

# CompLex



Senter for rettsinformatikk / Avdeling for forvaltningsinformatikk

Linn Høgåsen

## Prediksjoner av sykefraværslengde i NAV

En tverrfaglig case-studie om bruk av maskinlæring til å profilere

5/2021



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

Grunnlaget for denne rapporten er masteroppgaven min i forvaltningsinforma-  
tikk, som ble levert i juni 2021. Det er gjort minimale justeringer i innholdet  
siden levering av oppgaven: Kapittel 1.4 om undersøkelsesopplegget er kortet  
ned noe, og jeg har ellers rettet opp i enkelte, mindre skrivefeil.

Jeg vil gjerne takke veilederen min, Dag Wiese Schartum, som har rettet opp-  
merksomheten min mot en rekke interessante poenger i tilbakemeldinger på  
diverse utkast, og ellers gitt konstruktiv og god feedback. Jeg ønsker også å rette  
en takk til informantene mine i NAV, for spennende samtaler, raske svar på  
e-post, og tiden deres. I tillegg må jeg takke lesesalgjengen på Domus Acade-  
mica, My, og Andreas, som alle har bidratt med interessante innspill og motiva-  
sjon på sine måter.

Oslo, 28. november 2021

Linn Høgåsen

# Contents

Forord .....	3
1 Sammen drag .....	6
2 Innledning .....	7
2.1 Bakgrunn og aktualitet .....	7
2.2 Problemstillingene: Et rettslig, teknologisk og organisatorisk innblikk i profilering .....	8
2.2.1 Overordnet problemstilling og forholdet mellom delproblemstillingene .....	8
2.2.2 Problemstilling 1: Hvordan har NAV organisert arbeidet med å utvikle et profileringssystem? .....	9
2.2.3 Problemstilling 2: Hva er den tekniske innretningen til profileringssystemet? .....	10
2.2.4 Problemstilling 3: Hvordan har rettsprinsipper og rettsregler, herunder eventuelle rettslige usikkerheter, påvirket systeminnretningen? .....	11
2.3 Forskningsstatus .....	11
2.4 Undersøkelsesopplegget .....	13
2.4.1 Det overordnede undersøkelsesopplegget: Case studie .....	13
2.4.2 Dokument- og litteraturstudier .....	14
2.4.3 Intervjuer med informanter .....	15
2.4.4 Juridisk metode .....	16
2.4.5 Metodetriangulering .....	17
2.5 Oversikt over den videre framstillingen .....	17
3 Profilering: Teori og utfordringer .....	19
3.1 Teoretisk forklaring av profilering .....	19
3.2 Utfordringer knyttet til bruk av profilering i forvaltningen .....	21
4 Kontekst og rammer for utviklingsarbeidet i caset .....	24
4.1 Saksbehandlingskonteksten .....	24
4.2 Beslutningsstøttesystemet i kontekst .....	26
4.3 Rammer og føringer for systemutviklingsarbeidet .....	28
4.3.1 Forvaltningspolitiske og virksomhetsspesifikke føringer ...	28
4.3.2 Rettslig rammer .....	30

5	Arbeidet med prediksjon av sykefraværslengde .....	34
5.1	Fasene i utviklingsarbeidet .....	34
5.2	Organisering av utviklingsarbeidet .....	36
5.2.1	Teamet .....	37
5.2.2	Forholdet mellom teamet og andre organisatoriske enheter. ....	38
5.2.3	Dokumentasjon av teknisk og rettslig innhold .....	45
5.3	Foreløpig teknisk innretning av profileringsystemet .....	46
5.3.1	Dataene. ....	47
5.3.2	Analyse av dataene .....	52
5.4	Rettslige vurderinger som påvirker utviklingsarbeidet .....	54
5.4.1	Lovligheten til behandlingen .....	55
5.4.2	Krav til forklarbarhet og rettferdighet .....	60
5.4.3	Krav til dataminimering .....	62
5.4.4	Dokumentasjon .....	63
5.4.5	Rettslige uklarheter .....	66
5.5	Samlet drøftelse .....	67
6	Avsluttende bemerkninger .....	71
	Kilder .....	72

## Figurliste

Figur 1:	Sammenhengen mellom problemstillingene.....	9
Figur 2:	Profil og egenskaper .....	19
Figur 3:	Profilering basert på ekspertkunnskap/maskinlæring .....	21
Figur 4:	Beslutningsstøttesystem i saksbehandlingskonteksten .....	26
Figur 5:	Rammer og føringer for utviklingsarbeidet.....	28
Figur 6:	Fasene i utviklingsarbeidet.....	34
Figur 7:	Oversikt over organisatoriske enheter tilknyttet utviklingsarbeidet .....	39
Figur 8:	Teamet og ledelsen.....	40
Figur 9:	Dataflyten .....	47

# 1 Sammendrag

Denne rapporten dreier seg om NAVs arbeid med å utvikle et beslutningsstøttesystem som baserer seg på maskinlæring. Systemet skal brukes i forbindelse med vurderingen av behovet for dialogmøte 2 i sykefraværsoppfølgingen. Utviklingsarbeidet jeg har undersøkt er ikke ferdigstilt, og rapporten bør derfor leses som en utforskning og beskrivelse av en utviklingsprosess, ikke av et system i bruk.

Forskningsspørsmålene i studien er eksplorative, og forskningsdesignet er en case-studie. For å undersøke utviklingsarbeidet, tar jeg utgangspunkt i ideer fra *critical data studies*. Jeg har dermed undersøkt hvordan NAV jobber med å tilrettelegge og analysere datasett. I tillegg har jeg trukket linjer til hvordan organisering og rettslige vurderinger påvirker dette arbeidet. Et hovedpoeng har vært å synliggjøre utfordringer NAV møter på i utviklingsarbeidet.

Utviklingsteamet har enn så lenge brukt mye tid på å kartlegge og klargjøre datagrunnlaget de kan bruke, og mindre på å utvikle algoritmen som skal analysere datagrunnlaget. Organisatorisk er det primært teknologer, data scientister, og veiledere i NAV som påvirker systeminnretningen. Juristene i NAV har lite tid. Dette påvirker både omfanget av samarbeidet mellom teamet og jurister, men også de mange og prinsipielle vurderingene juristene må ta i et slikt arbeid. Ellers er det enkelte eksterne aktører som påvirker arbeidet, primært innenfor forklarbarhet.

Regelverket som gjelder for bruk av maskinlæring i saksbehandlingen til offentlig sektor er svært uklart. Det er stor usikkerhet rundt tolkning av det gjeldende regelverket, blant annet for forklarbarhet, rettferdighet, dataminimering, og hjemmelsgrunnlaget. Det er utfordrende for enkeltjurister i NAV å skulle ta slike omfattende og prinsipielle vurderinger. Jeg konkluderer derfor med å peke på behovet for at lovgiver tar tydelig stilling til disse spørsmålene, istedenfor å overlate det til rettsanvenderne.

## 2 Innledning

### 2.1 Bakgrunn og aktualitet

Kunstig intelligens og maskinlæring er på dagsorden i mange land og sektorer. Teknologiene skal bidra til å revolusjonere en rekke tjenester, både privat og offentlig, og regnes for å være store konkurransefortrinn i tiden fremover. Kina, USA, og andre land med svakere personvernlovgivning kan potensielt utkonkurrere de europeiske landene i kompetanse og bruk av kunstig intelligens. EU har derfor laget en felles strategi for de europeiske landene, som skal bidra til et særeuropeisk fokus på kunstig intelligens med menneskerettigheter i sentrum.<sup>1</sup>

I Norge har vi også stor tiltro til potensialet i kunstig intelligens og maskinlæring, og den nasjonale strategien for kunstig intelligens trekker fram høy kompetanse gjennom utvikling og bruk av kunstig intelligens som et hovedmål.<sup>2</sup> Det forventes også at offentlig forvaltning kan utnytte maskinlæring til å løse sine oppgaver både bedre og på helt nye måter.<sup>3</sup> Samtidig skal den kunstige intelligensen utvikles på en etisk måte, slik at utvikling og bruk ikke går på bekostning av den høye tilliten forvaltningen nyter i Norge.<sup>4</sup>

Parallelt med disse oppfordringene til bruk av kunstig intelligens, er det en rekke organisasjoner og personer som peker på utpregede risikoer ved bruk av maskinlæring. AlgorithmWatch publiserte for eksempel en rapport i 2020 om bruk av kunstig intelligens i Europa, der de særlig har trukket fram hvordan automatisering utfordrer menneskerettigheter og andre sentrale prinsipper i rettsstater.<sup>5</sup> FNs rapportør på fattigdom har også publisert en rapport om den digitale velferdsstaten, der han retter søkelyset mot digitalisering i velferdssektoren.<sup>6</sup> Han påpeker at det kan være fristende for mange stater å digitalisere for å spare inn på de store budsjettene som gjerne fordeles til denne sektoren.<sup>7</sup> Dermed testes ny teknologi i en sektor der tjenestene er rettet mot mennesker i særlig utsatte situasjoner, som derfor ofte har mindre ressurser og evner til å ivareta egne rettigheter.

---

1 Europakommisjonen (2020).

2 Kommunal- og moderniseringsdepartementet (2020) s. 33.

3 Ibid. s. 5.

4 Ibid. s. 56.

5 AlgorithmWatch (2020).

6 FN (2019) s. 4.

7 Ibid. s. 21.

Disse utfordringene bør være tankevekkere til norske myndigheter, som oppfordrer til bruk av kunstig intelligens. Dersom myndighetene skal kunne kontrollere bruken av kunstig intelligens, må det i hvert fall delvis skjer gjennom lovgivning. I tiden fremover blir det derfor helt sentralt å kunne regulere på en måte som legger til rette for bruk av kunstig intelligens, samtidig som rettighetene til enkeltmennesker ivaretas. Det forutsetter kunnskap om de problemstillingene som oppstår ved utvikling og bruk av kunstig intelligens, for å sikre at lovene i samfunnet adresserer de faktiske dilemmaene som oppstår.<sup>8</sup>

Denne rapporten bidrar til å kaste lys over de problemstillingene offentlige virksomheter kan møte ved utvikling av maskinlæring som skal brukes i saksbehandling. Helt konkret ser jeg på utvikling av maskinlæringsalgoritmer som grunnlag for profilering av sykmeldte i velferds- og arbeidsetaten (heretter NAV).

## **2.2 Problemstillingene: Et rettslig, teknologisk og organisatorisk innblikk i profilering**

### **2.2.1 Overordnet problemstilling og forholdet mellom delproblemstillingene**

Det overordnede spørsmålet i denne rapporten er **hvordan profilering basert på maskinlæring kan forankres i rettsregler**. Profilering kan forstås både som *bruk* av profileringsteknikker, men også *utvikling* av profileringsteknikker. Ettersom profilering basert på maskinlæring i begrenset grad brukes, har jeg ingen målsetting om å se nærmere på hvordan profilene brukes innenfor virksomheten i praksis. Når jeg i det følgende bruker begrepet «profilering», begrenser meg med andre ord til utviklingsarbeidet i forkant av praktisk anvendelse av profilene.<sup>9</sup> Når jeg henviser til et «profileringssystem», mener jeg en teknisk innretning som inkluderer bruk av profileringsteknikker.

For å finne svar på det overordnede spørsmålet mitt har jeg tre delproblemstillinger, som alle er eksplorative. I problemstilling 1 undersøker jeg **hvordan NAV organiserer arbeidet med å utvikle et profileringssystem**. Problemstilling 2 dreier seg om **den tekniske innretningen av profileringssystemet**. I problemstilling 3 undersøker jeg **hvordan rettsregler og rettsprinsipper, herunder eventuelle rettslige usikkerheter, påvirker profileringssystemet**. Mesteparten av rapporten er viet til å kartlegge hvordan datadrevet profilering kan gjennomføres i NAV, med vekt på problemstillinger NAV står overfor i et slikt arbeid, enten de skyldes organisatoriske, tekniske eller rettslige forhold.

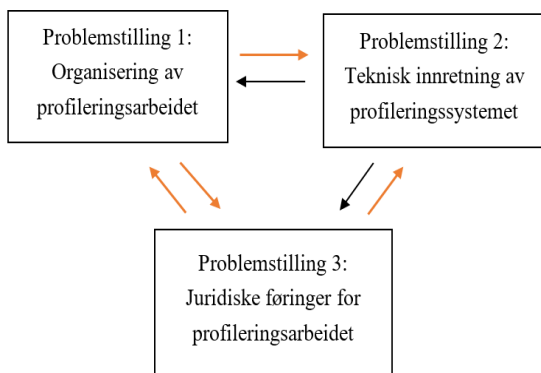
---

<sup>8</sup> Askland (2011) s. xiv.

<sup>9</sup> Se imidlertid avsnitt 2.2, der jeg nevner muligheten for dynamiske modeller. Da blir dette skillet mindre tydelig.



I figur 1 (nedenfor) har jeg laget en oversikt over hvordan de tre problemstillingene relaterer til hverandre. Det er piler som peker i begge retninger mellom alle boksene, for å illustrere at det i teorien er mulig å se på gjensidige påvirkningsforhold mellom alle de tre problemstillingene. Jeg har imidlertid begrenset meg til å se på de forholdene som er uthevet i oransje.



Figur 1: Sammenhengen mellom problemstillingene

Jeg har laget oransje piler fra problemstilling 1 og problemstilling 3 til problemstilling 2, fordi jeg forutsetter at rettslige og organisatoriske funn påvirker den tekniske innretningen til systemet, indirekte eller direkte. I tillegg er det oransje piler som peker begge veier mellom problemstilling 1 og problemstilling 3, fordi jeg peker på sammenhenger mellom tolkning av rettsregler og organisering av utviklingsarbeidet.

Jeg skal nå utdype hver enkelt av de tre delproblemstillingene.

### 2.2.2 Problemstilling 1: Hvordan har NAV organisert arbeidet med å utvikle et profileringsystem?

Jeg bruker ordet «organisering» i vid forstand, samtidig som jeg sikter til en avgrenset rekke temaer og aktiviteter. De temaene jeg tenker er spesielt relevante for å forstå profilene, er **mandat** og **formål** for prosjektet, **beslutningsprosesser** i prosjektet, og hvilken **kompetanse** som utnyttes i prosjektet. I tillegg ser jeg på hvordan systemet og de rettslige vurderingene **dokumenteres**. Målet er med andre ord å få et dekkende bilde for hvordan arbeidet foregår.

Grunnen til at jeg mener at momentene knyttet til organisering av arbeidet er viktig for å forstå hvorvidt profilene er forankret i rettslige prosesser, er fordi de menneskene som er involvert i et utviklingsarbeid påvirker produktet.<sup>10</sup> I dette ligger det to muligheter: Å granske holdninger og verdier blant dem som er involvert i utviklingsarbeidet, eller å få en oversikt over mer «offisiell» kompetanse og bakgrunn. Det er sistnevnte tilnærming jeg har valgt i denne undersøkelsen; jeg forutsetter med andre ord at det samlede kompetansegrunnlaget blant dem som deltar i utviklingen av profileringssystemet er utslagsgivende for den endelig tekniske innretningen. Jeg vier også spesiell oppmerksomhet til hvordan juridisk kompetanse er involvert i arbeidet.

Beslutningsprosesser er i tillegg viktig å kartlegge, for å forstå hvordan personene som er involvert i arbeidet jobber sammen, og hvilke typer argumenter som vektlegges i beslutninger. Dokumentasjon av innholdet i modellen og de rettslige vurderingene er også særlig viktig fordi det legger til rette for legalitetskontroll av systemet som utvikles.<sup>11</sup>

### **2.2.3 Problemstilling 2: Hva er den tekniske innretningen til profileringssystemet?**

Problemstilling 2 dreier seg om å undersøke den tekniske utformingen av profileringssystemet, med et særlig fokus på utfordringer og begrunnelsene for at systeminnretningen er som den er. Formålet med dette er å kunne trekke linjer til rettslige og politiske krav til systemet. Jeg understreker at systemet jeg undersøker er under arbeid, og at det derfor er enkelte deler ved profileringssystemet som er mindre gjennomarbeidet enn andre.

De konkrete aspektene ved innretningen jeg undersøker nærmere er **maskinlæringsalgoritmen** som er brukt som utgangspunkt for modellen, **informasjonstyper** som er inkludert i profilen, **datakilder** for dataene som brukes i modellen, og i hvilken grad **mønstre** i datasettene taler for seg selv. Da jeg startet dette arbeidet, planla jeg å se nærmere på noen av de mer verdiladede valgene som gjerne knyttes til maskinlæringsarbeider, som vektning av falske negative og falske positive og arbeidet med å håndtere skjevheter i datasettene.<sup>12</sup> Det ville i enda større grad ha gjort det mulig å kartlegge og diskutere de grunnleggende verdiladede valgene som legges til grunn for profilene. Det var imidlertid vanskelig fordi den konkrete tekniske utformingen knyttet til slike valg har kommet kort. Dette er derfor ikke behandlet i rapporten.

---

<sup>10</sup> Sml. Feng (2008) s. 106.

<sup>11</sup> Schartum (2018) s. 280.

<sup>12</sup> Weiskopf (2020) s. 9.

### 2.2.4 Problemstilling 3: Hvordan har rettsprinsipper og rettsregler, herunder eventuelle rettslige usikkerheter, påvirket systeminnretningen?

Gjennom problemstilling 3 gir jeg oversikt over hvilke rettsregler som påvirker utviklingen av profileringsystemet. Samtidig er det viktig å påpeke at rettsreglene ikke er sentrale *i seg selv*. Dette er med andre ord ikke en drøftelse av innholdet i rettsreglene, men en drøftelse av rettsreglenes påvirkning på systeminnretningen.

I tillegg forventet jeg ved oppstart av arbeidet at det finnes en del rettsregler som er svært uklare ved utvikling av profileringsystemet. Jeg ser derfor nærmere på hvordan disse uklarhetene påvirker systeminnretningen.

Jeg skiller mellom rettsregler som ramme, og rettsregler som innhold for profileringsarbeidet.<sup>13</sup> Det er imidlertid rettsregler som *ramme* for profileringsarbeidet som er mest fremtredende, og som jeg derfor vektlegger.

## 2.3 Forskningsstatus

Dette forskningsarbeidet befinner seg i mellomrommet mellom to akademiske strømninger: juridiske og samfunnsvitenskapelige vurderinger rundt bruk av maskinlæring for å profilere, og kartleggingsarbeider om bruk av kunstig intelligens i offentlig sektor. Jeg skal kort skissere hva som er status i begge disse forskningsfeltene, for å kunne argumentere for behovet for mitt eget forskningsopplegg.

