandles «bare dersom formålet med behandlingen ikke med rimelighet kan oppfylles på en annen måte.». Dette leder til at det ikke skal anvendes opplysninger man kan klare seg uten, og at dersom det finnes alternativer til behandling av opplysningene skal disse alternativene velges. I motsetning til de to andre kriteriene trekker ikke nødvendigvis adekvanskravet i retning av begrensning av mengden personopplysninger som behandles.¹⁴⁴ I stedet skal det sørges for at de personopplysningene som behandles er tilstrekkelige for å oppnå formålet med behandlingen.¹⁴⁵ Regelen kan bidra til å unngå en situasjon der mangelfulle opplysninger leder til at behandlingsformålet ikke realiseres, med den følgen at personopplysninger behandles unødvendig.¹⁴⁶

Dataminimeringsprinsippet anses å inneholde et krav om proporsjonalitet.¹⁴⁷ Kravet forutsetter at det må være et rimelig forhold mellom opplysningene som behandles og de formål den behandlingsansvarlige søker å oppnå med behandlingen.¹⁴⁸ Bruken av opplysninger må således ikke medføre et større inngrep overfor den registrerte enn det som er nødvendig. Det avgjørende vil være om bruken av personopplysninger vil være så inngripende, vidtgående eller risiko-

141 Schartum (2020) s. 93.

142 Jarbekk (red.) (2019) s. 144.

143 Schartum (2020) s. 92.

144 Schartum (2020) s. 92.

145 Motzfeldt og Abkenar (red.) (2019) s. 250.

146 Motzfeldt og Abkenar (red.) (2019) s. 250.

147 Eksempelvis Motzfeldt og Abkenar (red.) (2019) s. 142 og Datatilsynet (2018) s. 17.

148 Olsen (2020) s. 213.

fylt for de registrerte, at selv et saklig formål ikke kan legitimere behandlingen.¹⁴⁹

Dataminimeringsprinsippet henger tett sammen med riktighetsprinsippet. Uriktige opplysninger vil ofte ikke være adekvate eller relevante for behandlingsformålet. I tillegg vil det ofte måtte gjøres en avveining mellom prinsippene.¹⁵⁰ Eksempelvis kan behandlingsformålet kreve høy grad av opplysningskvalitet i henhold til riktighetsprinsippet. Dersom opplysningskvaliteten kan bedres ved at det sammenstilles opplysninger fra flere kilder, kan det tilsi at dataminimeringsprinsippet må tillegges mindre vekt.¹⁵¹

Også for dataminimeringsprinsippet vil behandlingsformålet være avgjørende for vurderingen av lovlig behandling av opplysninger. Som det fremgår i avsnitt 2.2 er en utfordring ved maskinlæring at det er vanskelig å vurdere behovet for opplysninger innledningsvis i et prosjekt. I tillegg kreves ofte store mengder opplysninger ved utvikling av maskinlæring.¹⁵² Hvilken betydning dette får ved anvendelsen av dataminimeringsprinsippet blir vurdert i avsnitt 3.4.4.

3.4.2.3 Prinsippet om forsvarlig saksbehandling, herunder utredningsprinsippet

Det følger av prinsippet om forsvarlig saksbehandling at forvaltningens avgjørelser skal bygge på et forsvarlig grunnlag. Når forsvarlighetsprinsippet vurderes i relasjon til utvikling av prediksjonsmodeller som skal brukes i beslutningsstøttesystemer, blir særlig den delen av forsvarlighetskravet som gjelder saksutredning sentral. Utredningsprinsippet krever at forvaltningens avgjørelser skal ha et forsvarlig faktisk og juridisk grunnlag.¹⁵³ Det grunnleggende hensynet er å sikre rettsriktige resultater, og på den måten beskytte individets rettigheter.¹⁵⁴ Fra hensynet til rettsriktighet kan det trekkes en linje til riktighetsprinsippet i PVF art. 5(1)(d) som også stiller krav til at opplysninger behandlingen produserer skal være riktige. Både det forvaltningsrettslige utredningsprinsippet og det personvernrettslige riktighetsprinsippet medfører derfor en plikt for forvaltningen til at beslutninger tas på et grunnlag som kan lede til riktige avgjørelser.

Utredningsprinsippet stiller krav til forvaltningens utredning, og må derfor medføre krav til forvaltningens beslutningsstøttesystemer. Forvaltningen kan kun anvende beslutningsstøttesystemer når de bidrar til å gi saksbehandleren et

149 Motzfeldt og Abkenar (red.) (2019) s. 251.

150 Schartum (2020) s. 95.

151 Schartum (2020) s. 95.

152 STOA (2020) s. 16.

153 Schartum (2018) s. 45.

154 NUT 1958:3 s. 160.

forsvarlig avgjørelsesgrunnlag. Som en følge av at opplysningskvaliteten vil ha stor betydning for kvaliteten til en prediksjonsmodell, må utredningsprinsippet medføre krav til det treningsmaterialet som anvendes. I denne sammenheng fungerer prinsippet om forsvarlig saksbehandling, herunder utredningsprinsippet, som en selvstendig rettslig skranke for treningsmaterialet som anvendes ved maskinlæring.