Det finnes en rekke dyktige forfattere som diskuterer profilering, gjerne med utgangspunkt i en eller annen form for teoretisk forståelse av profilering.<sup>14</sup> Stort sett vurderer forfatterne bruk av profilering på generell basis (riktignok innenfor en gitt sektor eller et gitt område), og påpeker hvilke særlige problemstillinger samfunnet bør ta stilling til når bruk av profilering blir mer utbredt. Det kan omfatte nedkjølingseffekten, skjeve algoritmer, mangel på gjennomskiktighet og innsyn, mangel på autonomi, osv.<sup>15</sup> Utgangspunktet for disse artiklene er gjerne enten omfattende bruk av data for å profilere i privat sektor,<sup>16</sup> eller bruk av profilering i helautomatisert saksbehandling for å fordele grunnleggende goder til

---

13 Se mer om dette i avsnitt 3.3.2 nedenfor.

14 Se f.eks. Haggerty (2000), Hildebrandt (2008) og Weiskopf (2020).

15 Yeung (2018), Pasquale (2015), Hildebrandt (2015), Lum (2016), Delacroix (2018), og Leese (2014).

16 Zuboff (2019), Weiskopf (2020).

enkeltpersoner.<sup>17</sup> Enkelte forfattere peker også på sammenhenger mellom disse to, der offentlige myndigheter utnytter profilene laget av private aktører.<sup>18</sup>

Slike diskusjoner er viktige, fordi dette er former for profilering som reelt sett foregår.<sup>19</sup> Samtidig risikerer artikler som tar utgangspunkt i dystopiske eksempler å gå glipp av viktige nyanser i konkrete utviklingsprosjekter som er mindre inngripende, men like fullt innebærer bruk av profilering og maskinlæring.<sup>20</sup> Diskusjoner rundt profilering som beslutningsstøtte, som motsetning til helautomatiserte beslutningssystemer, er for eksempel noe underbehandlet i faglitteraturen.<sup>21</sup> I disse prosjektene kan bastante argumenter og konklusjoner, trukket på basis av helautomatiserte beslutningssystemer, være lite treffende, og dermed forkastes. Risikoen er at slik beslutningsstøtte utvikles uten tilstrekkelig granskning, fordi den er mindre inngripende enn beslutningssystemene.

Den andre særlig aktuelle forskningstrenden er undersøkelser av bruk av kunstig intelligens i forvaltningen i Norge. Broomfield og Reutter kartlegger bruken av kunstig intelligens i forvaltningen generelt,<sup>22</sup> basert på en tilsvarende undersøkelse som har blitt gjort i Canada.<sup>23</sup> I tillegg til denne omfattende (og pågående) undersøkelsen, finnes det noe overordnet kartlegging av bruk av kunstig intelligens i kommunene i Norden.<sup>24</sup> Undersøkelsene tyder på at de fleste initiativer foreløpig fortsatt er på planleggingsstadiet.

De normative diskusjonene rundt bruken av profilering foregår dermed med utgangspunkt i dramatiske anvendelsesområder, som helautomatiserte beslutningssystemer, eller systemer som innebærer omfattende og usynlig sporing av enkeltpersoner på nettet. Konkret forskning på kunstig intelligens i forvaltningen i Norge er gjerne forbeholdt mer overordnede, kartleggende perspektiver. Kvantiteten ved slike omfattende kartleggingsarbeider går nødvendigvis på bekostning av dybdeforståelsen for konkrete anvendelser.

Forskningsrommet jeg utnytter er gapet mellom de normative og rettslige vurderingene, og konkrete vurderinger av teknologi i offentlig sektor. Det at kunstig intelligens i liten grad anvendes i Norge, er desto større grunn til å undersøke

---

17 Eubanks (2018), Büchi et al. (2020).

18 Se f.eks. Taylor (2018).

19 Se konkrete eksempler fra privat sektor i Panoptikon Foundation (2020) og Zuboff (2019), og fra offentlig sektor i AlgorithmWatch (2020) s. 160 flg., Lum (2016), og Eubanks (2018).

20 Veale et al. (2018) s. 3.

21 Se imidlertid Veale (2019).

22 Broomfield (2019) s. 2.

23 Redden (2018).

24 Nordisk ministerråd (2019).

utviklingsarbeidene som faktisk foregår. Disse prosjektene kommer til å legge føringer for framtidige anvendelser av kunstig intelligens i offentlig forvaltning i Norge, enten prosjektene er vellykkede eller ei.

For å undersøke dette forskningsrommet, tar jeg utgangspunkt i *critical data studies*, som dreier seg om å se på hvordan data produseres, analyseres og anvendes.<sup>25</sup> Utgangspunktet er at arbeidet med data ikke er verdinøytralt, men at det preges både av de konkrete personene som utvikler algoritmene og bredere samfunnsnormer.<sup>26</sup> Forskningen kommer som en reaksjon på at stordata og statistikk forbindes med nøytralitet.<sup>27</sup>

Mitt forskningsarbeid her er samtidig del av en mastergrad innenfor forvaltningsinformatikk, der vi jobber i krysningpunktet mellom jus, teknologi, og samfunnsvitenskap. Forvaltningen skal utføre sine myndighetsoppgaver i tråd med rettsregler, uavhengig av teknologien som brukes til å utføre oppgavene. Det må forventes at maskinlæringsmodeller som utvikles til bruk i offentlig forvaltning derfor er laget i tråd med gjeldende rettsregler. Derfor trekker jeg linjer fra studier av hvordan data produseres, analyseres, og anvendes, til hvilke rettsregler som kan eller skal påvirke dette arbeidet.

## 2.4 Undersøkelsesopplegget

### 2.4.1 Det overordnede undersøkelsesopplegget: Case studie

Alle forskningsspørsmålene mine er eksplorative, de innebærer en undersøkelse av dagsaktuelle hendelser, og hendelsene foregår utenfor min kontroll. I tillegg ønsker jeg å forstå hvilke beslutninger som har blitt tatt innenfor et prosjekt, og hvordan beslutningene blir tatt. Dermed er et case-studie en svært egnet måte å organisere undersøkelsesopplegget mitt.<sup>28</sup> Med case-studie som ramme for det metodiske opplegget mitt har jeg hatt muligheten til å gå i dybden på de forskjellige delene av et profileringsarbeid: organisering, teknisk innretning, og juridisk påvirkning på prosessen.

Jeg har kun ett case. Det skyldes hovedsakelig at det finnes få aktuelle prosjekter å undersøke. Profileringsarbeid basert på maskinlæring utvikles foreløpig i liten grad i norsk offentlig forvaltning, og mange av utviklingsprosjektene som finnes, har kommet såpass kort at de likevel er vanskelige å undersøke. Dette lave antallet potensielle undersøkelsesenheter betyr til gjengjeld at forskningsarbeidet mitt

25 Kitchin (2014).

26 Se f.eks. Introna (2007), Gitelman (2008), boyd (2012), Kaufmann (2020), og Redden (2018).

27 Sml. Feng (2008) s. 110.

28 Sml. Yin (2018) s. 9 og s. 14.

belyser et fenomen som er såpass unikt at et enkeltcase-studie likevel kan forsvares.<sup>29</sup> Når fenomenet samtidig er teoretisert og diskutert i stor grad,<sup>30</sup> kan mine empiriske funn innenfor én virksomhet bidra til å kaste lys over de teoriene som allerede finnes.

Case-studier, og særlig enkeltcase-studier, kritiseres generelt for å være vanskelige å generalisere.<sup>31</sup> Målet med denne rapporten er imidlertid ikke å si noe om hvordan profilering foregår generelt, men heller å systematisere arbeidet som gjennomføres i det aktuelle prosjektet, for å kunne knytte dette til annen, eksisterende teori. Jeg gjennomfører med andre ord en analytisk generalisering,<sup>32</sup> som kan brukes som utgangspunkt for andre profileringsprosjekter, som utgangspunkt for vurderinger av det spesifikke profileringsprosjektet jeg undersøker, eller som utgangspunkt for å vurdere behovet for nye rammeverk eller regulering.

Jeg skal nå gå gjennom de metodene jeg har brukt innenfor rammen av case-studiet, og drøfter deretter sammenhengen mellom metodene.

#### **2.4.2 Dokument- og litteraturstudier**

Jeg påpeker innledningsvis at når jeg her refererer til «dokumenter» favner begrepet bredt, og omfatter tradisjonelle, skriftlige dokumenter, men også bilder, podcaster, publiserte videoklipp, og lover. Slike dokumenter har vært en del av datagrunnlaget til alle de tre forskningsspørsmålene mine.

Det finnes noen offentlig tilgjengelige dokumenter om utviklingsarbeidet til NAV, og de jeg har brukt fremgår av litteraturlisten. For å få holdbare svar på spørsmålene min, anså jeg imidlertid tilgang til interne dokumenter som sentralt. Jeg fikk tilgang til alle dokumentene jeg etterspurte, unntatt kildekode til systemet. På grunn av manglende erfaring med koding, er det lite sannsynlig at kildekode ville gitt meg ytterligere forståelse for den tekniske innretningen, uten at jeg brukte for mye tid på den. Informanten min påpekte det samme. Mangel på innsyn i kildekode svekker imidlertid datagrunnlaget, ettersom en del informasjon om den tekniske innretningen til systemet kun er tilgjengelig skriftlig i kildekode.

Ellers var det en rekke relevante interne dokumenter jeg *ikke* etterspurte, som teams-logger og e-poster. Slike dokumenter inneholder gjerne mer personlige

---

29 Sml. *ibid.* s. 50.

30 Se mer om dette i avsnitt 1.3 over.

31 Yin (2018) s. 20.

32 Yin (2014) s. 68.

elementer, som kun er ment for konkrete mottakere.<sup>33</sup> Tilgang til slike kilder blir derfor mye mer inngripende enn tilgang til mer offisielle, interne kilder. Jeg valgte derfor å ikke etterspørre dokumentene, og det må forventes at jeg har gått glipp av en del interessante, interne funn. Jeg har imidlertid hatt mulighet til å avdekke enkelte slike diskusjoner gjennom intervjuer med nøkkelpersoner i prosjektet. Det er også verdt å merke seg at manglende «offisiell» dokumentasjon av slike interne diskusjoner er et interessant funn i seg selv.<sup>34</sup>

### **2.4.3 Intervjuer med informanter**

#### **2.4.3.1 Generelt om intervjuene**

Jeg gjennomførte tre semi-strukturerte intervjuer, der formålet var å få innsikt i erfaringene fra prosjektet, samt begrunnelser for valgene som er tatt i prosjektet. Alle intervjuene varte omtrent en time, og ble gjennomførte på Zoom med video og lyd.<sup>35</sup>

Jeg tok ikke opptak av intervjuene, for å unngå at informantene skulle føle ubehag og dermed potensielt holde tilbake.<sup>36</sup> I stedet for noterte jeg løpende underveis i intervjuene, og skrev et referat straks etter intervjuene. Ettersom jeg ikke tok lydopptak, har jeg ikke oversikt over hvordan jeg ordla oppfølgingsspørsmålene min. Det er derfor mer utfordrende å vurdere kvaliteten på datagrunnlaget mitt.

Etter alle intervjuene sendte jeg referater til informantene for gjennomgang, slik at informanten kunne korrigere eventuelle feil, og utdype dersom jeg hadde gått glipp av noen sentrale poenger.<sup>37</sup> Jeg fikk bekreftelser fra informantene på innholdet i alle de tre referatene, med utdypende kommentarer til det organisatoriske og det rettslige intervjuet.

#### **2.4.3.2 Utvalg av informanter**

Jeg har hatt to informanter; én data scientist, og én jurist. Data scientisten har fulgt utviklingsprosjektet siden det startet, og egnet seg som informant både for tekniske og organisatoriske elementer. Jeg gjennomførte dermed to intervjuer med denne informanten, hvilket gjør at dataene mine er preget spesielt mye av én person. Informanten hadde imidlertid god og relevant kunnskap, og var villig til å dele informasjon om prosjektet. Jeg mener derfor at informanten har gitt meg et bredt datagrunnlag.

---

33 Jacobsen (2015) s. 189.

34 Sml. også avsnitt 1.2.2.

35 Sml. *ibid.* s. 149.

36 Sml. Grønmo (2004) s 164.

37 Grønmo (2004) s. 164.

Intervjuet med juristen gjennomførte jeg veldig sent i forskningsarbeidet. Etter som juristen bare nylig har begynt å jobbe med problemstillingene jeg ønsket å diskutere, var det nødvendig å vente så lenge som mulig, slik at vedkommende hadde rukket å opparbeide seg så god kunnskap og forståelse for tematikken som mulig. Det innebærer imidlertid at datagrunnlaget jeg fikk gjennom intervjuet med denne informanten, stort sett er begrenset til de vurderingene som har blitt gjort siste halvdel av våren 2021.

Jeg har ikke gjennomført noe intervju med noen av fagpersonene i sykefraværsteamet.<sup>38</sup> Det var ikke før etter intervjuet med juristen at jeg skjønnte at jeg egentlig hadde behov for et intervju med en slik informant, og da manglet jeg tid. Det innebærer at jeg har et snevrere datagrunnlag enn ønskelig for enkelte aspekter både ved problemstilling 1 og 3.

#### **2.4.3.3 Drøftelser av konkrete intervjuer**

Under det tekniske intervjuet tok jeg en ganske fremtredende intervjuer-rolle. Det er derfor en risiko for at mine synspunkter har preget det datagrunnlaget jeg har samlet inn.<sup>39</sup> Samtidig satt jeg igjen med en del spørsmål etter intervjuet, fordi jeg ikke ba om utdypninger der det var behov. Det må derfor antas at funnene mine rundt den tekniske innretningen er noe begrenset.

Når det gjelder intervjuet rundt rettslige vurderinger, var også dette intervjuet utfordrende, men av andre grunner. Informanten startet med å gjennomgå en powerpoint om de rettslige vurderingene NAV diskuterer med Datatilsynet i forbindelse med arbeidet i den regulatoriske sandkassen. Dette var ikke irrelevant informasjon, men gikk like fullt på bekostning av det intervjuopplegget jeg hadde planlagt. Funnene mine gjenspeiler derfor særlig utfordringer knyttet til forholdet mellom personvernlovgivning og forvaltningslovgivning.

#### **2.4.4 Juridisk metode**

Juridisk metode innebærer å finne fram til gjeldende rett ved å anvende rettskilderegler.<sup>40</sup> Dette forskningsarbeid tar ikke primært sikte på å finne fram til gjeldende rett, men snarere å undersøke hvordan NAV jobber med å tolke rettsregler.<sup>41</sup> Omfattende og dyptgående anvendelse av juridisk metode er dermed hverken nødvendig eller egnet til å besvare forskningsspørsmålene mine. Når jeg har brukt rettskilder i dette arbeidet, har jeg istedenfor brukt dem til å forstå caset, og til å forså NAVs rettslige erfaringer fra utviklingsarbeidet. Rettskilde-

---

38 Se mer om disse i avsnitt 4.2.1.

39 Sml. Grønmo (2004) s. 192.

40 Boe (2020) s. 85.

41 Sml. avsnitt 1.2.4.



bruken i denne framstillingen er med andre ord overordnet, men like fullt basert på juridisk metode.

For å forstå lovhjemmelen som er utgangspunktet for profileringssystemet, har jeg for eksempel vurdert folketrygdloven § 8-7 a (heretter ftrl.) i lys av andre rettskilder, som rundskrivet til folketrygdloven, og en intern veileder for tolkning av paragrafen. I tillegg har jeg gjort noen søk etter relevante dommer i Lovdata, både i Trygderetten og domstolene. Jeg fant ingen relevante dommer. For å tolke enkelte bestemmelser i personvernforordningen, har jeg brukt uttalelser fra Personvernrådet.<sup>42</sup> Jeg har også brukt juridisk litteratur.

#### **2.4.5 Metodetriangulering**

Jeg har brukt ulike metoder og kilder til å supplere hverandre på flere måter. Jeg brukte offentlig tilgjengelige dokumenter som grunnlag for utforming av intervjuguidene, samt varierende grad av interne dokumenter. I etterkant av intervjuene har jeg generelt brukt dokumenter til å nyansere eller supplere informasjon fra informantene. Dette var spesielt viktig etter intervjuet om rettslige vurderinger. Det finnes veldig begrenset dokumentasjon av den tekniske innretningen til profileringsystemet, og jeg har derfor ikke hatt mulighet til å triangulere denne informasjonen nevneverdig.

Jeg har også triangulert mellom andre metoder. Kombinasjonen av juridisk metode, dokumentstudier og intervjuer har vært nødvendig for å forstå formålet med systemutviklingsarbeidet til NAV. Her har metodene delvis supplert hverandre, og delvis belyst enkelte potensielle motsetningsforhold. I tillegg har jeg sammenliknet informasjonen jeg har fått fra de to informantene, for å finne tilsvarende poenger eller motsetningsforhold.

## **2.5 Oversikt over den videre framstillingen**

For å tydeliggjøre hvordan profileringsbegrepet brukes i denne framstillingen, redegjør jeg for den teoretiske forklaringen av profileringsbegrepet i kapittel 2. Det omfatter både en forklaring av hva profilering er, og hvilke metoder som kan brukes for å profilere. I tillegg gjennomgår jeg et utvalg problemstillinger knyttet til bruk av profilering basert på maskinlæring i saksbehandling i offentlig sektor.

---

<sup>42</sup> Personvernrådet skal gi veiledning for tolkning av bestemmelsene i forordningen, for å sikre enhetlig anvendelse, jf. personvernforordningen artikkel 70.

I kapittel 3 redegjør jeg for føringene og rammene rundt systemutviklingsarbeidet. Det innebærer at jeg beskriver saksbehandlingskonteksten som profileringssystemet skal inngå i, samt de politiske føringene som tilsier at NAV skal utvikle et slikt system. Jeg presenterer også enkelte sentrale rettslige rammer som må vurderes ved utvikling av profileringssystemet.

Med kontekst, føring og rammer for systemutviklingsarbeidet i mente, presenterer jeg deretter funnene mine knyttet til hvert enkelt forskningsspørsmål i kapittel 4. Jeg setter disse fortløpende i sammenheng med teori og andre forskningsfunn, for å skape en flyt. I tillegg trekker jeg på nødvendige kildekritiske vurderinger der det er relevant,<sup>43</sup> slik at funnene leses i kontekst av mangler og utfordringer i undersøkelsesopplegget.