Når forsvarlighetsprinsippet generelt, og utredningsprinsippet spesielt, krever at beslutningsstøttesystemer skal være forsvarlige, blir spørsmålet hvilke krav som skal stilles for at et ML-basert beslutningsstøttesystem skal være forsvarlig. Dette har også betydning for anvendelsen av riktighetsprinsippet og datamini-meringsprinsippet. Vurderingen etter begge bestemmelsene må gjøres med utgangspunkt i hva som er «nødvendig for formålene». Når «formålene» må forstås med en forutsetning om at beslutningsstøttesystemer skal være forsvarlige, vil forsvarlighetsprinsippet fungere som tolkningsmoment ved klarleggingen av innholdet etter personvernprinsippene.

3.4.3 Forsvarlighet i ML-baserte beslutningsstøttesystemer

Spørsmålet i det følgende er hvilke krav som stilles for at forvaltningens beslutningsstøttesystemer skal være forsvarlige. Utgangspunktet er at kravene som stilles til forsvarlighet i saksbehandlingen beror på en skjønnsmessig vurdering med utgangspunkt i den enkelte sak.¹⁵⁵ Forsvarlighet kan altså ikke avgjøres absolutt, og terskelen for hva man aksepterer av mangler i beslutningsstøttesystemet vil variere beroende på de konkrete forholdene. Vurderingen av forsvarlighet i relasjon til beslutningsstøttesystemer vil derfor i stor grad bero på en risikovurdering, dvs. sannsynligheten for negative konsekvenser, samt hvor alvorlige disse er.¹⁵⁶ Dersom et beslutningsstøttesystem medfører en uakseptabel risiko, tilsier forsvarlighetsprinsippet at forvaltningen må iverksette kompenserende tiltak. Kravet om forsvarlig saksbehandling medfører dermed et krav om å iverksette kompenserende tiltak i den utstrekning beslutningsstøttesystemer medfører en uakseptabel risiko.¹⁵⁷

Forvaltningens plikt til å iverksette tiltak som kan redusere uakseptabel risiko har en parallell i personvernforordningen. Håndtering av risiko gjennom å iverksette tiltak er et sentralt virkemiddel for å sikre gjennomføring av forordningens formål. Dette kommer til uttrykk i reglene om den behandlingsansvarliges ansvar i art. 24(1), om innebygget personvern i art. 25, om behandlingssikkerhet i art. 32 og om personvernkonsekvensvurderinger i art. 35. Plikten til å

¹⁵⁵ Schartum (2018) s. 45.

¹⁵⁶ Schartum (2018) s. 45.

¹⁵⁷ Schartum (2018) s. 45. Se Boe (2018) s. 174 om aktivitetsplikt som krav for forsvarlighet i forvaltningens systeminnretning.

begrense risiko etter forordningen er knyttet til «fysiske personers rettigheter og friheter», jf. eksempelvis art. 24(1). «Rettigheter og friheter» henviser ikke bare til rettigheter etter forordningen, men anses også for å omfatte andre grunnleggende rettigheter og friheter.¹⁵⁸ Risiko for den registrertes rettssikkerhet i møte med forvaltningen har derfor også relevans etter personvernforordningen på en slik måte at den behandlingsansvarlige må iverksette kompensierende tiltak. Krav om forsvarlighet etter forvaltningsretten skal derfor vektlegges ved vurdering av personvernforordningens bestemmelser.

Med utgangspunkt i at vurderingen av forsvarlighet beror på risikoen knyttet til beslutningsstøttesystemer og den tilhørende plikten til å iverksette kompensierende tiltak, blir spørsmålet hvordan risikoen som er knyttet til utvikling av prediksjonsmodeller skal vurderes. Vurderingen må gjøres konkret for den enkelte prediksjonsmodell, men det er mulig å indikere generelle faktorer som vil ha betydning for risikoen forbundet med prediksjonsmodeller i beslutningsstøtte.

Et første moment er at beslutningsstøtteverktøyet potensielt vil ha betydning for et høyt antall mennesker i svært mange saker.¹⁵⁹ Konsekvensene av at systemet inneholder feil, for eksempel i form av uriktige prediksjoner, kan derfor bli store. Konsekvensen av feil må ses i sammenheng med at beslutningsstøttesystemer, i motsetning til rettslige beslutningssystemer, ikke vil bli brukt til å fatte enkeltvedtak automatisk.¹⁶⁰ Prediksjonen vil kun spille en rolle som veiledning eller bidrag til saksbehandlerens vurderinger, og vil ikke få direkte virkning for parten. Eksempelvis vil et beslutningsstøttesystem som peker ut individer for kontroll for trygdesvindler lede til at saksbehandler vurderer saken til individet, og ikke at personen ilegges reaksjoner direkte. Det vil derfor skje en viss kontroll av beslutningsstøttesystemets vurdering, med den følge at feil kan oppdages og korrigeres.

Betydningen av at beslutningsstøttesystemet ikke får direkte virkning for en part forutsetter at en saksbehandler gjør en selvstendig vurdering av saken. Der som prediksjonen tillegges vekt, eksempelvis fordi saksbehandleren stoler på at systemet er riktig, vil selv beslutningsstøttesystemet få direkte konsekvenser for parten. Dette må ses i sammenheng med at det kan være vanskelig å forstå hvorfor et beslutningsstøttesystem basert på maskinlæring kommer til et bestemt resultat.¹⁶¹ Kunstig intelligente systemer som anvendes av forvaltningen må ha

158 Jf. fortalepunkt 4.

159 Schartum, Jansen og Tranvik (2017) s. 109.

160 Schartum (2018) s. 25.

161 Datatilsynet (2018) s. 18.

funksjonalitet som gjør at de kan begrunne resultater,¹⁶² men det er uansett mulig at en saksbehandler tviler på egne forutsetninger til å kunne bedømme en sak bedre enn en prediksjonsmodell. I situasjoner med høyt arbeids- og tidspress, samt varierende kompetanse, øker sannsynligheten for at beslutningsstøttesystemet får stor betydning for utfallet av saken.

Et annet moment er at profilering i alminnelighet er ansett for å medføre risiko.¹⁶³ Profilering er ingen nøyaktig vitenskap. Statistiske korrelasjoner som gjelder for et flertall, kan ikke uten videre overføres på et individnivå. Det er også fare for at profilering kan lede til diskriminering eller forsterkning av sosiale ulikheter.¹⁶⁴ Faren knyttet til uriktige prediksjoner må ses på bakgrunn av at prediksjonen vil utgjøre en ny personopplysning om personen som det er vanskelig å verifisere.¹⁶⁵ Eksempelvis vil et individ som kategoriseres med en viss sannsynlighet for å bli ufør ha vanskeligheter med å bestride en slik personopplysning.

Ved vurdering av risiko må man også ta hensyn til virkeområdet for beslutningsstøtteverktøyet. Her vil det ha betydning om beslutningsstøtteverktøyet skal benyttes ved vurderingen av om et individ skal innvilges ytelser, om man skal underlegges kontroll eller om man skal markeres som sårbar og derfor ha et behov for oppfølging fra forvaltningen. Betydningen av prediksjonen for individet må vektlegges, herunder om prediksjonen får inngripende virkninger for personen. Dersom prediksjonen kan ha inngripende virkninger for individet, må risikoen anses som så høy at det må stilles svært strenge krav til kompensende tiltak som kan sikre at anvendelse av beslutningsstøtteverktøyet vil være forsvarlig.

Uavhengig av anvendelsesområdet for det konkrete beslutningsstøtteverktøyet, må det forutsettes at den potensielle risikoen knyttet til slike systemer er høy. Dette medfører at det må stilles strenge krav til et systems kvalitet og pålitelighet. Prinsippet om forsvarlig saksbehandling, herunder utredningsprinsippet, innebærer at det ikke kan aksepteres at slike system har kjente feil eller mangler. Dette leder til at datamaterialet må underlegges rettslige krav som sikrer at det endelige beslutningsstøttesystemet blir forsvarlig i den grad datamaterialet påvirker kvaliteten til systemet.

162 Jf. fvl. § 24, se også Datatilsynet (2021) og Lyngstad (2020).

163 Se avsnitt 2.2.1 ovenfor.

164 STOA (2020) s. 21.

165 Se avsnitt 2.3.

3.4.4 I hvilken utstrekning vil krav om kvalitet og minimering medføre begrensninger for bruk av data?

Spørsmålet i det følgende er i hvilken utstrekning de krav om kvalitet og minimering som følger av dataminimerings- og riktighetsprinsippet, samt prinsippet om forsvarlig saksbehandling begrenser forvaltningens bruk av data til utvikling av prediksjonsmodeller. Vurderingen av de rettslige bestemmelsene gjøres med utgangspunkt i at det stilles strenge kvalitetskrav til forvaltningens beslutningsstøttesystemer som en følge av forsvarlighetskravet. Avgjørende for de krav som følger av dataminimerings- og riktighetsprinsippet, samt forsvarlighetsprinsippet, er hvilken sammenheng det er mellom datamaterialet som anvendes og egenskapene til det endelige beslutningsstøttesystemet. Med andre ord må vurderingen se hen til hvilken betydning datamaterialet har for systemets forsvarlighet.

Egenskaper ved datamaterialet man anvender ved trening av maskinlæring kan ha betydning for prediksjonsmodellen på mange måter. En mulig tilnærming til vurderingen av rettslige krav som skal stilles til datamaterialet, er å skille mellom 1. betydning for systemets prediksjonsevner og 2. betydning for hvorvidt systemet vil ha skadelige egenskaper.

Betydningen av data for systemets prediksjonsevner refererer til hvordan de anvendte data har betydning for om systemet vil gjøre plausible prediksjoner som er nyttige for saksbehandler. Modellens prediksjonsevner forutsetter at treningsmaterialet inneholder *informasjon som kan gi modellen nødvendige inn-sikter*.¹⁶⁶ Dersom formålet er å utvikle en modell som angir sannsynlighet for at en arbeidsledig vil ha utfordringer med jobbsøking, vil trolig ikke informasjon om antall søsken og reisevaner gi modellen tilstrekkelig eller relevant kunnskap. På dette punktet vil dataminimeringsprinsippet krav til at opplysninger skal være relevante og adekvate forutsette at man benytter nødvendige opplysnings-typer, samt at man ikke benytter opplysninger man ikke kan begrunnes saklig i formålet.

Videre vil dataenes *kvalitet* ha betydning for modellens nøyaktighet.¹⁶⁷ Kvalitet kan referere til om datamaterialet er utdatert, kommer fra kilder som inneholder mange feil eller på andre måter har kvalitative mangler. I relasjon til dataenes kvalitet har særlig det personvernrettslige riktighetsprinsippet og det forvaltningsrettslige utredningsprinsippet betydning. Dataminimeringsprinsippet vil også spille inn ved at uriktige opplysninger kan medføre at kravet til adekvans og relevans ikke er oppfylt. Om ordlyden tas bokstavelig, stiller riktighetsprin-

¹⁶⁶ Datatilsynet (2018) s. 8.