Jeg oppsummerer funn og tilhørende drøftelser i avsnitt 4.5, for å kunne trekke noen bredere linjer på tvers av problemstillingene.

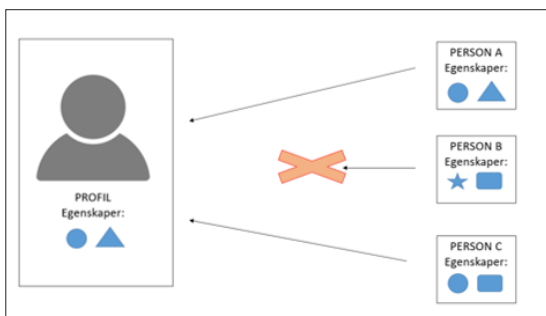
---

43 Basert på Jacobsen (2015) s. 227 flg. og Grønmo (2004) s. 122.

## 3 Profilering: Teori og utfordringer

### 3.1 Teoretisk forklaring av profilering

Å profilere betyr å lage en hypotese om hvilke egenskaper som kan knyttes til en bestemt gruppe mennesker.<sup>44</sup> Figur 2 viser en forenklet framstilling av en profil.



Figur 2. Profil og egenskaper

I figur 2 er profilen til venstre knyttet til konkrete egenskaper, som jeg har representert ved en trekant og en sirkel. Person A i figuren har de samme egenskapene som profilen, og passer derfor innenfor profilen. Person B har hverken en trekant eller en sirkel, og passer ikke innenfor profilen. Person C har en sirkel, mangler en trekant, men passer likevel innenfor profilen. De personene som passer innenfor en profil må med andre ord ikke dele *alle* egenskapene til profilen; det er tilstrekkelig at personene har mange nok like egenskaper som profilen.<sup>45</sup> Profiler er dermed en samlekategori for personer som har varierende grad av felles egenskaper med profilen.<sup>46</sup>

Det finnes forskjellige metoder som kan brukes for å bestemme hvilke egenskaper som skal knyttes til en profil. Når profilering skal brukes i formaliserte prosesser, som for eksempel saksbehandling, går det et skille mellom to hovedmåter

44 Hildebrandt (2008) s. 18.

45 Hildebrandt (2008) s. 21.

46 Se imidlertid González (2019) s. 610 flg. for en nærmere diskusjon rundt ulike typer profiler, blant annet profilering som kun knytter en profil til enkeltpersoner der personen og profilen deler alle egenskaper.

å lage profiler: ekspertkunnskaper og profilering basert på maskinlæring.<sup>47</sup> Profilering skjer i praksis ofte ved å kombinere begge metodene.<sup>48</sup> Nedenfor gjennomgår jeg likevel hver metode enkeltvis, for å tydeliggjøre forskjellene mellom dem.

Profilering basert på ekspertkunnskaper er den klassiske måten å profilere. Grunnantagelsen er at enkeltpersoner innehar enkelte egenskaper som egner seg for grupperinger, og som fagpersoner kan gjenkjenne og strukturere ved å trekke på sine fagkunnskaper og erfaringer fra feltet.<sup>49</sup> Et typisk eksempel er psykologer som bestemmer hvilke egenskaper som skal knyttes til konkrete psykiske lidelser, og som dermed grupperer enkeltpersoner etter disse egenskapene. Et mer relevant eksempel for denne rapporten er dersom fagpersoner i NAV lager en sjekkliste med egenskaper som tilsier at noen kommer til å være syk-meldte i lang tid. Sjekklisten kan deretter brukes til strukturert gruppering av enkeltpersoner, og eventuelt som utgangspunkt for automatiserte, regelbaserte systemer.

Maskinlæring og stordata er den «nye» måten å profilere. Slik profilering innebærer å samle inn store mengder data, som analyseres av en algoritme.<sup>50</sup> Algoritmen lærer på denne måten hvilke mønstre som finnes i dataene, og lager en *modell* som kan brukes til å sortere ut nye tilfeller. Grunnantagelsen er dermed at fortiden, slik den er representert i data, kommer til å gjenspeile fremtiden.<sup>51</sup> På denne måten gir profilene grunnlag for evidensbaserte beslutninger.<sup>52</sup> Til gjengjeld er representasjonen av virkeligheten gitt i form av *korrelasjoner*, ikke årsakssammenhenger. Det er dermed ikke knyttet faglige begrunnelser til profilene som fremgår av profileringsprosessen – begrunnelsen for å stole på profilene er store tall og statistikk.<sup>53</sup>

Jeg utelukker heretter andre former for maskinlæring enn overvåket maskinlæring.<sup>54</sup> Forskjellen på bruk av ekspertkunnskaper og maskinlæring til å profilere, kan da illustreres på følgende måte:

---

47 Weiskopf (2020) s. 14.

48 Se f.eks. Kaufmann (2017) s. 154.

49 Weiskopf (2020) s. 14.

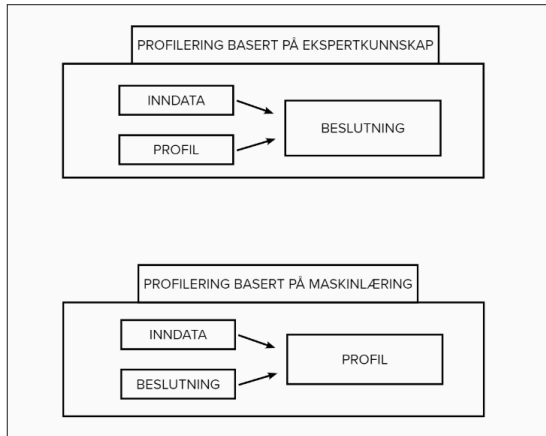
50 Favaretto (2019) s. 2.

51 Russell (2021) s. 1229.

52 Weiskopf (2020) s. 7.

53 Bench-Capon (2018) s. 2.

54 Se en oversikt over ulike typer maskinlæring i Russell (2021) s. 1205 flg.



Figur 3: Profilering basert på ekspertkunnskap/maskinlæring

Slik det fremgår av figur 3 innebærer ekspertbasert profilering at inndata (typisk personopplysninger) sammenliknes med en profil, som deretter leder til en beslutning om hvorvidt inndataene passer med profilen. Profilering basert på maskinlæring innebærer istedenfor at inndata sammenliknes med beslutninger, som deretter leder til en profil.

I den videre framstillingen bruker jeg «profilering» om profilering basert på maskinlæring, hvis ikke annet spesifiseres eksplisitt.

### 3.2 utfordringer knyttet til bruk av profilering i forvaltningen

Et (rettslig) beslutningssystem er helautomatisert, og innebærer dermed at resultatet systemet kommer fram til legges til grunn for vedtak.<sup>55</sup> Slike systemer kan være særlig utfordrende, fordi det ikke er noen mennesker som har kontroll over vedtakene. Denne framstillingen dreier seg imidlertid om bruk av profilering som beslutningsstøtte, hvilket vil si at systemet kun gir støtte til en saksbehandler som tar den endelige vurderingen.<sup>56</sup> Jeg skal nå kort redegjøre for enkelte trekk ved slik profilering som likevel er utfordrende. Det er verdt å påpeke at framstillingen nedenfor er preget av at jeg har begrenset plass, og der-

<sup>55</sup> Schartum (2018) s. 25.

<sup>56</sup> Ibid. s. 25.

for må nøye meg med å påpeke et utvalg problemstillinger. Jeg har heller ikke anledning til å gå i dybden på de enkelte problemstillingene.

Den første problemstillingen jeg påpeker, er knyttet til formålet med profilering. Kontroll kan være et fristende profileringsformål, fordi bruk av maskinlæring og profilering kan gjøre det billigere å gjennomføre omfattende og systematisk kontroll. Dette kan imidlertid være utfordrende i et demokratisk samfunn, der myndighetene ikke bør ha en «ubegrenset» kontrolladgang.<sup>57</sup> Selv om innsamling og bruk av store mengder personopplysninger til andre profileringsformål enn kontroll også kan være inngripende, må det imidlertid antas at andre typer formål minsker potensialet for en del av de negative konsekvensene som gjerne trekkes fram.<sup>58</sup> Jeg påpeker derfor behovet for å vurdere profileringsarbeidet i lys av formålet.

Videre er maskinlæring til profileringsformål grunnleggende sett en form for kunnskapsutvikling, der maskiner leter gjennom dataene for å finne potensielle sammenhenger.<sup>59</sup> Samtidig er myndighetsutøvende forvaltningsorganer underlagt departementer, både administrativt og faglig.<sup>60</sup> Dersom forvaltningsorganet selv (ved hjelp av en maskinlæringsalgoritme) kommer fram til hvilke egenskaper ved en person som tilsier at vedkommende faller innenfor eller utenfor en profil, er det en risiko for at forvaltningsorganet «hopper over» departementenes styringsmakt. Istedenfor at grunnleggende forståelse som er hentet ut av datasettene rapporteres til overordnet organ eller Stortinget (for eksempel i forbindelse med forslag til lovendringer), anvender forvaltningen kunnskapen om korrelasjoner direkte. Det må derfor forutsettes at en maskinlæringsalgoritme brukes innenfor rammen av de rettsregler som allerede finnes, der forvaltningen har fått en viss åpning til å vurdere hvilke egenskaper som skal passe innenfor eller utenfor profilen.

Dersom det først er bestemt at profilering basert på maskinlæring skal brukes i forvaltningen, er det en rekke utfordringer å vurdere, også når profilene fungerer som beslutningsstøtte. Den første handler om at teknologi gjerne forstås uavhengig av sin historie.<sup>61</sup> Dette er problematisk ved bruk av profiler, fordi både den tekniske innretningen og dataene som brukes til å trene modeller er formet av mennesker.<sup>62</sup> Dersom ikke menneskelig påvirkning synliggjøres, både til brukere av profilen og til dem som profileres, er det en risiko for at bruken av

---

57 Schartum (2016) s. 75.

58 Sml. FN (2019) s. 20.

59 Hildebrandt (2008) s. 30.

60 Graver (2019) s. 162.

61 Sml. Feenberg (2010) s. 7.

62 Thylstrup (2019) s. 2 og Muller (2019) s. 4.

profilene preges av *automation bias*, hvilket innebærer at profilene tilegnes for mye vekt.<sup>63</sup> For en saksbehandler som bruker profilene, kan det innebære at beslutningene som baserer seg på profilene blir feil, mens det for den som blir profilert kan være utfordrende å klage på beslutningen.

En annen utfordring med bruk av maskinlæring til profilering, er at en del maskinlæringsmodeller er det som kalles «svarte bokser». Dette er maskinlæringsmodeller som er så komplekse at det er umulig for utviklere, saksbehandlere og borgere å forstå hvordan profilene er laget.<sup>64</sup> Det kan også skyldes at den interne logikken til systemet er taushetsbelagt, fordi det regnes som en forretningshemmelighet, eller at åpenhet om den interne logikken går på bekostning av formålet med systemet (slik argumentet ofte lyder for systemer som skal brukes til å kontrollformål).<sup>65</sup> Ved bruk av svarte bokser kan saksbehandlere i NAV risikere å måtte stole blindt på beslutningsstøtten. Det blir også svært utfordrende for saksbehandler å gi en begrunnelse for beslutningen saksbehandler tar.

En tredje utfordring knyttet til maskinlæring er muligheten for å lage *dynamiske* modeller. Det betyr at modellen oppdaterer seg fortløpende etter hvert som den får tilgang på nye data.<sup>66</sup> Når profilene er dynamiske, blir det vanskeligere å gjennomføre demokratisk kontroll av profilene.<sup>67</sup> Det kan også bli vanskelig å koble prediksjoner til den modellen som faktisk ble brukt til å lage prediksjonen, slik at forklarbarhetsutfordringer blir enda mer fremtredende.

Avslutningsvis påpeker jeg at profilering generelt innebærer å strukturere et ledd i saksbehandlingsprosessen, som kan innebære en endring i rettsstillingen til enkeltborgere. Dersom vurderingene pleide å være helt skjønnsmessige, men går over til å være basert på profiler, kan enkelte borgere som ikke passer inn i «normen» bli særlig utsatt.<sup>68</sup> Bruken av en maskinlæringsmodellen endrer også hvordan informasjon brukes i praksis; fra at en saksbehandler bruker informasjon litt sporadisk, bruker en algoritme informasjonselementene konsekvent. Informasjonselementer som var mindre viktige i praksis ved manuell behandling, kan potensielt bli mer viktig i automatisert behandling. Dersom det ikke allerede utnyttes regelbasert profilering i prosessen der man vurderer å introdusere profilering basert på maskinlæring, er det derfor grunn til å være ekstra oppmerksom på hvordan profilene endrer prosessen.

---

63 Sml. Yeung (2018) s. 516.

64 Rudin (2019) s. 206.

65 Se f.eks. *NCJM and Others v. State of the Netherlands* avsnitt 6.49.

66 Weiskopf (2020) s. 17.

67 Yeung (2018) s. 517.

68 Sml. Schartum (2018) s. 120.

## 4 Kontekst og rammer for utviklingsarbeidet i caset

### 4.1 Saksbehandlingskonteksten

NAV er underlagt Arbeids- og sosialdepartementet, og har ansvar for å fordele ytelser ved for eksempel sykefravær, arbeidsløshet, og uførhet.<sup>69</sup> I tillegg har NAV ansvar for oppfølging av sykmeldte, og andre tjenester til mennesker utenfor arbeid. Ettersom borgere som benytter seg av tjenestene og ytelsene fra NAV gjerne er i særlig utsatte situasjoner, har NAV større makt overfor den enkelte enn mange andre offentlige virksomheter.<sup>70</sup>

Maskinlæringsmodellen NAV jobbet med våren 2021, og som jeg har som case, innebærer prediksjoner av sykefraværslengde. Prediksjonene skal inngå som beslutningsstøtte innenfor sykefraværsoppfølgingsarbeidet til NAV, og jeg skisserer derfor i det følgende kort saksbehandlingskonteksten. Ettersom maskinlæringsmodellen jeg ser på i denne sammenheng utelukkende gjelder tjenestområdet til NAV, avgrensner jeg mot alle deler av sykefraværsoppfølgingen som innebærer fordeling av ytelser, f.eks. aktivitetsplikten ved uke 8.<sup>71</sup>

Sykefraværsoppfølging skjer på bakgrunn av lovkrav i både arbeidsmiljøloven og folketrygdloven. I første omgang skal arbeidsgiver og arbeidstaker selv forsøke å finne en måte å få den sykmeldte tilbake i arbeid.<sup>72</sup> Det innebærer i praksis at arbeidsgiver og arbeidstaker sammen skal utarbeide en oppfølgingsplan, som skal inneholde konkrete tiltak som kan bidra til å få arbeidstaker tilbake i arbeid. Aktørene er også pålagt å gjennomføre et første dialogmøte,<sup>73</sup> der de diskuterer oppfølgingsplanen. Både arbeidstaker og arbeidsgiver kan involvere NAV i dette stadiet, dersom de trenger ekstern bistand.<sup>74</sup>

Dersom den sykmeldte ikke er tilbake i arbeid innen 16 uker, skal NAV vurdere om arbeidsgiver, den sykmeldte, og eventuelt sykmelder<sup>75</sup> bør innkalles til dialogmøte 2.<sup>76</sup> Dialogmøte 2 skal gjennomføres senest innen 26 ukers sykefra-

---

69 NAV (2020) s. 6.

70 Sml. Graver (2019) s. 65.

71 Jf. ftrl. § 8-4 andre ledd.

72 Jf. aml. § 6-4.

73 Jf. aml. § 6-4 fjerde ledd første punktum.

74 Jf. aml. § 4-6 fjerde ledd femte punktum.

75 Jf. Felles datakatalog (2020).

76 NAV (2019).



vær,<sup>77</sup> og er innført for å styrke oppfølgingen av langtidssykemeldte, slik at sannsynligheten for at de kommer tilbake i arbeid øker.<sup>78</sup> Dialogmøte 2 forutsettes med andre ord å kunne være utslagsgivende for sykefraværslengden til den enkelte.<sup>79</sup>

Ved vurderingen av behovet for dialogmøte 2 sender saksbehandler i NAV (heretter veileder) en melding til arbeidstaker og arbeidsgiver, og spør om partene mener det er behov for et nytt dialogmøte. Dersom noen av dem ber om at dialogmøte 2 skal gjennomføres, eller at det skal gjennomføres tidligere enn uke 26, skal NAV innkalle til et møte.<sup>80</sup> Dersom ingen av dem etterspør dialogmøte 2, åpner loven for at veileder hos NAV kan velge å ikke invitere aktørene til dialogmøte 2. Det forutsetter imidlertid at dialogmøte 2 anses som «åpenbart unødvendig».<sup>81</sup>

Unntaket fra kravet om dialogmøte 2 skal generelt tolkes snevert.<sup>82</sup> I rundskrivet til bestemmelsen fremkommer det enkelte konkrete eksempler som kan gi unntak for kravet om å gjennomføre dialogmøte 2. Den sykmeldte kan for eksempel være så syk at vedkommende ikke kommer til å kunne komme tilbake i arbeid, uavhengig av dialogmøtet. Det kan også være tilfeller der den sykmeldte ikke kan delta på grunn av innleggelse på en helseinstitusjon. Det påpekes at et eventuelt unntak må basere seg på *dokumenterte* opplysninger som tilsier at dialogmøte 2 er unødvendig.<sup>83</sup>

Det er også to eksempler som knytter seg til sannsynlig friskmelding innen henholdsvis 28 og 39 uker.<sup>84</sup> Begge disse to vurderingene forutsetter dokumenterte begrunnelser for veilederens vurdering av sykefraværslengde. For å kunne gi unntak fra dialogmøte 2 på bakgrunn av forventet friskmelding innen 28 uker, må for eksempel veileder ta en grundig gjennomgang av tidligere medisinske opplysninger, som legens vurdering, progresjon og behandlingsopplegg.<sup>85</sup> I tillegg skal veileder vurdere om det er grunn til å stole på at sykmelder har satt et realistisk friskmeldingstidspunkt eller ikke. En slik vurdering kan basere seg på veilederens erfaringer med sykmelderens praksis. Det kan for eksempel dreie seg om hvorvidt sykmelder ofte forlenger sykmeldingen etter hvert som syke-

77 Jf. ftrl. § 8-7 a andre ledd første punktum.

78 Prop.89 L. (2010–2011), s. 33.

79 Se for øvrig en kommentar om manglende kunnskapsgrunnlag for denne antakelsen i Prop. 1 S (2020–2021) s. 84.

80 Jf. ftrl. § 8-7 a andre ledd, andre punktum.

81 Jf. ftrl. § 8-7 a andre ledd, første punktum.

82 Prop. 89 L (2010–2011) s. 34.

83 Jf. NAV (01.05.1997) s. 77.

84 Jf. NAV (01.05.1997) s. 77.

85 Jf. NAV (01.05.1997) s. 77.

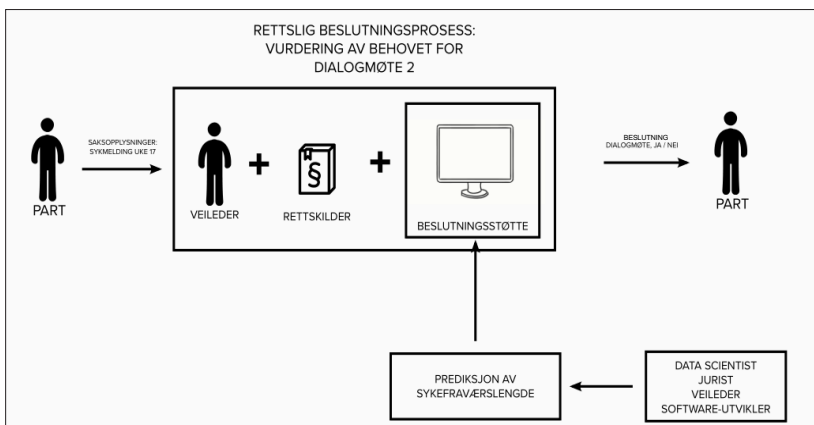
fraværet utvikler seg, eller om friskmeldingstidspunktet ofte forholder seg fast gjennom sykefraværsløpet.