¹⁶⁷ Teknologirådet (2018) s. 19.

sippet i art. 5(1)(d) krav om at opplysningene skal stemme overens med virkeligheten slik den er. Dette medfører at de opplysninger forvaltningen anvender i datamaterialet må være sannferdige. Ved maskinlæring benytter man ofte store mengder opplysninger, og sannferdigheten av den enkelte opplysning vil derfor ha liten betydning. I relasjon til maskinlæring vil riktighetsprinsippet altså ikke medføre et absolutt krav om at opplysningene må være korrekte. Det kan allikevel ikke aksepteres at treningsmaterialet har slike mangler at det går utover prediksjonsmodellen nøyaktighet.

For det tredje vil prediksjonsmodellens evne til å fange komplekse sammenhenger være påvirket av *mengden* treningsdata man benytter i utviklingen. Ofte vil tilgang til mer data bety at algoritmen kan integrere flere nyanser i modellen.¹⁶⁸ Dersom dette bidrar til mer treffsikre modeller, kan det gi insentiv til å bruke mer data i utviklingsfasen. Her vil dataminimeringsprinsippet ha betydning som rettslig skranke. Som angitt, vil prinsippet få betydning både for antall registrerte forvaltningen kan innhente data fra, hvor mange opplysningskategorier de kan anvende og for antallet opplysninger innenfor den enkelte kategori. Kravet om at opplysninger skal være relevante og begrenset til det som er nødvendig for formålet medfører at forvaltningen må søke å anvende så lite opplysninger som mulig. Et eksempel kan være at forvaltningen anvender opplysninger om inntekt fra 5 år i stedet fra 20 år dersom man uansett kan oppnå formålet med behandlingen. Utfordringen knyttet til maskinlæring er at det kan være vanskelig å vurdere hvilke opplysninger som vil være nødvendige for å utvikle en bestemt modell.¹⁶⁹ I en slik situasjon tilsier dataminimeringsprinsippet at forvaltningen på forhånd må skaffe kunnskap om området modellen skal fungere i og hvilke opplysninger som faktisk er nødvendige og relevante for å trene en slik modell.¹⁷⁰ Samtidig må forvaltningen etterstrebe å anvende den minste mengden opplysninger som er mulig, og heller utvide treningsmaterialet trinnvis til modellen oppnår kvaliteter man ønsker.¹⁷¹ Forvaltningen må også anvende syntetiske testdata¹⁷² hvis dette er et alternativ til personopplysninger, samt anonymisere eller pseudonymisere opplysninger om det lar seg gjøre.¹⁷³

Krav om dataminimering vil ikke gjelde i den grad minimering av opplysninger vil kunne påvirke prediksjonsmodellens forsvarlighet. Forvaltningen må anvende de opplysningene som er nødvendig for å utvikle en forsvarlig prediksjonsmodell, selv om man benytter store mengder personopplysninger fra

168 Teknologirådet (2018) s. 18.

169 Se avsnitt 2.1.

170 Datatilsynet (2018) s. 17.

171 Motzfeldt og Abkenar (red.) (2019) s. 143.

172 Ringnes (2019).

173 Motzfeldt og Abkenar (red.) (2019) s. 141-142.

mange registrerte. Proporsjonalitetskravet i dataminimeringsprinsippet markerer allikevel en yttergrense for hva som kan aksepteres av inngrep overfor den registrerte. Dersom utviklingsprosjektet er knyttet til et beslutningsstøtteverktøy som i begrenset grad vil lede til positive virkninger for forvaltningen eller borgeren, kan det ikke aksepteres at man behandler opplysninger av en art eller i et slikt omfang at det medfører et stort inngrep overfor den registrerte. Eksempelvis vil det neppe være proporsjonalitet mellom anvendelse av et stort antall personopplysninger for å utvikle en modell som kan predikere når det vil være kø på et NAV-kontor.

Som angitt ovenfor vil datamaterialet også ha betydning for om prediksjonsmodellen vil ha uønskede og skadelige egenskaper. Fordi modellen vil reflektere dataene som er benyttet, vil forskjeller i samfunnet, ekskludering eller andre uttrykk for diskriminering som gjenspeiles i treningsgrunnlaget også gjenspeiles i prediksjonsmodellen.¹⁷⁴ En mulig fare er at man benytter tidligere saksbehandlingspraksis som treningsmaterialet for algoritmen. Selv om det er et krav om at forvaltningens avgjørelser skal baseres på lovlige kriterier i samsvar med myndighetsmisbrukslæren, er det allikevel en fare for at menneskelige vurderinger bevisst eller ubevisst er påvirket av holdninger som kan lede til sympatier eller antipatier. Hvis en gruppe i samfunnet historisk er blitt behandlet fordelaktig eller ufordelaktig av forvaltningen, vil denne ulikheten kunne bli reflektert og reproduisert i en prediksjonsmodell. Dersom det er en fare for at et utvalg tidligere saksbehandlingspraksis kan reflektere usaklige eller upartiske vurderinger, tilsier både riktighets- og dataminimeringsprinsippet, samt utredningsprinsippet at opplysningene ikke benyttes ved utviklingen.

For å unngå urettferdige eller diskriminerende prediksjonsmodeller er det også viktig at treningsgrunnlaget speiler befolkningen som helhet.¹⁷⁵ Om en gruppe i samfunnet ikke er representert i treningsdataene vil erfaringer knyttet til deres opplysninger ikke være med å forme modellen. Resultatet kan være uriktige prediksjoner for den gruppen som ikke er representert. Modellens presisjon vil også bli påvirket hvis én gruppe er overrepresentert i treningsdataene.¹⁷⁶ Ved utvikling av en prediksjonsmodell tilsier særlig kravet til adekvans at det skal sørges for tilstrekkelig representasjon innenfor de ulike opplysningstypene man benytter ved treningen, det være seg kjønn, aldersgrupper, yrkesgrupper osv.