Basert på rettskildegrunnlaget kan det virke som at en prediksjon av sykefravær enten ved uke 26, uke 28, eller uke 39 er særlig aktuelle. Prediksjonen av sykefravær ved uke 26 kan eventuelt begrunnes i at dialogmøte 2 skal holdes innen uke 26, mens prediksjon av uke 28 eller 39 kan begrunnes direkte i ordlyden i rundskrivene. Jeg påpeker imidlertid at det er visse betingelser knyttet til veileders vurdering av sannsynlighet for friskmelding, både for uke 28 og uke 39.

## 4.2 Beslutningsstøttesystemet i kontekst

Beslutningsstøttesystemet NAV utvikler skal hjelpe veileder å vurdere behovet for dialogmøte 2. Systemet er avgrenset til å kun innebære en sannsynlighet for om den sykmeldte vil forbli sykmeldt etter et visst punkt eller ikke. Prediksjonene er med andre ord et sett med profiler; enten kan den sykmeldte plasseres innenfor profilen «sykmeldt» eller «ikke-sykmeldt» ved et spesifikt tidspunkt. Denne profilen sendes til en veileder, som deretter tar en helhetlig, endelig beslutning basert på den tilgjengelige informasjonen.

Når den rettslige beslutningsprosessen (eller også saksbehandlingsprosessen) involverer et beslutningsstøttesystemet kan prosessen forstås på følgende måte:



Figur 4: Beslutningsstøttesystem i saksbehandlingskonteksten

Helt til venstre i figur 4 er parten, som har oversendt diverse saksopplysninger knyttet til sykefraværet. Jeg påpeker at det er flere aktører som sender saksopplysninger til NAV i denne sammenheng, der særlig sykmelder og arbeidsgiver er de mest sentrale. Jeg kommer tilbake til sykmelderens betydning i avsnitt 4.3.1, men skisserer for enkelhets skyld ikke vedkommende i denne figuren. Poenget er at saksopplysningene som er grunnlag for beslutningsprosessen hos NAV oversendes til NAV. Saksopplysningene består blant annet av sykmeldingene til den sykmeldte.

Den rettslige beslutningsprosessen som er illustrert i midten av figur 4 involverer en veileder hos NAV, aktuelle rettskilder for vurderingen av behovet for dialogmøte 2, samt et beslutningsstøttesystem. Veileder vurderer saksopplysningene opp mot rettskildene, og får i tillegg en prediksjon av sykefraværslengde fra beslutningsstøttesystemet. Basert på en helhetlig vurdering av den tilgjengelige informasjonen, bestemmer veileder om den sykmeldte skal kalles inn til dialogmøte 2 eller ikke. Denne vurderingen oversendes til parten.

I tillegg til å skissere saksbehandlingsløpet, har jeg markert hvordan beslutningsstøtten blir laget, nede til høyre i figur 4. Beslutningsstøtten må tross alt bygges før det kan utnyttes av veileder i konkrete rettslige beslutningsprosesser. Dette utviklingsarbeidet involverer flere aktører, som data scientister, softwareutviklere, designere, jurister og veiledere. I tillegg involverer utviklingsarbeidet store mengder data, som skal bidra som grunnlaget for prediksjonene.<sup>86</sup> Poenget er altså at slike beslutningsstøttesystemer bygges, og at det derfor kan være nødvendig å se nærmere på *hvordan* det enkelte beslutningsstøttesystem bygges.

Det er dette arbeidet med utvikling av maskinlæringsmodellen som er hovedtematikken for de resterende delene av denne framstillingen. Når modellen skal tas i bruk, er det en rekke problemstillinger som må vurderes. Et avgjørende spørsmål er for eksempel hvorvidt veileder faktisk bruker prediksjonene som beslutningsstøtte, eller om vedkommende vektlegger prediksjonene for mye eller for lite.<sup>87</sup> Dette er et felt NAV selv har vektlagt i sin vurdering av personvernkonsekvenser, og som skal testes ved produksjonssetting av systemet. I denne rapporten avgrenser jeg imidlertid mot slike vurderinger, og nøyer meg med å kommentere dem der det eventuelt er relevant for å belyse problemstillinger som gjelder for utviklingsarbeidet.

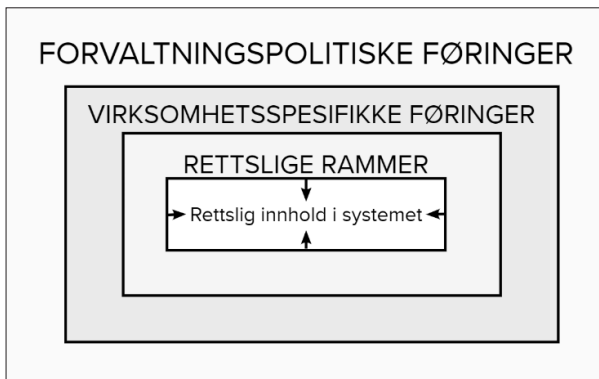
---

86 Sml. avsnitt 2.1.

87 Sml. avsnitt 2.2.

### 4.3 Rammer og føringer for systemutviklingsarbeidet

Når offentlig sektor utvikler systemer som skal introduseres i saksbehandlingsrutiner, finnes det en rekke forskjellige styrende dokumenter (både rettslige og politiske) å forholde seg til. Grunnleggende sett skal forvaltningen være gjenstand for demokratisk styring.<sup>88</sup> Jo mer automatisert rettsanvendelsen skal være, jo flere rettslige krav må systemutviklingsprosessen dekke.<sup>89</sup> Schartum deler disse opp i rettslige rammer, virksomhetsspesifikke føringer, og forvaltnings-spesifikke føringer.<sup>90</sup> Jeg har illustrert denne sammenheng i figur 5, nedenfor.



Figur 5: Rammer og føringer for utviklingsarbeidet

I midten av figur 5 er den delen av regelverket som er gjenstand for en automatiseringsprosess, som i denne sammenheng er gjennomgått i avsnitt 3.1 over. Automatisering av disse reglene må skje innenfor rettslige rammer, og virksomhetsspesifikke og forvaltningspolitiske føringer. Alle tre kan påvirke det rettslige innholdet, men hvis det er motstrid mellom de rettslige rammene og andre føringer, har de rettslige rammene alltid forrang. Jeg kommer i det følgende til å behandle rettslige rammer for seg. Først redegjør jeg imidlertid kort for forvaltningspolitiske og virksomhetsspesifikke føringer.

#### 4.3.1 Forvaltningspolitiske og virksomhetsspesifikke føringer

I nasjonal strategi for kunstig intelligens påpeker Regjeringen at Norge har gode forutsetninger for å bli ledende innenfor bruk av kunstig intelligens, herunder

<sup>88</sup> NOU 2019: 5 s. 141.

<sup>89</sup> Se Schartum (2017) s. 87.

<sup>90</sup> Schartum (2018) s. 85.

maskinlæring.<sup>91</sup> Som ledd i å utvikle kompetanse og et fortrinn innenfor bruk av kunstig intelligens, forventer Regjeringen at offentlig sektor skal ta i bruk teknologien, og maskinlæring som beslutningsstøtte trekkes fram som et eksempel på et aktuelt bruksområde.<sup>92</sup> Brukertilpasning av tjenester trekkes også eksplisitt fram i strategien, og i andre forvaltningspolitiske dokumenter som Digitaliseringsrundskrivet.<sup>93</sup> Det er med andre ord en rekke politiske signaler som tyder på at offentlig forvaltning, herunder NAV, bør utnytte maskinlæring, også som beslutningsstøtte. Generelt forutsetter disse politiske dokumentene at forvaltningen utvikler «ansvarlig kunstig intelligens», som blant annet innebærer at utviklingsarbeidene skal respektere menneskerettigheter, bygge på etiske retningslinjer, og være sikre.<sup>94</sup>

For NAV gjelder i tillegg noen virksomhetsspesifikke føringer, som blant annet poengterer behovet for å digitalisere og brukerrette sykefraværsoppfølgingen til NAV.<sup>95</sup> Dette handler overordnet sett om mest mulig effektiv bruk av fellesskapets ressurser,<sup>96</sup> men kan også ses i lys av muligheter for å forbedre oppfølging av de som er sykmeldt. Tilpasning av sykefraværsløpet for å sikre oppfølging av den enkelte sykmeldte, istedenfor å sende alle sykmeldte gjennom det samme løpet, har for øvrig vært en målsetting i mange år.<sup>97</sup> NAV er også tydelige på at de ønsker å utvikle ansvarlig kunstig intelligens.<sup>98</sup>

Budsjettmessig har NAV gjennomført omfattende satsning på å automatisere deler av sine tjenester, både innenfor sykefraværsoppfølgingen og ytelser. I 2019 ble det blant annet ferdigutviklet et beslutningsstøttesystem for sykmeldere.<sup>99</sup> Ellers er det primært modernisering på ytelsesområdet som har vært i fokus, med utskiftning av en rekke eldre saksbehandlingssystemer.<sup>100</sup> Utvikling av et profileringsystem som beslutningsstøtte i sykefraværsoppfølgingen til veiledere hos NAV, er på sin side ikke eksplisitt trukket fram i budsjetteringer fra Arbeids- og sosialdepartementet.

Slik det fremgår av de politiske dokumentene, er det uansett et gjennomgående tema at tjenestene til NAV skal være effektive, og utformet med tanke på bruker-

91 Kommunal- og moderniseringsdepartementet (2020) s. 5.

92 Ibid. s. 53.

93 Kommunal- og moderniseringsdepartementet (2021) punkt 1.

94 Kommunal- og moderniseringsdepartementet (2020) s. 56.

95 Meld. St. 33 (2015–2016) s. 32.

96 Ibid. s. 38.

97 Sml. prop. 39 L (2014–2015) s. 48.

98 Se f.eks. Digitaliseringsdirektoratet (2021) og Teknologirådet (2021).

99 Dette systemet skulle bidra til å skape en mer enhetlig sykmeldingspraksis på tvers av sykmeldere, se prop. 1 S (2020–2021) s. 85.

100 Ibid. s. 130.

tilpasset oppfølging. Et viktig spørsmål er i hvilken grad de politiske forventningene gjenspeiles i de konkrete rettsreglene som gjelder for utviklingsarbeidet.<sup>101</sup>

### 4.3.2 Rettslig rammer

Det finnes mange konkrete lover som kan eller bør fungere som rammer for utvikling av et rettslig beslutningsstøttesystem i offentlig forvaltning. Jeg har ikke anledning til å gjennomgå alle potensielle krav til utvikling av en maskinlæringsmodell, men kommer til å prioritere å gi et overblikk over noen førende prinsipper fra personvern- og forvaltningsretten som har betydning for lovligheten av maskinlæringsmodeller. I tillegg kommenterer jeg kort de rettslige rammene som gjelder for beslutningssystemer, versus beslutningsstøttesystemer. For en helhetlig oversikt over regelverk som fungerer som rammer for systemutvikling i offentlig forvaltning, se kapittel 6 i Schartum (2018). Jeg understreker at gjennomgangen nedenfor kun omfatter gjeldende regelverk. Forslag til en mulig, fremtidig EU-forordning om kunstig intelligens, herunder til bruk for profilering, kommenteres kort i avsnitt 4.4 nedenfor.

I personvernforordningen artikkel 5 nr. 1 fastsettes de overordnede prinsippene for behandling av personopplysninger, som både har selvstendig rettslig betydning og er utgangspunktet for tolkninger av de øvrige bestemmelsene i forordningen.<sup>102</sup> Det innebærer blant annet at personopplysninger skal behandles etter prinsipper om lovlighet, åpenhet og rettferdighet, formålsbegrensning, data-minimering, lagringsbegrensning, opplysningskvalitet, og konfidensialitet. Ettersom personvernforordningen inneholder flere rettslige rammer enn forvaltningsloven,<sup>103</sup> tar jeg utgangspunkt i prinsipper fra personvernretten, og trekker på forvaltningsrettslige prinsipper der det er relevant.

Det første prinsippet jeg gjennomgår er prinsippet om lovlighet, åpenhet og rettferdighet, jf. personvernforordningen artikkel 5 nr. 1 bokstav a. Kravet om at behandlingen skal være lovlig innebærer at personopplysninger må behandles i tråd med de øvrige bestemmelsene i personvernforordningen.<sup>104</sup> Det bør imidlertid også forstås som et krav om at behandlingen skjer i samsvar med *andre* rettsregler.<sup>105</sup> Prinsippet er dermed naturlig å se i sammenheng med legalitetsprinsippet fra forvaltningsretten, som innebærer at all myndighetsutøvelse må

---

101 Schartum (2018) s. 85.

102 Jf. Skullerud (2018) om personvernforordningen artikkel 5. Prinsippene har for øvrig vært en del av personvernteori i lang tid før de ble lovfestet i personvernforordningen, se f.eks. Schartum (2016) s. 114.

103 Schartum (2018) s. 127.

104 Skullerud (2018) om personvernforordningen artikkel 5 nr. 1 bokstav a.

105 Sml. Schartum (2020) s. 89.

ha hjemmel i lov.<sup>106</sup> Både behandlingen av personopplysninger og myndighetsutøvelsen må med andre ord være i tråd med gjeldende rettsregler.

Kravet om rettferdighet i personvernforordningen artikkel 5 nr. 1 bokstav a innebærer at sammenhengen mellom formålet med behandlingen og de personopplysningene som behandles skal være rimelig for de registrerte.<sup>107</sup> Det kan også tolkes som et krav til at motstridende interesser, ved vurdering av enkeltbestemmelser i forordningen, skal vurderes forholdsmessig.<sup>108</sup> Prinsippet kan forstås i lys av forvaltningsrettslige prinsipper om nøytralitet- og saklighet. Prinsippene innebærer blant annet at forvaltningen ikke kan ta vilkårlige eller urimelige avgjørelser,<sup>109</sup> for eksempel ved å ta hensyn til irrelevante saksopplysninger i forbindelse med et vedtak.

Kravet til åpenhet etter personvernforordningen artikkel 5 nr. 1 bokstav a innebærer at behandling av personopplysninger skal være forutsigbar og forståelig for de registrerte.<sup>110</sup> Det kan være naturlig å se dette prinsippet i kontekst av kontradiksjonsprinsippet fra forvaltningsretten. Prinsippet forutsetter langt på vei forklarbare beslutninger, for at borgere skal kunne supplere eller motsi saksgrunnlaget til en beslutning.<sup>111</sup> Det er vanskelig å få til hvis ikke veileder selv forstår begrunnelsen for en beslutning.<sup>112</sup> I tillegg er det relevant å se hen til offentlighetsprinsippet, som innebærer at forvaltningen skal drives på en åpen måte for å sikre demokratisk kontroll av forvaltningens virke.<sup>113</sup>

Prinsippet om formålsbegrensning innebærer at forvaltningen kun skal behandle personopplysninger til det formålet de ble samlet inn for, jf. personvernforordningen artikkel 5 nr. 1 bokstav b. Dersom forvaltningen skal behandle personopplysninger til et annet formål, må forvaltningen vurdere om det nye formålet er forenlig med det gamle formålet, jf. personvernforordningen artikkel 6 nr. 4. Det innebærer blant annet at forvaltningsorganet skal ta hensyn til om det finnes noen forbindelser mellom det nye og det gamle formålet, forholdet mellom forvaltningsorganet og de personene opplysningene handler om, om personopplysningene er sensitive, og om opplysningene er krypterte og/eller pseudonymiserte. Dersom det nye formålet ikke er forenlig med det gamle formålet, må det nye formålet enten være forankret i lov eller basere seg på sam-

---

106 Jf. GrL. § 113.

107 Skullerud (2018) om personvernforordningen artikkel 5 nr. 1 bokstav a.

108 Sml. Schartum (2020) s. 89.

109 Graver (2019) s. 123.

110 Skullerud (2018) om personvernforordningen artikkel 5 nr. 1 bokstav a.

111 Graver (2019) s. 413.

112 Se avsnitt 2.2.

113 Offl. § 1.

tykke.<sup>114</sup> Samtykke er neppe aktuelt gitt det ubalanserte maktforholdet mellom et forvaltningsorgan og borgerne, og behandlingen må derfor hjemles i lov.

Prinsippet om dataminimering innebærer at personopplysningene som brukes skal være begrenset til det som er nødvendige for formålet de behandles for, jf. personvernforordningen artikkel 5 nr. 1 bokstav c. I praksis innebærer dette at forvaltningsorganet bør vurdere om formålet kan oppnås ved bruk av færre personopplysninger.<sup>115</sup> Det kan blant annet begrunne at forvaltningsorganet pseudonymiserer eller anonymiserer personopplysningene. Prinsippet om dataminimering bør ses i sammenheng med prinsippet om lagringsbegrensning, jf. personvernforordningen artikkel 5 nr. 1 bokstav e. Lagringsbegrensningsprinsippet innebærer at opplysninger ikke bør kunne knyttes til enkeltpersoner lenger enn det som er nødvendig for formålet. Det er likevel lov til å lagre personopplysningene lenger for formål knyttet til arkivering, forskning, og statistikk. Det er verdt å merke seg at personopplysninger som oppbevares til slike formål, ikke nødvendigvis kan brukes til nye, «praktiske» formål,<sup>116</sup> jf. også formålsbegrensningsprinsippet som nevnt over. Dette kan både ha betydning for lagring av prediksjonene fra en maskinlæringsmodell, men også for de opplysningene NAV ønsker å bruke til å trene modellen.

Opplysningskvaliteten skal være så god som formålet med behandlingen krever, jf. personvernforordningen artikkel 5 nr. 1 bokstav d. Dette henger sammen med forvaltningsprinsippet om forsvarlig saksbehandling.<sup>117</sup> Det innebærer at opplysningene som brukes til å trene maskinlæringsmodellen må være gode nok til treningsformålet, og at prediksjonene må være av god nok kvalitet til bruk i saksbehandling.

Prinsippet om konfidensialitet i personvernforordningen artikkel 5 nr. 1 bokstav f innebærer at personopplysninger skal behandles på en måte som sikrer integriteten og konfidensialiteten til personopplysningene. Det innebærer blant annet en plikt til å sikre personopplysningene mot uautorisert innsyn.<sup>118</sup> I tilknytning til konfidensialitet er taushetsplikten fra forvaltningsloven § 13 første ledd nr. 1 relevant, som i praksis innebærer at opplysninger om enkeltpersoners personlige forhold ikke uten videre kan deles åpent innenfor forvaltningsorga-

---

114 Enkelte formål vil generelt likevel være tillatt, som f.eks. formål knyttet til arkivering, forskning, og statistikk, jf. personvernforordningen artikkel 5 nr. 1 bokstav b.

115 Skullerud (2018) om personvernforordningen artikkel 5 nr. 1 bokstav c.

116 Schartum (2020) s. 97.

117 Sml. Graver (2019) s. 102.

118 Skullerud (2018) om personvernforordningen artikkel 5 nr. 1 bokstav f.



net.<sup>119</sup> Disse prinsippene begrenser potensielt muligheten til å dele nødvendige treningsdata med utviklingsteamet.

Avslutningsvis påpeker jeg også skillet som går mellom rettslige rammer som gjelder for beslutningssystemer, og rettslige rammer som gjelder for beslutningsstøttesystemer.<sup>120</sup> Profilerings er definert i personvernforordningen artikkel 4 nr. 4, og omfatter automatisert behandling av personopplysninger med formål om å komme fram til eller forutsi visse personlige aspekter ved en fysisk person. Det betyr at begge typene profilering jeg var innom i kapittel 2 kan være omfattet av begrepet, og at profilering basert på maskinlæring generelt er omfattet av begrepet. De rettslige konsekvensene av at en behandling omfattes av personvernforordningen artikkel 4 nr. 4 er likevel begrenset. Personvernforordningen artikkel 22, som har «profilering» som del av navnet til bestemmelsen, kommer for eksempel kun til anvendelse ved helautomatisert behandling av personopplysninger.<sup>121</sup> Profilerings som beslutningsstøtte er dermed generelt unntatt. Denne bestemmelsen er utgangspunktet for enkelte andre krav i personvernforordningen, som en styrket rett til informasjon i personvernforordningen artikkel 15 nr. 1 bokstav h. Det finnes med andre ord enkelte særlige krav som gjelder for helautomatisert behandling av personopplysninger, herunder profilering. For delautomatisert behandling av personopplysninger, er derimot rettskildegrunnlaget mye spedere, og disse behandlingene må istedenfor vurderes basert på de mer generelle rettslige føringene som finnes.

---

119 Graver (2019) s. 377.

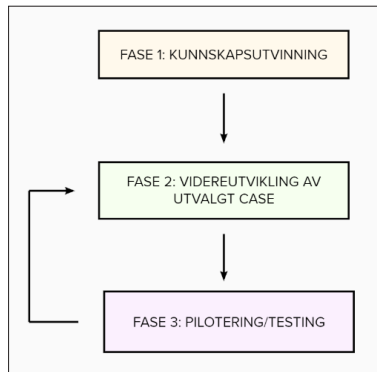
120 Se kort om dette i kapittel 2 over.

121 Se en nærmere drøftelse av bestemmelsen i Mendoza og Bygrave (2017) s. 87.

## 5 Arbeidet med prediksjon av sykefraværslengde

### 5.1 Fasene i utviklingsarbeidet

Jeg skal nå gjennomgå hvordan arbeidet med utvikling av et system for prediksjoner av sykefraværslengde foregår i NAV. For å gi et overordnet bilde av prosessen, starter jeg med å gjennomgå de overordnede fasene i utviklingsprosessen til NAV. Navnene på fasene tilsvarer ikke ordbruken til NAV, da det etter det jeg vet ikke finnes konsekvente betegnelser på fasene (utover henvisninger til fase 1, 2 og 3). Fasene er illustrert i figur 6 nedenfor.



Figur 6: Fasene i utviklingsarbeidet

Fase 1 i figur 6 har jeg kalt kunnskapsutvinningsfasen. Det innebærer å utforske det omfattende datagrunnlaget NAV sitter på om sykefraværsoppfølging i Norge. Formålet med fasen er å få et kunnskapsgrunnlag for å vurdere hvilke deler av sykefraværsløpet som kan forbedres, enten ved utvikling av nye verktøy, nye arbeidsmetoder, eller nye tiltak for oppfølging. Fasen er med andre ord en forberedende fase, som legger grunnlaget for forbedring av oppfølgingsløpet. For å få til dette, har NAV brukt tradisjonelle, statistiske modeller, som reflekterer et stillbilde av sammenhenger i datasettene. Data scientisten forklarte at analysene forutsetter et omfattende og tidkrevende arbeid, der data scientister må gjennomføre en rekke operasjoner på dataene, som å rette opp feil, formattere dataene, omkode variabler, og fange opp avvik.

Fase 2 i figur 6 er videreutvikling av et konkret utvalgt case, som er valgt på bakgrunn av kunnskapen som er utvunnet i fase 1. I denne sammenheng innebærer caset å lage prediksjoner av sykefraværslengde som beslutningsstøtte. Slik jeg har forstått data scientisten, er det først i fase 2 at arbeidet rettes mot et konkret utviklingsformål, som motsetning til bred kunnskapsutvinning.

Det er verdt å merke seg at arbeidet med prediksjoner av sykefraværslengde ikke baserer seg på et formelt mandat, og at det heller ikke er knyttet til noen formell satsning fra Arbeids- og sosialdepartementet. Prosjektet stemmer imidlertid godt overens med mer generelle forvaltningspolitiske føringer.<sup>122</sup> Formålet med modellen er, så vidt jeg har skjønt, å *effektivisere* saksbehandlingen, ved at kun de sykmeldte som har behov for dialogmøte 2 innkalles. I tillegg er det tenkt at effektivisering i praksis *forbedrer* saksbehandlingen, fordi (1) den bidrar til at veilederens praksis blir mer enhetlig, (2) den frigir veilederens tid til dem som har behov for dialogmøte 2, og (3) den unngår å bruke den sykmeldtes (og andres) tid når vedkommende ikke har behov for dialogmøte 2. Et viktig poeng er at modellen ikke har kontroll av borgerne som hovedformål.<sup>123</sup>

For å kunne bruke prediksjonene i praksis, må flere elementer komme på plass gjennom arbeidet i fase 2. Utviklingsteamet må hente ut mer spissede data som er beregnet til bruk til casets konkrete formål, som motsetning til dataene som ble hentet ut i forbindelse med fase 1. Teamet må også lage en maskinlæringsalgoritme som kan analysere dataene, og maskinlæringsalgoritmen og dataene må fungere innenfor NAVs samling av ulike saksbehandlingssystemer. Denne delen av arbeidet er også det jeg har sett mest på i denne undersøkelsen, og jeg kommer tilbake til nærmere funn i hovedavsnittene nedenfor (4.2, 4.3 og 4.4).

Fase 3 i figur 6 er pilotering av løsningen. Pilotering skal bidra til å teste både feil i datasystemet og beslutningsstøtten. Jeg har forstått det slik at pilotering samtidig regnes som produksjonssetting av løsningen, og dermed innebærer at løsningen tas i bruk i saksbehandling. Det er imidlertid bare et fåtall utvalgte veiledere som skal delta i piloteringen, så det rulles ikke ut til veiledere generelt. Et sentralt formål med å pilotere løsningen, er å få tilbakemeldinger fra veiledere på om prediksjonene, samt forklaringene på prediksjonene, virker rimelige for veilederne. I tillegg skal pilotering sikre at modellene teamet bygger har tilstrekkelig presisjon, samt at prediksjonene reelt sett brukes som beslutningsstøtte.<sup>124</sup> Jeg påpeker for øvrig at NAV per mai 2021 ikke har igangsatt denne fasen, på

---

122 Se mer om dette i avsnitt 3.3.1.

123 Sml. avsnitt 2.2.

124 Se kort om dette i avsnitt 3.2 over.

grunn av manglende sikkerhet på det rettslige grunnlaget.<sup>125</sup> NAV befinner seg altså stadig i fase 2.

I figur 6 er det en pil fra fase 3 til fase 2. Dette skyldes at prediksjonene, ved testing, kan fungere dårligere i praksis enn det virket som ved utvikling. Dette oppdages for eksempel ved at veiledere melder om rare prediksjoner, eller at systemet melder om at prediksjonskraften til modellen går ned. I så fall gjenopptar teamet arbeidet med innhenting av data og videreutvikling av algoritmen. Utviklingsarbeidet er med andre ord planlagt som en smidig prosess, der testing og utvikling skal skje iterativt. I hvilken grad modellen på sikt skal være dynamisk, er etter min forståelse ikke bestemt.<sup>126</sup>

Jeg skal nå gjennomgå funnene mine. Funnene er plassert slik at de svarer til forskningsspørsmålene mine.<sup>127</sup> Jeg starter med å diskutere organisering av utviklingsarbeidet, ser deretter på den tekniske innretningen, og avslutter med en gjennomgang av rettslige problemstillinger NAV står overfor.

## 5.2 Organisering av utviklingsarbeidet

I forbindelse med utarbeidelse av forskningsspørsmålene mine, operasjonaliserte jeg «organisering» til å særlig omfatte mandat og formål for prosjektet, beslutningsprosesser, og organisering av arbeidskraft og kompetanse internt og eksternt. I tillegg ønsket jeg å se nærmere på rutinene for dokumentasjon. Jeg hadde med andre ord ambisjoner om å få en bred oversikt over hvordan utviklingsarbeidet foregår og forankres i organisasjonen. I gjennomgangen nedenfor, ser jeg på disse aspektene ved utviklingsarbeidet.

Innledningsvis er det verdt å påpeke at datagrunnlaget mitt er preget av at jeg kun har snakket med én person i teamet, det vil si en av data scientistene. Selv om data scientisten har fulgt prosjektet siden start,<sup>128</sup> og derfor sitter på veldig relevant informasjon, betyr det at dataene mine om organisering nødvendigvis primært gjenspeiler data scientistens forståelse av organisering. Informanten har også deltatt i utarbeidelse av mange av de interne dokumentene jeg bruker som supplerende kilder,<sup>129</sup> og det må antas at kildene jeg har brukt derfor heller ikke er helt uavhengige.<sup>130</sup> Når det gjelder forholdet mellom teamet og jurister

---

125 Se nærmere om dette i 4.4 under.

126 Sml. avsnitt 2.2.

127 Se avsnitt 1.2.

128 Se avsnitt 1.4.3.2.

129 Se avsnitt 1.4.2.

130 Jacobsen (2015) s. 231.

har jeg supplert informasjonen fra data scientisten med informasjon fra juristen jeg intervjuet. Totalt sett er det likevel grunn til å være oppmerksom på at gjennomgangen sannsynligvis er svært preget av data scientistens syn.

### 5.2.1 Teamet

Jeg begynner med å se nærmere på teamet som jobber med utvikling av beslutningstøtte i sykefraværsoppfølgingen. Det er et tverrfaglig team, men består likevel primært av software-utviklere. I tillegg består teamet av to data scientister, én designer, og to fagpersoner fra Arbeids- og tjenesteavdelingen hos NAV. Det er en av fagpersonene som er produkteier, og som dermed formelt sett har ansvar for å sikre at prosjektet resulterer i et produkt som har nytteverdi i saksbehandlingen.<sup>131</sup>

Data scientistene har etter min forståelse hovedansvar for den tekniske utformingen av selve prediksjonene, herunder både algoritmen og datasettene som skal brukes til å trene modeller. Jeg har kun intervjuet en av data scientistene i teamet, og har derfor begrenset kunnskap om hva de andre personene i teamet jobber med. Et hovedproblem er manglende datagrunnlag om fagpersonenes rolle i teamet.<sup>132</sup> Av e-post fra data scientisten, fremgikk det imidlertid at ingen av fagpersonene har juridisk bakgrunn, men at begge har jobbet som veiledere innenfor sykefraværsoppfølging tidligere. De har med andre ord praktisk erfaring med anvendelse av den aktuelle lovhjemmelen, og med den bredere saksbehandlingskonteksten profileringsystemet skal fungere innenfor.<sup>133</sup>

Innenfor teamet er det stor grad av dialog, og derfor mulig å ta uformelle, løpende beslutninger. De to data scientistene jobber spesielt tett, og har daglig kontakt rundt dataene og modellen. Teamet har ellers faste, ukentlige møter med utveksling av erfaringer og spørsmål i løpet av en arbeidsuke, og har også et fast møte annenhver uke for å vurdere fremgangen siden forrige møte. I praksis har dermed teamet flere arenaer til å diskutere og konkludere på problemstillinger den enkelte teamdeltaker møter på i sitt arbeid. Det må også forventes at en del vurderinger tas uformelt over e-post- og chatteløsninger.<sup>134</sup>

Beslutningene teamet tar internt omfatter for eksempel valg rundt prioriteringer av utviklingsoppgaver, samt definering og avgrensninger av informasjons-

---

131 Digitaliseringsdirektoratet (2019).

132 Jf. også avsnitt 1.4.3.2.

133 Se avsnitt 3.1.

134 Dette fremgikk også av den innledende samtalen med informanter. Ettersom samtalen var mindre formell, sendte jeg aldri et referat til sjekk av informantene, og datagrunnlaget er derfor mindre sikkert enn øvrig informasjon i denne framstillingen.

typer til bruk i løsningen.<sup>135</sup> De fleste beslutninger som tas i arbeidet er innom teamet til diskusjon, men enkelte beslutninger kan også tas av enkeltpersoner alene. For data scientistene gjelder dette for eksempel faglige vurderinger rundt hvordan datavask skal gjennomføres, i de tilfellene data scientistene ikke har grunn til å tenke at noen andre i teamet kjenner årsaken bak mangler i data-settene.

I intervju påpekte også data scientisten at formelle beslutningsprosesser typisk skjer i møte med jurister, fordi juristene gjerne har en mer formalisert arbeidsmetodikk. Juristen jeg intervjuet forklarte tilsvarende at rettslige vurderinger etter hennes erfaring gjerne dokumenteres. Vurderinger knyttet til personvernkonsekvensvurderinger, rettslig grunnlag, og pilotering av løsningen, er dermed mer formaliserte. Disse beslutningene involverer jurister, ledelsen, og/eller personvernombudet i NAV, og er derfor preget av flere meninger og uttalelser enn de interne diskusjonene i teamet.

Hovedvekten av kompetansegrunnlaget i teamet er kompetanse som sikrer et teknisk resultat med god kvalitet, både på den mer tradisjonelle systemutviklingsfronten (designeren og software-utviklerne), og på datafronten (data scientistene). Fagkompetanse som sikrer at produktet teamet utvikler er relevant for NAV er også representert gjennom de to fagpersonene i teamet. Til gjengjeld er det ingen jurister som deltar i teamet til daglig. En av forutsetningene for dette masterarbeidet, er at de personene som utvikler et system har avgjørende betydning for hvordan systemet innrettes.<sup>136</sup> Ettersom det ikke er noen jurister involvert direkte i teamet, er det derfor sentralt å se nærmere på hvordan jurister ellers er involvert, for å sikre at utviklingen skjer i samsvar med gjeldende rettsregler. Det blir desto mer interessant fordi beslutningene som tas med jurister, gjerne er mer formaliserte, og dermed sannsynligvis oftere dokumenterte.

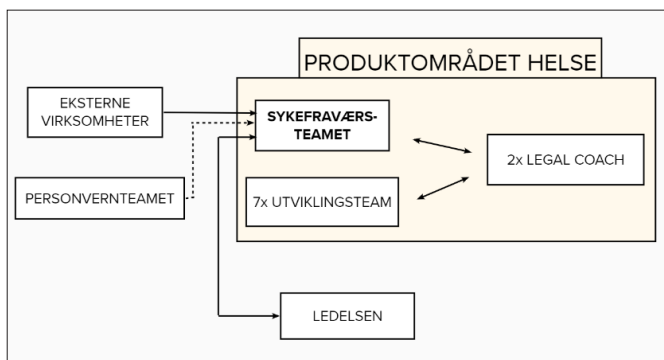
## **5.2.2 Forholdet mellom teamet og andre organisatoriske enheter**

Teamet som jobber med utvikling av profileringssystemet er den viktigste organisatoriske enheten for å forstå kompetansegrunnlaget som utnyttes i utviklingsarbeidet. Også andre organisatoriske enheter påvirker imidlertid arbeidet. I figur 7 nedenfor har jeg illustrert forholdet mellom relevante organisatoriske enheter. Sykefraværsteamet er uthevet i fet skrift i figuren, for å markere at dette teamet er kjernen i framstillingen.

---

135 Se mer om dette i avsnitt 4.3.1.2 under.

136 Se avsnitt 1.2.2.



Figur 7: Oversikt over organisatoriske enheter tilknyttet utviklingsarbeidet

Slik det fremgår av figur 7 befinner sykefraværsteamet seg innenfor produktområdet helse (heretter PO Helse). PO Helse er en organisatorisk enhet i NAV som består av åtte tverrfaglige team (7 i tillegg til sykefraværsteamet), som alle jobber med ulike systemløsninger innenfor helse i NAV. I tillegg jobber det to jurister innenfor produktområdet som legal coacher, som har spesialisert seg på rammeverk som gjelder for systemutvikling i NAV. Eksterne virksomheter, ledelsen, personvernteamet og juridisk seksjon er organisert utenfor produktområdet.