174 Teknologirådet (2018) s. 48.

175 STOA (2020) s. 21.

176 STOA (2020) s. 21.

3.4.5 Oppsummering

Dataminimering må hensyntas når man skal benytte opplysninger til å utvikle prediksjonsmodeller. I den utstrekning det er en fare for at treningsmaterialet inneholder opplysninger som reflekterer sosial ulikhet, tilsier krav om adekvans og relevans at forvaltningen begrenser bruken av disse opplysningene. Dataminimeringsprinsippet tilsier også at forvaltningen i alminnelighet skal søke å begrense omfanget av opplysninger så mye som mulig. Dette kan man sikre gjennom tiltak som trinnvis utvikling av modellen, syntetiske testdata, anonymisering eller pseudonymisering.

Dataminimering kan allikevel ikke gis gjennomslag på en slik måte at det går utover forsvarligheten til systemet. Forsvarlighetskravet tilsier at kvalitet og tilstrekkelighet bør vektlegges på bekostning av at behandlingen av personopplysninger begrenses så mye som mulig. Forsvarlighet forutsetter at datamaterialet er tilstrekkelig for å kunne gi prediksjonsmodellen nødvendige innsikter. Man må også sikre at modellen ikke blir upresis for enkelte grupper som en følge av manglende representativitet. Krav til tilstrekkelighet vil derfor gå foran minimering i den utstrekning dette har betydning for systemets forsvarlighet. Proporsjonalitetskravet vil allikevel utgjøre en yttergrense. Det kan ikke aksepteres at det behandles en uforholdsmessig mengde personopplysninger sett opp mot formålet med beslutningsstøtteverktøyet.

3.5 Samlet drøftelse

Fremstillingen viser at rettslige reguleringer i forvaltningsretten og personvernretten har avgjørende betydning for forvaltningens bruk av opplysninger til trening av maskinlæring. De forvaltningsrettslige og personvernrettslige reguleringene vil spille inn på systemutviklingen samtidig. I det følgende forsøker jeg å belyse hvordan det skjer et samspill mellom personvernretten og forvaltningsretten når begrensninger knyttet til bruk av opplysninger skal klarlegges.

Formålsbegrensningsprinsippet og taushetspliktsreglene kan betegnes som materielle skranke med betydning for om opplysningene overhodet skal kunne anvendes. Reglene har likhetstrekk ved at de har som utgangspunkt at opplysninger lovlig kan benyttes til det formål de er samlet inn for. Taushetsplikten vil ikke gjelde overfor en saksbehandler som benytter opplysningene i forbindelse med enkeltsaksbehandlingen. Tilsvarende vil formålsbegrensningsprinsippet bare aktualiseres i den utstrekning den behandlingsansvarlige benytter opplysninger til nye formål. Reglene har allikevel ulik utforming og ulike begrunnelser. Taushetsplikten gjelder egenskaper ved opplysningene i seg selv, mens for-

målsbegrensningsprinsippet er knyttet til formålene opplysningene er samlet inn for. Dette medfører at opplysningene lovlig vil kunne behandles som treningsdata som en forenelig viderebehandling, men allikevel ikke kunne benyttes dersom de er underlagt taushetsplikt. Reglene bygger også på ulike hensyn. Et utslag av dette viser seg i anvendelsesområdet for bestemmelsene. Taushetspliktsreglene har som hovedformål å beskytte interessene til den som må gi opplysninger til forvaltningen, og er knyttet til situasjonen der «andre» kan få adgang eller kjennskap til opplysningen. Formålsbegrensningsprinsippet skal sikre forutberegnelighet og kontroll for den registrerte i forbindelse med behandling av personopplysninger, og er knyttet til viderebehandlingssituasjonen. Reglene reflekterer at forvaltningsretten og personvernretten bygger på ulike hensyn. Forvaltningsretten omhandler møtet mellom individ og forvaltningsorganer, og har således ivaretagelse av rettssikkerhet og beskyttelse av individet som fokus.¹⁷⁷ Personvernretten skal ivareta den registrertes rettigheter og friheter i forbindelse med behandling av personopplysninger, og inneholder derfor regler som skal sikre forutberegnelighet og autonomi.¹⁷⁸

Et annet viktig skille mellom formålsbegrensningsprinsippet og taushetsplikreglene er at formålsbegrensningsprinsippet mister betydning som materiell skranke i det behandlingsformålet angis ved innsamlingen av opplysningene. Dersom særlovgivningen uttrykker at opplysninger kan benyttes til å utvikle maskinlæring i medhold av art. 6(1)(e), jf. art 6(3), vil viderebehandlingsforbudet ikke aktualiseres for opplysninger som samles inn. Det er trolig at særlovgivningen i økende grad vil inneholde slike bestemmelser etter hvert som maskinlæring blir en viktigere del av forvaltningens virksomhet. A-opplysningsloven er et eksempel i det § 6a (1) angir at «skattemyndighetene kan behandle innhentede personopplysninger for å utvikle og teste IT-systemer». En slik lovfesting av formål vil allikevel ikke gi unntak fra taushetsplikreglene «med mindre bestemmelsen klart fastsetter eller forutsetter at taushetsplikten ikke skal gjelde», jf. fvl. § 13 f (2). Forvaltningslovutvalget har foreslått nye taushetsplikregler i NOU 2019:5, men forslaget inneholder ikke bestemmelser som klart forutsetter at taushetspliktige opplysninger skal kunne benyttes til utvikling av IT-systemer.¹⁷⁹ Det må derfor gjøres særskilt lovvedtak for at taushetsplikten ikke skal gjelde ved trening av maskinlæring.