Det er piler mellom hver enhet som illustrerer relevante påvirkningsforhold. Jeg kommer til å diskutere de ulike påvirkningsforholdene i det følgende, med vekt på forholdet mellom sykefraværsteamet og ledelsen, eksterne virksomheter og legal coachene. Jeg påpeker i tillegg forholdet mellom legal coacher og øvrige team i PO Helse, men kun i den grad det er nødvendig for å forstå legal coachenes forhold til sykefraværsteamet. Pilen mellom personvernteamet og sykefraværsteamet er stiplet, fordi personvernteamet er oppløst.

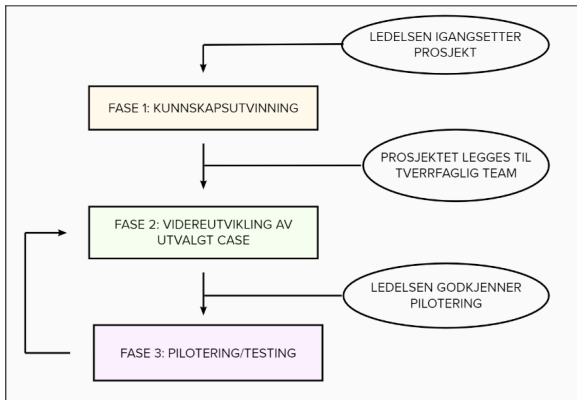
Juristen jeg intervjuet forklarte at det også jobber fire fagjurister i produktområdet, og at disse jobber direkte med teamene. Data scientisten sa istedenfor at teamet kun skulle forholde seg til legal coachene ved rettslige spørsmål. Etter som data scientisten jobber direkte i teamet, og har jobbet i produktområdet lengst, har jeg valgt å vektlegge data scientistens vurdering tyngst, og ikke skissert fagjuristene i figur 7. Det bør uansett legges til grunn at figur 7 neppe er en fullstendig oversikt over alle de organisatoriske enhetene som påvirker teamet. De enkelte team-medlemmene tilhører for eksempel egne fagavdelinger, (IT, Arbeids- og tjenesteavdelingen, etc.), og disse er ikke skissert. Heller ikke juridisk seksjon eller faggruppen for data science er skissert. Figuren inneholder

imidlertid de viktigste påvirkningsforholdene jeg har fått høre om gjennom intervjuer med informanter i NAV.

Jeg skal nå gå gjennom hvert enkelt påvirkningsforhold som er skissert i figur 7.

### 5.2.2.1 Teamet og ledelsen

Det første viktige påvirkningsforholdet å behandle er forholdet mellom teamet og ledelsen. Jeg har illustrert de påvirkningspunktene jeg har funnet mellom teamet og ledelsen i figur 8 nedenfor. Figuren tilsvarer figur 6, men inneholder enkelte viktige tidspunkter som illustrerer ledelsens involvering i utviklingsarbeidet. Disse er ført inn i figuren i sirklene til høyre.



Figur 8: Teamet og ledelsen

Slik det fremgår i figur 8 startet hele prosjektet som et initiativ fra ledelsen, nærmere bestemt fra fagledelsen i form av sykefraværskontoret. Sykefraværskontoret har ansvar for å gjennomføre Regjeringens sykefraværspolitik, og å gi faglige innspill for politikkutvikling til Arbeids- og sosialdepartementet.<sup>137</sup> De hadde hypoteser om sammenhenger i dataene NAV sitter på om sykefraværsoppfølging, og initierte derfor et samarbeid med AI-labben<sup>138</sup> i NAV, for å teste ut disse hypotesene. På dette tidspunktet var med andre ord fagledelsen veldig involvert i prosjektet.

<sup>137</sup> Karrierestart.no (2020).

<sup>138</sup> AI-labben er ikke lenger en organisatorisk enhet i NAV.



Ved overgangen til fase 2, ble prosjektet flyttet over i det tverrfaglige teamet. På dette punktet gikk prosjektet over til å bli styrt mer nedenfra og opp, og ledelsen involverte seg i mindre grad. Det må imidlertid antas at det konkrete prosjektet som ble valgt ut ved overgangen til fase 2 ble valgt ut i samsvar med ledelsen. Data scientisten har i intervjuer påpekt at det kan være behov for at ledelsen involverer seg i større grad igjen, for å få litt mer fortgang i arbeidet. Dette er etter min forståelse ikke som følge av en konkret tidsramme for prosjektet, men fordi data scientisten opplever at lite involvering av ledelsen gjør det vanskeligere å komme videre i prosjektet i det hele tatt.

For å kunne igangsette fase 3, må ledelsen godkjenne pilotering av løsningen. Ledelsen godkjenner imidlertid bare pilotering dersom det rettslige grunnlaget er på plass. På grunn av store usikkerheter på behandlingens rettslige grunnlag, har pilotering derfor ikke begynt.<sup>139</sup> Data scientisten påpekte også at det kanskje ville vært lettere for juristene å gi klarsignal for det rettslige grunnlag dersom de visste at ledelsen ville støtte dem.

#### **5.2.2.2 Teamet og interne jurister**

Det andre svært viktige påvirkningsforholdet mellom organisatoriske enheter, er forholdet mellom sykefraværsteamet og jurister. I figur 7 har jeg skissert de to aktuelle organisatoriske enhetene, legal coachene og personvernteamet. Data scientisten forklarte at sykefraværsteamet tidligere har trukket mye på personvernteamets kompetanse omkring kravene som følger av personvernforordningen. Personvernteamet ble imidlertid avviklet høsten 2020.<sup>140</sup> Nå skal teamet forholdet seg til legal coachene ved rettslige spørsmål, og jeg bruker derfor den resterende plassen til å behandle teamets forhold til legal coachene.

Legal coach-rollen er en egen type juridisk rolle, som skal jobbe tett på utviklingsmiljøene i NAV for å bidra inn i smidige prosesser.<sup>141</sup> Ideen bak legal coach-rollen er at teamet skal ha tilgang på jurister med spisskompetanse på rammeregler som gjelder for systemutviklingsprosjekter. Juristen jeg snakket med, forklarte at rollen kan være særlig utfordrende fordi legal coachene sitter i team med utviklere som ikke nødvendigvis har så mye forhold til lovgivers intensjon, eller andre sentrale rettslige prinsipper (som f.eks. maktfordelingsprinsippet). Da må legal coachene balansere ønsket om å være fremoverlent med å sikre at borgernes rettigheter blir ivaretatt. Legal coachene må også ta fortløpende, midlertidige vurderinger for å holde følge med de teknologiske endringene. Juristen påpekte at det er et ganske tydelig skille mellom legal coacher

---

139 Se mer om dette i avsnitt 4.4.

140 Informanten nevnte at et alternativt team, bestående av interne jurister, kanskje skulle startes.

141 Grønning (2019).

og mer tradisjonelle jurister, der legal coachene ser behovet for (regelverks)ending i større grad enn tradisjonelle jurister. De tradisjonelle juristene er på sin side mer rigide i regelanvendelsen.

Det er enkelte utfordringer knyttet til forholdet mellom legal coachene og teamet. Den første utfordringen er at organiseringsmodellen forutsetter at teamene selv vet når de trenger å involvere en jurist. Det er ikke gitt at teknologer vet når de må koble på jurister. Juristen påpekte imidlertid at det er helt sentralt at teamet stoler på at legal coachene kan hjelpe, og at legal coachene stoler på at teamet tar kontakt ved behov, for at organiseringsmodellen skal fungere. Korona-situasjonen med hjemmekontor har gjort samarbeidet mer utfordrende. Når juristen og teamene jobber fysisk i de samme lokalene, er det både lavere terskel for at teamene stiller spørsmål til legal coachene, og at legal coachene fanger opp relevante samtaler.

I tillegg har NAV noen verktøy som kan hjelpe både legal coachene og teamet til å fange opp når juridisk kompetanse er nødvendig i et utviklingsarbeid. Det første er en liste over etterlevelseskrav som NAV forvalter, og som inneholder alle rettslige krav som gjelder for NAVs saksbehandlingssystemer.<sup>142</sup> Etterlevelseskrav er med andre ord synonymt med det jeg i denne framstillingen kaller «rettslig ramme».<sup>143</sup> Ved avsjekking av krav fra denne listen, kan det bli tydelig at det er enkelte momenter ved løsningen som burde vært diskutert med en legal coach. I tillegg er arbeidet med personvernkonsekvenser en omfattende prosess, der manglende juridiske vurderinger også kan fanges opp. NAV har utviklet en fast mal for gjennomføring av personvernkonsekvensvurderinger, som teamene gjennomgår og fyller ut med jurister. Konsekvensvurderingen skal også kommenteres av personvernombudet i NAV, som tilføyer kommentarer. Her vil med andre ord rettslige utfordringer bli tydeligere for teamet.

Den andre utfordringen med forholdet mellom legal coacher og teamet, er at organiseringsmodellen forutsetter at legal coachene har tid til å hjelpe teamet. I PO Helse jobber det to legal coacher i 80%-stillinger. Mens enkelte produktområder er nede i ett produktteam, består PO Helse av hele åtte team, se figur 7. Av disse åtte teamene er juridisk bistand til det teamet som jobber med ny sykepengeløsning en hovedprioritet for juristene, fordi ny sykepengeløsning er en politisk satsning fra Arbeids- og sosialdepartementet.<sup>144</sup> Arbeidet med sykefraværsteamet er istedenfor en form for overskuddsaktivitet for legal coachene,

---

142 Datakatalogen (u.å.).

143 Se avsnitt 3.3.2.

144 Prop. 1 S (2020–2021) s. 130.

som i praksis innebærer at det blir begrenset tid til å vurdere de mange rettslige problemstillingene.<sup>145</sup>

Data scientisten forklarte at teamet har vært tydelige overfor juristene på at de ikke kan ta noen juridiske vurderinger på egenhånd, og mangelen på juridisk kompetanse i det daglige arbeidet er derfor spesielt merkbart. Data scientisten forklarte at teamet i praksis må skrive opp spørsmål de ønsker å drøfte med legal coacher i en backlog, og drøfte flere problemstillinger samtidig når legal coachene først har kapasitet. Samarbeidet med legal coacher skjer med andre ord på ujevn basis, og virker til tider å avhenge mer av kapasitet enn behov. Samarbeidet er til gjengjeld veldig tett i arbeidsintensive faser, for eksempel i forbindelse med utarbeidelse av personvernkonsekvensvurderinger, eller i forbindelse med arbeidet med den regulatoriske sandkassen til Datatilsynet.<sup>146</sup>

### 5.2.2.3 Teamet og eksterne partnere

I tillegg til den interne organisering i NAV er det enkelte eksterne virksomheter og personer jeg ønsker å trekke fram, se figur 7. Disse eksterne aktørene er viktige å kommentere av to grunner. For det første jobber de direkte med teamet i forbindelse med utvikling av profileringsystemet, og det må antas at de derfor påvirker systeminnretningen. For det andre trekkes eksterne virksomheters involvering i systemutviklingsarbeidet ofte fram i litteraturen som særlig risikabelt for demokratiske verdier.<sup>147</sup> Dersom eksterne virksomheter for eksempel utvikler kode eller andre tillegg til systemløsningen, kan forvaltningsvirksomheten miste kontroll over det tekniske innholdet i saksbehandlingssystemet som utvikles. Dette blir enda mer utfordrende dersom den eksterne virksomheten i praksis bestemmer hvordan myndighetsutøvelsen til forvaltningsorganet skal utføres.

Big Insight, ved Norsk Regnesentral, bidrar med å utvikle konkrete, tekniske metoder for å kunne forklare prediksjoner i maskinlæringsmodellen.<sup>148</sup> Dette er den eneste eksterne virksomheten som skriver kode som brukes direkte i løsningen til NAV. Jeg har ikke diskutert samarbeidet inngående med informanter i NAV, og er derfor forsiktig med å trekke konklusjoner om hva samarbeidet faktisk innebærer.

NAV har også et nyoppstartet samarbeid med forskningsgruppen «AI4Users». Samarbeidet har ikke kommet langt nok til at data scientisten jeg intervjuet hadde oversikt over hva samarbeidet ville innebære. AI4Users jobber imidlertid

145 Se nærmere om disse i avsnitt 4.4.

146 Se mer om dette i avsnitt 4.2.2.3 under.

147 Se Schartum (2018) s. 77 flg. og Redden (2018) s. 9.

148 Big Insight (u.å.)

med å finne ut av hvilke forklaringer brukere av maskinlæringsmodeller trenger,<sup>149</sup> og det er grunn til å tro at noe tilsvarende er fokus for samarbeidet med NAV. Forskerne skal trekke på erfaringene og meningene til «vanlige brukere».<sup>150</sup> Hvorvidt disse er å regne som veiledere i NAV eller de sykmeldte vet jeg ikke.

Sykefraværsteamet involverer Frischsenteret. Det er et forskningsmiljø som jobber med effektanalyser, og som kan bidra med veiledning rundt hvordan teamet skal organisere forsøk. Dette gjelder for eksempel pilotering av løsningen, men også andre, mindre endringer i løsningen. Inntrykket mitt er at denne kompetansen kan bidra til å gi teamet bedre forståelse for systemets påvirkning på veiledere, men at det ikke kommer til å ha noen direkte effekt på innholdet i systemet.

NAV involverer også noen eksterne aktører med juridisk kompetanse. NAV, ved sykefraværsteamet, deltar for eksempel i Datatilsynets regulatoriske sandkassen for kunstig intelligens, der de diskuterer kravene til forklarbarhet, rettferdighet og rettslig grunnlag.<sup>151</sup> NAV og Det juridiske fakultet i Oslo har også så smått diskutert et samarbeid, som skal omfatte problemstillinger knyttet til forklarbarhet.<sup>152</sup> Disse aktørene kan bidra med inngående juridiske drøftelser og avklaringer, som kan påvirke den tekniske systeminnretningen.

Sykefraværsteamet har vurdert å involvere det sentrale brukerutvalget til NAV.<sup>153</sup> Det sentrale brukerutvalget består av representanter for personbrukerne, arbeidsgivere, og KS, samt utvalgte ledere i NAV.<sup>154</sup> Formålet med et møte med disse representantene er å få tilbakemeldinger på utviklingsarbeider som særlig påvirker brukergruppene til NAV.<sup>155</sup> Møtet skulle egentlig gjennomføres i 2020, men ble satt på vent da korona inntraff. Dermed har teamet ikke gjennomført noe møte med brukerutvalget når dette forskningsarbeidet ferdigstilles i mai 2021. Data scientisten påpekte at brukerutvalget uansett er en «enkel» form for brukerinvolvering, fordi det kun involverer representanter for de sykmeldte, og at det beste ville vært å involvere de sykmeldte direkte. Det forutsetter imidlertid mye planlegging for å sikre at teamet har en god gjennomføringsplan, og teamet planlegger foreløpig ingen slik direkte involvering av de sykmeldte.

---

149 AI4Users (u.å.1).

150 AI4Users (u.å.1).

151 Datatilsynet (u.å.1).

152 Digitaliseringsdirektoratet (2021).

153 Jf. også personvernforordningen artikkel 35 nr. 7.

154 NAV (2020a).

155 NAV (2015).

Avslutningsvis nevner jeg at alle de eksterne virksomhetene som er involvert per dags dato, unntatt Frischsenteret, jobber enten utelukkende med forklarbarhet eller delvis med forklarbarhet. Dette er med andre ord et populært forsknings-tema, både teknisk, juridisk og praktisk, og virker som det er den delen av systeminnretningen som foreløpig ser ut til å bli mest påvirket av eksterne virksomheter. Når det gjelder Big Insight, som utvikler kode som brukes i løsningen, mener jeg at involveringen av Big Insight sannsynligvis ikke innebærer noen stor demokratisk utfordring. Det skyldes særlig at forskningsgruppen etter min forståelse lager kildekode som gjelder et snevert aspekt ved løsningen (forklarbarhet), samtidig som de er et forskningsinitiativ uten særlige økonomiske insentiver. Det er dermed liten grunn til å tro at kildekode (1) ikke kan undersøkes og forstås av teknologer i NAV eller (2) påvirker de delene av modellen som innebærer saksbehandling. Jeg påpeker også at NAV generelt virker interesserte i å diskutere løsninger og problemstillinger med eksterne virksomheter, idet de deltar og diskuterer problemstillinger i en rekke offentlige fora.<sup>156</sup>

### **5.2.3 Dokumentasjon av teknisk og rettslig innhold**

Dokumentasjon kan betegne mye, og kan skje på mange nivåer. I denne framstillingen har jeg primært forsøkt å kartlegge de dokumentene som (1) gir innsikt i hvordan profileringsystemet fungerer og er bygget opp, og (2) gir innsikt i de rettslige vurderingene NAV har tatt. Førstnevnte innebærer forklaringer av teknisk/statistisk art, mens sistnevnte kan forstås som en del av den rettslige systemdokumentasjon av løsningen.<sup>157</sup>

Data scientisten forklarte at den mest fullstendige dokumentasjonen av systeminnretningen slik den ser ut per dags dato er kildekode, samt kommentarer i kildekode. I tillegg fremgikk det av intervjuet at det sannsynligvis ikke er mulig for noen andre enn teamet som jobber med systemet å forstå denne dokumentasjonen. Generelt nedprioriteres dokumentasjon av selve løsningen til fordel for å lage en fungerende løsning. Samtidig forklarte data scientisten at han begynner å se et større behov for å dokumentere, for å kunne forklare enkelte deler av løsningen til jurister.

De delene av løsningen som teamet prioriterer å dokumentere, er gjerne knyttet til rettslige vurderinger. Personvernkonsekvensvurderingene er de mest omfattende dokumentene. De er skrevet i tre runder, i forbindelse med de forskjellige fasene av utviklingsarbeidet til NAV.<sup>158</sup> Disse dokumentene er laget av teamet selv, andre jurister, og er også kommentert av personvernombudet. Ettersom

<sup>156</sup> AICAST (2019), LØRN.TECH (2019), Datatilsynet (2021), Teknologirådet (2021) og Digitaliseringsdirektoratet (2021).

<sup>157</sup> Jf. Schartum (2018) s. 278.

<sup>158</sup> Se også figur 6.

kommentarene til vurderingene blir registrert og lagret, reflekterer de interne diskusjoner, ikke bare den offisielle konklusjonen til NAV. I tilknytning til personvernkonsekvensvurderingene har teamet også utarbeidet en oversikt over informasjonselementer som brukes i løsningen, og en illustrasjon av den tekniske dataflyten.