De resterende rettslige reguleringene som er behandlet i rapporten stiller krav til kvaliteter ved treningsmaterialet som benyttes ved utvikling av prediksjonsmodeller. Avgjørende ved anvendelse av de rettslige reguleringene er hvilke krav

177 Graver (2019) s. 299.

178 Skullerud (2019) s. 35.

179 NOU 2019:5 s. 26-28.

som stilles for at forvaltningens beslutningsstøttesystemer skal være forsvarlige. Vurderingen av forsvarlighet må ta utgangspunkt i den risikoen som følger med utvikling og bruk av modellen. Dette beror særlig på det konkrete anvendelsesområdet for prediksjonsmodellen, men risikoen assosiert med profilering og beslutningsstøtte generelt tilsier at det må stilles strenge krav til systemet for at det skal være forsvarlig å benytte det ved saksbehandling. Vektleggingen av ulovfestede rettsprinsipper illustrerer den manglende forvaltningsrettslige lovreguleringen av systemutvikling. Når det ikke kan utledes krav til datamaterialet med utgangspunkt i lov, får forsvarlighetsprinsippet betydning som indirekte krav til treningsmaterialets kvalitet, samt som tolkningsmoment ved anvendelse av personvernprinsippene.

4 Rettspolitiske vurderinger

Vektleggingen av forsvarlighetsprinsippet i denne rapporten gir opphav til spørsmål om det burde lovfestes krav til utvikling av beslutningsstøtteverktøy som skal brukes i forvaltningens saksbehandling. Forvaltningslovutvalget har avvist at det bør lovfestes generelle systemkrav til innretningen av forvaltningsvirksomhet i forvaltningsloven. I stedet vises det til at det kan lovfestes minstekrav i særlovgivningen i den utstrekning det er behov for det. Det fremheves videre at personvernforordningen vil sette visse krav til innretningen av IKT-systemer.¹⁸⁰

Det kan stilles spørsmål ved i hvilken utstrekning personvernforordningen er egnet som rettsikkerhetsgaranti. Rettsikkerhet vil være et hensyn som skal vektlegges ved anvendelse av personvernforordningens bestemmelser,¹⁸¹ og man vil derfor kunne sikre forsvarlighet ved å anvende forsvarlighetsprinsippet som tolkningsmoment. Jeg anser det allikevel som problematisk å benytte et ulovfestet prinsipp som garanti for sentrale rettsikkerhetshensyn. Som påpekt i avsnitt 1.3 er innholdet i forsvarlighetsprinsippet uklart. Dette medfører at det vil være utfordrende for de som skal utvikle prediksjonsmodellen å orientere seg om hvilke rettslige krav modellen må imøtekomme. I en slik situasjon vil IT-prosjekter i forvaltningen måtte sikres tilstrekkelig juridisk kompetanse for å kunne harmonisere personvernrett og forvaltningsrett i tråd med forsvarlighetsprinsippet.

Som en følge av de særlige rettsikkerhetshensynene som aktualiseres ved utvikling av ML-baserte beslutningsstøtteverktøy, mener jeg at det bør lovfestes minstekrav til utvikling av slike systemer. Dette gjelder særlig dersom verktøyet skal benyttes ved myndighetsutøvelse. Reguleringen bør konkretisere de deler av forsvarlighetsprinsippet som har betydning for systemutvikling slik at forsvarlighet ikke vil bero på skjønsmessige vurderinger alene. Eksempler på slike reguleringer kan være krav til at datamaterialet skal sikre representativitet for ulike grupper i samfunnet, krav til at personopplysninger ikke skal benyttes med mindre det er nødvendig og regler om testing og kvalitetssikring av systemet. Reglene bør harmoniseres med personvernforordningen slik at anvendelsen av reguleringene blir så enkel som mulig. Etter Forvaltningslovutvalgets

180 NOU 2019:5 s. 159.

181 Se avsnitt 3.4.3.

forslag vil slike reguleringer måtte tas inn i særlovgivningen til det enkelte forvaltningsorgan.

Rapporten har videre avdekket at rettslige reguleringer spiller en viktig rolle når systemer basert på maskinlæring skal utvikles og tas i bruk. Reguleringene som er behandlet i rapporten har potensiale til å begrense forvaltningens mulighet for å utvikle systemer basert på kunstig intelligens. Når reglene ikke er utformet med tanke på utvikling og bruk av kunstig intelligens, er det mulig at deler av de rettslige begrensningene ikke er tilsiktet. Europakommisjonen har begynt arbeidet med å utvikle et rettslig rammeverk for kunstig intelligens innenfor EU, og presenterte 21. april 2021 et forordningsforslag for regulering av kunstig intelligens.¹⁸² Forslaget viser at Kommisjonen har valgt en risikobasert tilnærming, der de rettslige reguleringene vil tilpasses den risiko som assosieres med anvendelsen av den konkrete KI-teknologien.¹⁸³ Som en følge av at EU har påbegynt arbeidet med å regulere kunstig intelligens, vil det trolig bli enklere å anvende regler i eksisterende regelverk. Slik Kommisjonen fremhever i forslaget, er det viktig å finne en balanse mellom reguleringer som adresserer risikoen som følger med utvikling og bruk av kunstig intelligens og faren for å unødvendig hindre og begrense teknologisk utvikling.¹⁸⁴ Hvordan EU håndterer denne balansegangen vil ha betydning for Europas rolle i det teknologiske kappløpet. Det vil også ha betydning for hvorvidt utvikling og bruk av kunstig intelligens vil skje på bekostning av rettssikkerhet.

182 Europakommisjonen (2021) s. 1.

183 Europakommisjonen (2021) s. 3.

184 Europakommisjonen (2021) s. 3.

Litteraturliste

Internasjonale rettskilder

Konvensjoner

Personvernforordningen 2016/679 EUROPAPARLAMENTS- OG RÅDSFORORDNING (EU) 2016/679 av 27. april 2016 om vern av fysiske personer i forbindelse med behandling av personopplysninger og om fri utveksling av slike opplysninger samt om oppheving av direktiv 95/46/EF (generell personvernforordning) [GDPR]

Artikkel 29-gruppen

A29-gruppen (2018) Article 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (2018). [Tilgjengelig på: <https://ec.europa.eu/newsroom/article29/items/612053>] hentet: 13.05.2021

A29-gruppen (2013) Article 29 Data Protection Working Party, *Opinion 03/2013 on purpose limitation*, (2013) [Tilgjengelig fra: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf] hentet: 13.05.2021

EU-kommisjonen

COM (2021) *Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*

Norske rettskilder

Lover

1967 Lov 10. februar 1967 Lov om behandlingsmåten i forvaltningssaker (forvaltningsloven)

2012 Lov 22. juni 2012 nr. 43 Lov om arbeidsgivers innrapportering av ansettelses- og inntektsforhold m. m (a-opplysningsloven)

- 2017 Lov 16 juni 2017 nr. 51 Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven)
- 2018 Lov 15. juni 2018 nr. 38 Lov om behandling av personopplysninger (personopplysningsloven)

Forarbeider

- Ot.prp.nr.3 (1976–1977) Om lov om endringer i lov 10 februar 1967 om behandlingsmåten i forvaltningssaker (regler om taushetsplikt m. m)
- Prop.56 LS (2017–2018) Lov om behandling av personopplysninger (personopplysningsloven) og samtykke til deltakelse i en beslutning i EØS-komiteen om innlemmelse av forordning (EU) nr. 2016/679 (generell personvernforordning) i EØS-avtalen
- NUT 1958:3 Innstilling fra Komiteen til å utrede spørsmålet om mer betryggende former for den offentlige forvaltning.
- NOU 2019:5 Ny forvaltningslov – Lov om saksbehandlingen i offentlig forvaltning (forvaltningsloven)

Litteratur

- Aarli og Mæhle (2018) Aarli, Ragna og Synne Sæther Mæhle. *Juridisk metode i et nøtteskall*. 1. utg., Oslo: Gyldendal, 2018.
- Bendiksen og Norman Hansen (2019) Bendiksen, Christian og Eirik Norman Hansen. *Når juss møtes AI: rettslig regulering av kunstig intelligens*. 1. utg., Oslo: Gyldendal, 2019.
- Boe (2018) Boe, Erik Magnus. *Forsvarlig forvaltning*. Oslo: Universitetsforlaget, 2018.
- Christensen og Berg (2019) Christensen, Johan og Berg, Ole T. «Velferdsstat.» i Store norske leksikon på snl.no. (04.09.2019) <https://snl.no/velferdsstat> hentet 05.03.2021
- Datatilsynet (2018) Datatilsynet. *Kunstig intelligens og personvern*. Rapport, januar 2018. Oslo: 2018. <https://www.datatilsynet.no/globalassets/global/dokumenter-pdf/rettigheter-og-plikter/rapporter/rapport-om-ki-og-personvern.pdf>

- Datatilsynet (2021) Datatilsynet. «NAV» (08.03.2021) <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/pagaende-prosjekter2/nav/> hentet 13.05.2021
- de Terwangne (2020) de Terwangne, Cécile de. «Article 5. Principles relating to the processing of data» I *The EU General Data Protection Regulation (GDPR) A Commentary*. Christopher Kuner, Lee A. Bygrave og Christopher Docksey (red.). Oxford University Press UK – OSO, 2020, 309-320.
- Eckhoff og Smith (2018) Eckhoff, Torstein og Eivind Smith. *Forvaltningsrett*. 11. utgave. Oslo: Universitetsforlaget, 2018.
- Kommunal og moderniseringsdepartementet (2020) *Nasjonal strategi for kunstig intelligens*. (14.01.2020) <https://www.regjeringen.no/contentassets/1febbbb2c4fd4b-7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf>
- Graver (2019) Graver, Hans Petter. *Alminnelig forvaltningsrett*. 5. utgave. Oslo: Universitetsforlaget, 2019.
- Henriksen (2019) Henriksen, Ragne. «Kan maskinlæring løse alle problemer?» (23.94.2019) <https://www.visma.no/blogg/maskinlaering-lose-problemer/> hentet 13.05.2021
- Fredriksen og Mathisen (2019) Fredriksen, Halvard Haukeland og Gjermund Mathisen. «EU-rett som norsk rettskilde.» I *Juridisk metode og tenkemåte*. Alf Petter Høgberg og Jørn Øyrehagen Sunde (red.). 1. utgave, Oslo: Universitetsforlaget, 2019, 386-480.
- ICO (2017) ICO (Information commissioner's Office). *The Big Data, AI, and Machine Learning report*. (publisert i 2014, oppdatert i 2017) <https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>
- Innjord (2014) Innjord, Frode A. «Forvaltningsretten.» I *Knophs oversikt over Norges rett*. Ragnar Knoph red., 14. utg., Oslo: Universitetsforlaget., 2014, s. 572-592.
- Jarbekk (2019) Jarbekk, Eva (red.). *Personopplysningsloven og personvernforordningen (GDPR) med kommentarer*. 1. utgave, Oslo: Gyldendal, 2019.