I tillegg til vurderingene som fremgår i personvernkonsekvensvurderingene, er vurderingene av det rettslige grunnlaget til løsningen dokumentert i et eget dokument. Det samme gjelder oversikten over etterlevelseskrav. Begge disse dokumentene er under arbeid, og oppdateres fortløpende etter hvert som NAV gjør nye vurderinger.

Jeg forutsetter at det kan være mangler i datagrunnlaget mitt, og at det kan finnes ytterligere dokumentasjon hos NAV enn det jeg har greid å finne, særlig når det gjelder rettslige vurderinger. Framstillingen av dokumentasjonen bør derfor ikke forstås som en fullstendig oversikt over den dokumentasjonen som finnes, men heller som en gjengivelse av de dokumentene jeg har greid å finne i løpet av undersøkelsen.<sup>159</sup> Det er også verdt å påpeke at arbeidet med prediksjoner av sykefraværslengde er et prosjekt som fortsatt er under utvikling. Dokumentasjonen som finnes per dags dato er derfor mer begrenset enn dokumentasjonen som vil finnes av løsningen senere. Det planlegges flere dokumenter som ikke enda er ferdigstilt, for eksempel en oversikt over informasjon som skal gis til veiledere, arbeidet som er gjort med dataminimering, og hvordan datakvalitetssjekker gjennomføres.

### **5.3 Foreløpig teknisk innretning av profileringsystemet**

Jeg skal nå gjennomgå funnene mine av den tekniske innretningen av profileringsystemet som NAV bygger, med vekt på de tekniske/praktiske problemstillingene data scientistene står overfor. Jeg starter med å gå gjennom hvilke datakilder sykefraværsteamet bruker, og ser deretter på de konkrete dataene som inngår i løsningen. Til slutt ser jeg på algoritmen som skal analysere dataene, og de konkrete prediksjonene som er resultatet av analysene. Disse setter jeg i profileringskonteksten jeg jobber ut fra. Innledningsvis påpeker jeg også at den tekniske innretningen jeg går gjennom er under arbeid, og derfor midlertidig.

Det er verdt å påpeke at det finnes svært lite dokumentasjon av den tekniske innretningen, og at jeg mye av det jeg presenterer i avsnitt 4.3 derfor er basert utelukkende på informasjon fra intervjuer med data scientisten. Unntaket er en

---

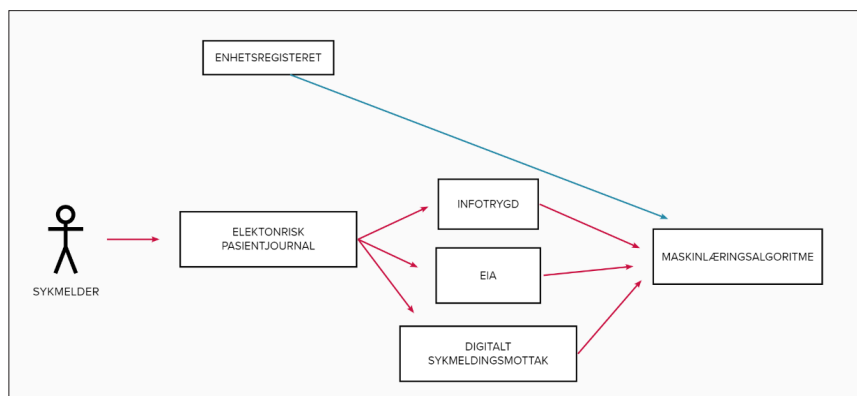
<sup>159</sup> Se også avsnitt 1.4.2.

illustrasjon av den interne dataflyten, og en oversikt over informasjonselementer.<sup>160</sup> Data scientisten har imidlertid en sentral rolle i utviklingsarbeidet,<sup>161</sup> og det må antas at dataene fra intervjuet derfor er veldig relevante. Jeg har ingen grunn til å tro at informanten har gitt meg feilaktig informasjon. Mangler ved både dataflyten og oversikten over informasjonselementer ble for eksempel påpekt og supplert av data scientisten i intervjuene. Det er imidlertid en mulighet at jeg har misforstått noen av poengene til informanten.<sup>162</sup> Ettersom jeg fikk en godkjenning på referatet,<sup>163</sup> antar jeg at det i så fall ikke er snakk om alvorlige feil.

### 5.3.1 Dataene

#### 5.3.1.1 Interne og eksterne datakilder

I figur 9 har jeg illustrert datakildene som brukes til å trene maskinlæringsalgoritmen. De røde pilene representerer flyt av sykmeldinger, mens den blå pilen representerer flyt av data om arbeidsgiver.



Figur 9: Dataflyten

Slik det fremgår av figur 9, har jeg startet oversikten over dataløpet til sykmeldingene ved sykmelder. Jeg kunne også startet dataløpet ved den sykmeldte selv, som er den originale kilden til de dataene sykmelder registrerer i sykmeldingene.<sup>164</sup> Likevel starter jeg dataløpet ved sykmelder, fordi det er sykmelderne som registrerer innholdet i sykmeldingene som brukes av sykmeldingsteamet. Jeg

160 Se avsnitt 4.2.3.

161 Se avsnitt 1.4.3.2.

162 Se avsnitt 1.4.3.3.

163 Se avsnitt 1.4.3.1.

164 Sml. Kaufmann (2020) s. 236.

har heller ingen relevante data til å drøfte overgangen fra den sykmeldtes faktiske situasjon til de dataene som er registrert om vedkommende. Jeg nøyer meg derfor med å påpeke at dataene som registreres i sykmeldingene som NAV mottar er sykmelders tolkning av den sykmeldtes helsesituasjon.

Uttrekk av data fra sykemeldingene gir flere utfordringer. Den første gjelder registrering av sykmeldinger i elektroniske pasientjournaler (heretter EPJ-systemer). Slik det fremgår av figur 9, rapporterer sykemelderne inn sykmeldinger via EPJ-systemer som er knyttet opp mot det digitale sykmeldingsmottaket til NAV.<sup>165</sup> Disse EPJ-systemene har ulike grensesnitt for registrering, slik at de for eksempel inneholder varierende grad av fritekstfelt. I praksis er det derfor forskjell på innrapporteringsrutinene for sykmeldinger, som gjør at sykmeldingene inneholder ulike typer feil. Ifølge data scientisten er typiske feil at sykmelder har blandet sammen feltene for registrering av diagnosebeskrivelse og diagnosekode. Dette er likevel feil som gjelder et mindretall av sykmeldinger.

Figur 9 viser at det er tre interne systemer for mottak av sykmeldinger: Infotrygd, EIA og Digitalt sykmeldingsmottak. I dag brukes kun Digitalt sykmeldingsmottak for registrering av sykmeldinger. Maskinlæringsalgoritmen til sykefraværsteamet forutsetter imidlertid store mengder data, hvilket i praksis innebærer data fra de siste 10 årene. I løpet av disse årene, har NAV brukt tre forskjellige systemer for sykefraværmottak. Disse systemene er organisert på ulikt vis, og data scientisten trakk fram behovet for å kartlegge systemene for å kunne bruke dataene.<sup>166</sup> Det innebærer for eksempel å forstå forholdet mellom kolonner og rader i systemene, samt filformatene til dataene. Dette er spesielt utfordrende i Infotrygd, som er et 40 år gammelt system under utfasing. Fordi de fleste utviklerne som jobbet med utvikling av dette systemet har sluttet i NAV, har sykefraværsteamet måttet trekke på kunnskapen fra veiledere som har brukt systemet for å kunne unytte dataene. Teamet har ikke greid å finne systemdokumentasjon.

Den siste datakilden jeg har ført opp i figuren er Enhetsregister. Det kan være grunn til å gjøre et skille mellom interne og eksterne datakilder, fordi virksomheten selv gjerne har mer kontroll over interne datakilder. Det gjelder både kvalitet på dataene, men også over hvordan begrepene avgrenses og defineres, og om innholdet i begrepene endrer seg.<sup>167</sup> I tillegg kan ikke virksomheten bestemme hvilke data som skal samles inn til de eksterne datakildene.<sup>168</sup> Enhetsregisteret er imidlertid en nasjonal felleskomponent, som er laget for at andre

<sup>165</sup> Prop. 1 S (2020–2021) s. 84.

<sup>166</sup> Gjennomføres både i fase 1 og fase 2, jf. figur 6.

<sup>167</sup> Sml. f.eks. Digitaliseringsdirektoratets Seksjon for Informasjonsforvaltning (2019).

<sup>168</sup> Schartum (2018) s. 111.



virksomheter skal kunne utnytte dataene i registeret.<sup>169</sup> Jeg forventet derfor ikke at teamet møte på noen særlige utfordringer ved bruk av denne datakilden, til tross for at den var ekstern. Data scientisten bekreftet dette inntrykket, og forklarte at den eneste hindringen for bruk av disse dataene bestod i å forstå hva informasjonstypene betyr. Ettersom begrepene er eksplisitt definerte i API-dokumentasjonen til Enhetsregisteret,<sup>170</sup> var ikke dette en stor problemstilling.

### 5.3.1.2 Informasjonstyper og informasjonelementer

Datakildene gir teamet tilgang til sykmeldinger om de sykmeldte, samt noe informasjon om arbeidsgivere. Utgangspunktet er altså at teamet kan benytte alle data om alle sykefraværsløp som er registrert hos NAV, med utgangspunkt i datakildene jeg har gjennomgått over. For å sikre at algoritmen trenes på relevante data, har teamet imidlertid definert to grunnleggende egenskaper ved relevante data: dataene må være knyttet til et sykefraværsløp med en lengde på 17 uker eller mer, og dataene må ha vært tilgjengelig for veileder ved uke 17. Jeg skal kort skissere begrunnelsene for dette, rett nedenfor.

Den første egenskapen for at et sykefraværsløp skal inkluderes i treningsdataene til sykefraværsteamet følger naturlig av formålet med prediksjonene, nemlig å være beslutningsstøtte til veiledere for å vurdere personer som allerede er sykmeldt 17 uker. For at maskinlæringsalgoritmer skal kunne trene gode modeller, må treningsdataene representere den samme typen personer som modellen deretter skal brukes til å predikere.<sup>171</sup> I følge NAVs egne beregninger, gjengitt i personvernkonsekvensvurderingene, innebar det i 2018 omtrent 9% av alle sykefraværstilfeller som har blitt innrapportert til NAV.

Den andre egenskapen som begrenser aktuelle sykefraværsløp er at dataene må ha vært tilgjengelige ved 17 ukers sykefravær. Dersom teamet trekker på data som ikke var tilgjengelig før etter uke 17, trenes modellen på et mer omfattende datagrunnlag enn det modellen vil ha tilgang på i reelle sykefraværstilfeller som ligger til saksbehandling. Dette utsetter modellen for overtilpasning, altså at modellen fungerer bedre på treningsdata enn ekte data.<sup>172</sup> Data scientisten jeg intervjuet trakk fram dette som en spesielt stor utfordring i de historiske sykmeldingsløpene, fordi sykmeldere kan overskrive dateringer av sykmeldinger manuelt, slik at det er uklart hvilke meldinger som var tilgjengelig på hvilket tidspunkt.

---

169 Brønnøysundregistrene (2020).

170 Sml. Brønnøysundregistrene (2021).

171 Berk (2019) s. 22.

172 Russell (2021) s. 1210.

Når omfanget av dataene som brukes til trening av sykefraværsløp er begrenset av de grunnleggende egenskapene som nevnt over, kan teamet begynne å hente ut data knyttet til aktuelle sykefraværsløp. Sykmeldingene er mest omfangsrrike, og blir oppdatert flere ganger i løpet av ett enkeltsykefraværsløp. Hvert sykefraværsløp er dermed knyttet til flere sykmeldinger. Sykmeldingene består av 13 overskrifter hver, med underoverskrifter. Det er dermed mulig å trekke ut svært mye forskjellig informasjon fra disse dokumentene. Det gjelder for eksempel den sykmeldtes situasjon, som diagnose, sykmeldingsgrad, og friskmeldingsplan. Av sykmeldingene fremgår også noe informasjon om sykmelder, som for eksempel yrke (f.eks. allmennlege, psykolog eller kiropraktor). Det er også mulig å utlede sykmeldingspraksis, ved å følge samme sykmelder over flere sykmeldingsløp.

Oversikten over dataene som samles inn om de sykmeldte er definert på to nivåer: som informasjonselementer og som informasjonstyper. Informasjonselementene er overordnede, brede kategoriseringer, mens informasjonstypene er helt konkrete aspekter ved den sykmeldte som kan utledes fra informasjonselementene.<sup>173</sup> «Data» er verdiene til slike informasjonstyper. Til sammenlikning er oversikten over informasjonselementene en liste på 13 elementer, mens informasjonstypene kan favne opp til 300 unike, om enn noe overlappende, kolonner. Informasjonselementene er listet opp som ledd i dokumentasjon av modellen, og brukes i forbindelse med utredningen av personvernkonsekvenser. Informasjonstypene er bare tilgjengelig ved å se direkte på kildekode. Som nevnt i kapittel 1, har jeg kun fått tilgang til oversikten over informasjonselementene.

Dataene som hentes ut fra sykmeldingene må deretter vaskes, defineres, og avgrenses. NAV har for eksempel ingen oversikt over når et sykefraværsløp er avsluttet; lengden må derfor bestemmes basert på øvrige data fra sykmeldingene. Dermed er dette et tydelig eksempel på at data scientistene ikke bare henter inn data, men også *former* data til det angitte formålet.<sup>174</sup> Data scientistene former både inngangsverdier og utgangsverdier i testsettene.<sup>175</sup> Inngangsverdier er de uavhengige variablene som skal brukes til å predikere noe, som for eksempel diagnose, alder, etc. Utgangsverdier er «svaret» på prediksjonen, det vil si sykefraværslengde i denne sammenheng. Jeg skal kort skissere praktiske problemstillinger knyttet til begge.

---

173 Jeg har hentet begrepet «informasjonselementer» fra NAV, mens «informasjonstyper» er mitt begrep. Sistnevnte ble blant annet referert til som «variabler» og «kolonner» i intervjuene, mens jeg har samlet referansene i ett begrep.

174 Muller (2019) s. 4.

175 Sml. Russell (2021) s. 1205.

Utgangsverdiene brukes til å vurdere prediksjonskraften til treningsdataene, og påvirker dermed hvordan alle de øvrige variablene i datasettet vektet. Data scientisten forklarte at en av de viktigste oppgavene teamet gjør for å sikre gode prediksjoner, er å sikre at denne verdien er korrekt. Dessverre er det også den mest kompliserte verdien for teamet å sammenstille. Det er flere grunner til det, men det skyldes særlig at dataløpet er rotete, som følge av at sykmelderne har lav terskel for å gjøre endringer i sykmeldinger. Meldingene som erstatter hverandre kan inneholde motstridende informasjon om sykefraværet til den sykmeldte. Det er utfordrende for sykefraværsteamet å lage kode som kan håndtere slike motstridende meldinger på en god måte, uten å for eksempel konkludere på friskmelding for tidlig eller for sent. I tillegg har sykemelderne mulighet til å manuelt overskrive dateringer av sykmeldinger som registreres hos NAV. Dette er også noe sykemelderne gjør i stor grad. Da kan det se ut som et sykefravær ble avsluttet på et tidligere eller et senere tidspunkt enn det som faktisk var tilfelle.

Når det gjelder inngangsverdier til algoritmen, mener jeg det er grunn til å tro at enkelte av disse er «ferdige», slik at de ikke formes noe særlig av data scientis-tene.<sup>176</sup> Det gjelder for eksempel verdiene fra Enhetsregisteret, som allerede er definerte og avgrensede informasjonstyper. Ellers beskrev data scientisten jeg intervjuet at han og teamet bruker fantasien og kvalifisert gjetning for å finne fram til gode informasjonstyper å bruke i modellen. Arbeidet med å definere og lage avgrensede og relevante informasjonstyper skjer ved å prøve og feile. Et eksempel data scientisten trakk fram, var avgrensning av verdien «jobb». En sykmeldt kan for eksempel tenkes å ha «to og en halv» jobb. Hvorvidt dette skal rundes opp eller ned, eller eventuelt forbli «to og en halv», er en type problemstilling teamet står overfor jevnlig.

Data scientisten forklarte at øvrige informasjonstyper enn utgangsverdien er mindre avgjørende for den endelige løsningen enn avgrensninger av sykefraværslengden. Dette er på mange måter åpenbart: for å sikre et best mulig produkt er det helt grunnleggende å vite at «fasiten» er riktig. Samtidig påpekte informanten at det er viktig at den sykmeldte kan kjenne seg igjen i informasjonen NAV har samlet inn i form av inngangsverdier. I tillegg gir mer presise inngangsverdier gjerne mer presise prediksjoner.

Ettersom mye av arbeidet med modellen så langt har handlet om å vurdere hvilke informasjonstyper teamet kan få tilgang til, ved å hente ut og strukturere data fra sykmeldinger, har sykefraværsteamet enda ikke gjennomført noen systematisk vurdering av hvilke informasjonstyper som kan fjernes fra en «endelig

---

176 Muller (2019) s. 5.

versjon» av maskinlæringsmodellen.<sup>177</sup> Det finnes dermed heller ingen dokumenterte vurderinger rundt hvilke informasjonstyper som er valgt vekk, eller hva som er begrunnelsen for at de er valgt vekk.

Til tross for at det ikke er valgt vekk informasjonstyper enda, er det helt klart at både inngangsverdier og utgangsverdier formes av data scientistene og teamet for øvrig. Av spesielt stor betydning er utformingen av utgangsverdiene, som påvirker hvordan alle øvrige informasjonstyper utnyttes i modellen. Utgangsverdien kan bli feil på to måter: utgangsverdien kan ende opp med å ikke gjenspeile lengden på sykefraværsløpene slik de er registrert hos NAV (av grunner jeg har skissert rett over), eller de kan ende opp med å ikke gjenspeile de reelle sykefraværsløpene.

Grunnen til at jeg trekker frem manglende gjenspeiling av reelle sykefraværsløp er at spørsmålet korresponderer med debatten rundt bruk av registrerte arrestasjoner for å angi kriminalitet i justissystemer i USA.<sup>178</sup> Det kan tenkes at det er visse strukturelle forskjeller i rapportering av sykefraværslengde i sykmeldingene mellom kvinner og menn, eller mellom ulike etniske grupperinger i Norge. NAV vet for eksempel at sykefraværsløpet gjerne er lengre for kvinner enn for menn.<sup>179</sup> Da er det en risiko for at prediksjonene blir skjeve basert på beskyttede kategorier. Jeg har ikke diskutert dette inngående med informanter i NAV, så jeg nøyer meg med å påpeke denne sammenhengen.