- Kotschy (2020) Kotschy, Waltraut. «Article 6. Lawfulness of processing» I *The EU General Data Protection Regulation (GDPR) A Commentary*. Christopher Kuner, Lee A. Bygrave og Christopher Docksey (red.). Oxford University Press UK – OSO, 2020, 321-344.
- Kotschy (2021) Kotschy, Waltraut. «Article 6. Lawfulness of processing» I *The EU General Data Protection Regulation (GDPR): A Commentary – 2021 Update*. Christopher Kuner, Lee A. Bygrave og Christopher Docksey (red.). Oxford University Press, 2021, s. 71-76. https://fdslive.oup.com/www.oup.com/academic/pdf/law/GDPRCommentary_ArticleUpdates.pdf
- Lyngstad (2020) Lyngstad, Cathrine Pihl. «Forklarbar AI i praksis.». *Nettseminar om forklarbar kunstig intelligens*. [Powerpoint presentasjon] Tekna Big Data og Tekna Forskerne. (15.05.2021) s. 4 <https://www.tekna.no/contentassets/4df948a6c2b64e-03808355f13acdac3f3-cathrine-pihl-lyngstad-forklarbar-ai-i-praksis.pdf>
- Løland, Berset og Haff (2017) Løland, Anders, Anders Berset og Ingrid Hobæk Haff. «Er maskinlæring framtida i Skatteetaten?» *Praktisk økonomi & finans* 03/2017 (Volum 33) Side: 344-352. DOI: <https://doi-org.ezproxy.uio.no/10.18261/issn.1504-2871-2017-03-06>
- Motzfeldt og Abkenar (Red.) (2019) Motzfeldt, Hanne Marie og Azad Taheri Abkenar (red.). *Digital Forvaltning – utvikling af sagsbehandlende løsninger*. 1. utg., København: Djøf Forlag, 2019.
- Olsen (2020) Olsen, Birgitte Kofod. *Håndbog i dataansvarliged*. 1. utgave, Djøf forlag, 2020.
- Ringnes (2019) Ringnes, Ida F. «Kunstig intelligens kan hjelpe NAV med å gi bedre tjenester.». (18.03.2019) <https://memu.no/artikler/kunstig-intelligens-kan-hjelpe-nav-med-a-gi-bedre-tjenester/> hentet 14.05.2021
- Schartum (2015) Schartum, Dag Wiese. *Den elektronisk forvaltning og loven*. Oslo: Kommuneforlaget, 2015.
- Schartum (2018) Schartum, Dag Wiese. *Digitalisering av offentlig forvaltning – fra lovtekst til programkode*. Bergen: Fagbokforlaget, 2018.
- Schartum (2020) Schartum, Dag Wiese. *Personvernforordningen – en lærebok*. 1. utg., Fagbokforlaget 2020.

- Schartum, Jansen og Tranvik (2017) Schartum, Dag Wiese, Arild Jansen og Tommy Tranvik. *Digital forvaltning. En innføring*. Bergen: Fagbokforlaget, 2017.
- Sejersted (2011) Sejersted, Fredrik mfl. *EØS-rett*. 3.utgave. Oslo: Universitetsforlaget, 2011
- Skullerud (2019) Skullerud, Åste Marie Bergseng mfl. *Personopplysningsloven og personvernforordningen (GDPR) Kommentartutgave*. Oslo: Universitetsforlaget, 2019.
- Stalla-Bourdillon og Knight (2018) Stalla-Bourdillon, Sophie og Alison Knight. «Data Analytics and the GDPR: Friends or Foes? A Call for a Dynamic Approach to Data Protection Law» (September 13, 2018). I *Data Protection and Privacy: The Internet of Bodies*. R. Leenes mfl. Red., Oxford: Hart publishing, 2018. Hentet fra SSRN: <https://ssrn.com/abstract=3248976>
- STOA (2020) STOA – The Panel for the Future of Science and Technology ved Giovanni Sartor og Franscesca Lagioia. «The impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence.» PE 641.530 – June 2020. Brussel: 2020. DOI: 10.2861/293
- Teknologirådet (2018) Teknologirådet. *Kunstig intelligens – muligheter, utfordringer og en plan for Norge*. Oslo: September 2018. <https://teknologiradet.no/wp-content/uploads/sites/105/2018/09/Rapport-Kunstig-intelligens-og-maskinlaering-til-nett.pdf>
- Valevatn (2017) Valevatn, Joakim. «Når data utpeker svindlere og hjelper skoleleie og arbeidsledige». (30.01.2017) <https://teknologiradet.no/nar-data-utpeker-svindlere-og-hjelper-skoleleie-og-arbeidsledige/> hentet 18.05.2021.
- Woxholt (2011) *Forvaltningsloven med kommentarer*. 5. utgave. Oslo: Gyldendal, 2011.
- Öman (2019) Öman, Sören. *Dataskyddsförordningen (GDPR) m.m En kommentar*. Stockholm: Gula biblioteket, 2019.