### 5.3.2 Analyse av dataene

Dataene som hentes ut fra datakildene skal deretter analyseres ved hjelp av en maskinlæringsalgoritme, for å predikere sykefraværslengde. Det finnes en rekke forskjellige maskinlæringsalgoritmer, med hver sine fordeler og ulemper. Ettersom NAV lager komplette treningsdatasett, med både inngangs- og utgangsverdier, er algoritmer innenfor overvåket læring mest aktuelle.<sup>180</sup>

Sykefraværsteamet har ikke gjort et systematisk arbeid med valg av algoritme enda. Enn så lenge er det *gradient boosting* som brukes for å analysere dataene. Gradient boosting regnes generelt som en algoritme som fungerer godt til å predikere både klassifikasjoner og regresjon når input-dataene er formaliserte, slik tilfellet er for sykefraværsteamet.<sup>181</sup> Ulempen er imidlertid at gradient boosting er en såkalt «black box»-algoritme, hvilket er utfordrende i et forklarbarhetsper-

177 Se mer om dette i avsnitt 4.4.3.

178 Se f.eks. Lum (2016).

179 Digitaliseringsdirektoratet (2021).

180 Russell (2021) s. 1205.

181 Ibid. s. 1297.

spektiv.<sup>182</sup> Algoritmevalget har ikke hatt betydning for utformingen av løsningen ellers, så foreløpig er det uproblematisk å bytte algoritme. Jeg ser derfor ikke nærmere på betydningen valget av algoritmen har for den øvrige systeminnretningen.

Et sentralt spørsmål er hvorvidt algoritmen skal trene modeller på dataene alene, eller om utviklere i tillegg skal tilpasse modellen til ekspertkunnskaper.<sup>183</sup> Data scientisten jeg intervjuet forklarte at endringer i sykmeldingsgrad som går stabilt nedover gjerne trekker sykefraværslengden ned, mens ustabile endringer i sykmeldingsgrad trekker sykefraværslengden opp, også når den siste rapporterte sykmeldingsgraden er lav. Dersom teamet kun måler sykmeldingsgraden fra den første sykmeldingen mot sykmeldingsgraden i den siste sykmeldingen, fanger ikke algoritmen opp slike svingninger. Arbeidet med å gjenkjenne slike deler av datagrunnlaget, og tilpasse algoritmen deretter, har skjedd i samarbeid med veiledere som kjenner til typiske sammenhenger i datagrunnlaget. Sykefraværsteamet kombinerer dermed ekspertkunnskaper og data, istedenfor å bare la dataene «snakke for seg selv».<sup>184</sup>

Resultatet av analysene basert på valgene over er en maskinlæringsmodell, som inneholder «regler» for å predikere nye tilfeller.<sup>185</sup> Maskinlæringsmodellen predikerer sykefraværslengde, inntil maksimalt 365 dager. Beslutningsstøtten veileder mottar, er likevel en prediksjon av sykmelding ved et konkret angitt tidspunkt, som for eksempel kan være uke 26, uke 28, eller uke 36.<sup>186</sup> Dersom prediksjonen av sykefraværslengde for eksempel tilsier at den sykmeldte vil forbli sykmeldt i 130 dager til, kan denne prediksjonen prosesseres for å vurdere om den sykmeldte fortsatt vil være sykmeldt ved uke 26 (63 dager).

Bakgrunnen for valget om prediksjon av lengde over tidspunkt, er at konkrete tidspunkter er kunstig satt av NAV. En prediksjon av sykefraværslengde er dermed mer robust, ettersom en prediksjon av sykefraværslengde kan brukes som grunnlag for å vurdere sykefraværet til den sykmeldte ved ethvert konkret tidspunkt NAV måtte ønske å predikere.

Det er verdt å merke seg at modellens prediksjonskraft måles etter modellens evne til å predikere sykefravær ved et angitt tidspunkt, ikke evnen til å predikere sykefraværslengde. Selv om evnen til å predikere lengde kan korrelere med evnen til å predikere sykefravær ved et angitt tidspunkt, er ikke disse nødven-

182 Se kort om dette i avsnitt 2.2 over, og nærmere drøftelser i avsnitt 4.4.3 nedenfor.

183 Kaufmann (2017) s. 154.

184 Sml. avsnitt 2.1.

185 Sml. avsnitt 2.1.

186 Jf. gjennomgangen av saksbehandlingskonteksten, avsnitt 3.1.

digvis sammenfallende. Som beslutningsstøtte for NAVs veiledere, er det imidlertid modellens evne til å predikere ved et angitt tidspunkt som er det essensielle. Det betyr at modellens evne til å predikere sykefraværestidspunktet er gjenstand for diverse tester, for eksempel av nøyaktighet, eller antall falske negative (FN) og falske positive (FP) prediksjoner. Nøyaktighet og FN/FP kan også vurderes på tvers av grupper,<sup>187</sup> for å teste modellen etter rettferdighetsmetriker.<sup>188</sup>

I en profileringskontekst, er det denne endelige vurderingen av hvilke egenskaper om tilsier at den sykmeldte vil forbli sykmeldt ved et angitt tidspunkt som er profilen.<sup>189</sup> Når modellen predikerer hvorvidt noen kommer til å forbli sykmeldte eller ikke, er det altså reglene som er nedfelt i modellen som tilsier hvilke tilfeller som skal falle utenfor eller innenfor profilen. Hvert steg jeg har gjennomgått over, er steg som er nødvendige for å lage denne profilen.

## 5.4 Rettslige vurderinger som påvirker utviklingsarbeidet

Jeg skal nå gjennomgå NAVs foreløpige juridiske vurderinger rundt utvikling av profileringssystemet, og hvordan de har påvirket den tekniske innretningen til profileringsystemet. Det sentrale i gjennomgangen er hvordan rettsreglene påvirker profileringsystemet. Jeg understreker innledningsvis at de rettslige vurderingene nedenfor er midlertidige, og av den grunn kan ha endret seg innen denne artikkelen publiseres. Det som fremkommer nedenfor bør derfor ikke forstås som konklusjoner NAV har trukket, men som diskusjoner rundt potensielle måter å forstå rettsreglene.

I tillegg må jeg innledningsvis kommentere at datagrunnlaget mitt er noe snevert. De interne dokumentene jeg har fått tilgang til er preget av at de er under arbeid,<sup>190</sup> og selv om de er autentiske, er de ikke nødvendigvis oppdaterte. Jeg har derfor primært basert meg på informasjon fra intervjuene, mens dokumentene bare har fungert som supplerende materiale. Samtidig har juristen jeg intervjuet kun jobbet med rettslige vurderinger knyttet til profileringsystemet i noen måneder.<sup>191</sup> Det er derfor grunn til å tro at enkelte vurderinger som ikke kom opp i intervjuet med juristen, kan ha vært behandlet tidligere, av andre jurister i NAV. Juristen har for eksempel ikke vært med å vurdere etterlevelseskraav i løsningen, som favner bredere enn de vurderingene vi var innom i inter-

187 Hellman (2020) s. 835.

188 Se nærmere om dette i avsnitt 4.4.3 under.

189 Sml. avsnitt 2.1.

190 Se avsnitt 1.4.2.

191 Se generelt om dette i avsnitt 1.4.3.2.

vjuet. Jeg påpeker også at juristen ikke har ansvar for å tolke ftrl. § 8-7 a. Til tross for utfordringene med datagrunnlaget mener jeg at dataene dekker flere store, rettslige problemstillinger, og at de som fremgår er interessante nok (uten at de nødvendigvis dekker alle rettslige problemstillinger NAV har vurdert).

Ellers er gjennomgangen nedenfor preget av at det er personvern vurderinger som er kjernen i NAVs rettslige diskusjoner rundt denne modellen. Det skyldes nok rett og slett at de fleste relevante rettslige rammer for systemutvikling fremkommer av personvernregelverket.<sup>192</sup> Juristen jeg intervjuet var imidlertid tydelig på at en del av de rettslige kravene som fremkommer i personvernforordningen, også må tolkes i lys av øvrige krav til forvaltningen, spesielt forvaltningsloven og forvaltningsrettslige prinsipper. NAV har derfor måttet bruke en del tid på å vurdere spørsmål som for eksempel hvorvidt innkalling til dialogmøte 2 er et enkeltvedtak eller ikke.

En avgjørelse må innebære rettigheter eller plikter overfor den personen avgjørelsen retter seg mot for at det skal være et enkeltvedtak.<sup>193</sup> Juristen påpekte at ikke alle ledd i et sykefraværsløp kan være enkeltvedtak; da er det en risiko for at flyten i sykefravær oppfølgingen blir borte, fordi veileder bare må sitte å vurdere klager på vedtakene. Jeg påpeker for ordens skyld at NAV foreløpig vurderer det slik at vurdering av dialogmøte 2 ikke er et enkeltvedtak i seg selv, men må forstås som en del av et større saksbehandlingsløp. Vurderingen er imidlertid stadig gjenstand for intern diskusjon. Enn så lenge betyr det at øvrige bestemmelser i forvaltningsloven som knytter seg til enkeltvedtak,<sup>194</sup> for eksempel retten til klage, ikke vil gjelde for denne vurderingen.

Framstillingen nedenfor dekker ulike problemstillinger juristen og jeg var innom i løpet av intervjuet om rettslige problemstillinger. I tillegg kom det opp enkelte relevante rettslige problemstillinger under intervjuet med data scientisten, og disse er også behandlet nedenfor. Der det er tilfelle, kommenteres det eksplisitt.

#### **5.4.1 Lovligheten til behandlingen**

Vurderingen av om sykefraværsteamet har rettslig grunnlag for behandlingen av personopplysninger i profileringssystemet er grunnleggende for å vurdere systemets lovlighet. Arbeidet med prediksjoner av sykefraværslengde forutsetter et rettslig grunnlag både etter personvernforordningen artikkel 6 nr. 1 og 9 nr. 2. Ettersom prediksjonene skal brukes som ledd i myndighetsutøvelse, må disse

---

192 Se avsnitt 3.3.2.

193 Jf. fvl. § 2 bokstav b.

194 Se fvl. kapittel IV, V og VI.

rettslige grunnlagene basere seg på lovhjemmel, jf. personvernforordningen artikkel 6 nr. 3 og artikkel 9 nr. 2 b eller g. For å kunne vurdere lovligheten av behandlingen, må NAV derfor se på forholdet mellom den planlagte behandlingen og lovhjemmelen de baserer behandlingen på. Juristen forklarte at dette forutsetter dialog med fagpersoner, som har best kunnskap om den aktuelle hjemmelen.<sup>195</sup>

Etter min forståelse vurderer NAV foreløpig ftrl. § 8-7 a som supplerende rettslig grunnlag,<sup>196</sup> både for utvikling og bruk av prediksjonene. Vurderingene av om det supplerende rettslige grunnlaget er tilstrekkelig må gjøres separat for behandling av personopplysninger til saksbehandling, og behandling av personopplysninger til trening av maskinlæringsmodellen. Juristen trakk også fram behandlingens forholdsmessighet som en viktig vurdering, altså om den planlagte behandlingen er står i et rimelig forhold til det inngrepet behandlingen innebærer for de registrerte.<sup>197</sup> Jeg skal nå kort skissere utfordringer knyttet til vurdering rundt rettslig grunnlag for begge behandlingene.

#### **5.4.1.1 Supplerende rettslig grunnlag**

Den ene måten NAV behandler personopplysninger på i tilknytning til profileringssystemet, er den behandlingen av personopplysninger som skjer når den sykmeldtes personopplysninger brukes til å vurdere sykmeldingslengden til vedkommende. Juristen påpekte at et viktig poeng er at prediksjonene må brukes i saksbehandlingen uten at de går på bekostning av lovgivers intensjon, altså at dialogmøte 2 som hovedregel skal gjennomføres. NAV forstår ellers foreløpig lovhjemmelen som teknologinøytral, fordi bestemmelsen hverken tilsier at vurderingen må gjøres manuelt av en veileder i NAV, eller at bestemmelsen åpner for maskinlæring.

Ved vurderinger av forholdsmessigheten til behandlingen trekkes det på den ene siden fram at behandlingen av personopplysningene til sykmeldte innebærer en behandling av helseopplysninger, som dermed i utgangspunktet er inngripende.<sup>198</sup> På den annen side trekkes det fram at de mulige negative konsekvensene av verktøyet er begrenset ettersom beslutningsstøtten uansett kan overstyres av veileder, og den sykmeldte uansett har rett til å kreve et dialogmøte.<sup>199</sup> Det forventes også at profileringssystemet innebærer en kvalitetsgevinst for den enkelte bruker, ved at beslutningene rundt behovet for dialogmøte 2 oftere blir riktigere.

<sup>195</sup> Sml. avsnitt 4.2.1.

<sup>196</sup> Se en nærmere gjennomgangen av denne i 3.1 over.

<sup>197</sup> Sml. Schartum (2016) s. 118.

<sup>198</sup> Jf. personvernforordningen artikkel 9 nr. 1.

<sup>199</sup> Se avsnitt 3.1.



Den andre måten NAV behandler personopplysninger på i forbindelse med profileringssystemet, er når maskinlæringsalgoritmen bruker personopplysningene til en rekke tidligere sykmeldte personer til å trene en maskinlæringsmodell.<sup>200</sup> Det er mindre klart hvordan ftl. § 8-7 a kan begrunne denne typen behandling av personopplysningene til mennesker som ikke lenger følges opp av NAV. Juristen påpekte at det kanskje kan være mulig å trekke paralleller til veilederes bruk av tidligere saker for å «lære seg» hvilke typer sykefravær som gjerne varer lenge, og hvilke som gjerne varer kort. Et annet alternativ NAV utforsker, er å trekke paralleller til forskning i forbindelse med utviklingsarbeidet, fordi Datatilsynet har hatt noen vage uttalelser om dette i tidligere veiledninger.<sup>201</sup> NAV har bedt om en oppfølging av dette i den regulatoriske sandkassen.

Forholdsmessighetsvurderingen er ifølge juristen vanskeligere for trening av modellen enn saksbehandlingen, fordi personopplysningene tilhører mennesker som ikke vil få noen direkte gevinst ut av at sykmeldingsoppfølgingen forbedres. Samtidig innebærer også denne delen av arbeidet bruk av helseopplysninger, som derfor gjør behandlingen spesielt inngripende.<sup>202</sup> I forbindelse med arbeidet med den regulatoriske sandkassen til Datatilsynet, har tilsynet også påpekt at helsesektoren har fått en egen lovhjemmel for å bruke pasientopplysninger til maskinlæringsformål. NAV skal derfor se nærmere på dette.

Utfallet av vurderingen av rettslige grunnlag, både for trening og bruk av modellen, har uansett store konsekvenser for sykefraværsteamet. Dersom NAV kommer fram til at de ikke har et rettslig grunnlag, kan ikke modellen brukes. Det faktum at NAV fortsatt jobber med vurderinger rundt det rettslige grunnlaget, til tross for at arbeidet med utvikling av beslutningsstøtten på undersøkelsestidspunkt (dvs. våren 2021) har foregått i flere år, sier sitt om hvor utfordrende vurderingen er å konkludere.

En grunn til at det er utfordrende å vurdere rettslig grunnlag, er ifølge juristen at det mangler rettskilder som angir hvor eksplisitt supplerende rettslige grunnlag må være, særlig ved behandling av helseopplysninger. Både data scientisten og juristen påpekte også at NAV har blitt ekstra påpasselige med vurderinger av det rettslige grunnlaget i etterkant av CV-saken, der NAV trodde de hadde et supplerende rettslig grunnlag for deling av CVer, noe som etterpå viste seg å være en feilaktig vurdering.<sup>203</sup> Data scientisten forklarte at slik usikkerhet rundt det rettslige grunnlaget får utviklingsprosessen til å stoppe litt opp. Han mente

200 Jf. avsnitt 4.3.1.

201 Sml. Datatilsynet (2018) s. 16.

202 Personvernforordningen artikkel 9 nr. 1.

203 NAV (2021).

derfor at det kunne være behov for å koble ledelsen på utviklingsarbeidet, for å ta en autorativ avgjørelse, eller i hvert fall gi tydelig støtte til juristenes vurderinger.<sup>204</sup> Både juristen og data scientisten håpet at samarbeidet med Datatilsynet i den regulatoriske sandkassen ville være oppklarende. Det kan uansett være grunn til å vurdere om lovgiver generelt bør legge noen tydeligere føringer for hvilke hjemmelskrav som stilles til maskinlæring i norsk offentlig sektor.<sup>205</sup>

Basert på intervjuet med juristen og supplerende interne dokumenter virker det på meg som at NAV så langt primært har sett på det rettslige grunnlaget for bruk av profileringssystemet i saksbehandling. Det kan derfor være grunn til å tenke at denne vurderingen er sikrere enn vurderingen av rettslig grunnlag for behandling av personopplysninger i forbindelse med trening av modellen. Et særlig poeng som ikke ble trukket frem i intervjuet (men som jeg heller ikke spurte eksplisitt om), er prinsippet om formålsbegrensning. Når formålet med innsamling av personopplysningene i utgangspunktet var å behandle enkelt-saker, må det vurderes om gjenbruk av personopplysningene til å trene en maskinlæringsmodell er forenlig med dette originale formålet, jf. personvernforordningen artikkel 6 nr. 4.<sup>206</sup> Dersom trening av modellen ikke er forenlig med det originale formålet, må det nye formålet begrunnes i en (egen) lovhjemmel.

I intervju med data scientisten fremgikk det at teamet må gjennom en (relativt uformell) intern søknadsprosess for å få tilgang til de interne datakildene til NAV. Søknaden innebærer blant annet at teamet må kunne peke til et reelt behov for tilgang til opplysningene, samt at nødvendige juridiske vurderinger er tatt og dokumentert. Dette ble ikke satt i en eksplisitt rettslig kontekst i intervjuer. Jeg antar imidlertid at det knytter seg til NAVs plikt til å sikre konfidensialitet i personvernforordningen artikkel 5 nr. 1 bokstav f, og taushetsplikten for NAV-ansatte, jf. forvaltningsloven § 13 første ledd nr. 1.<sup>207</sup>

#### 5.4.1.2 Nødvendighet

For at behandlingen av personopplysninger skal være lovlig, må de personopplysningene som brukes til behandlingen være nødvendige for å utføre oppgaven i allmennhetens interesse eller utøve offentlig myndighet, jf. personvernforordningen artikkel 6 nr. 1 bokstav c og e. Nødvendigheten av personopplysningene må altså begrunnes i det supplerende rettslige grunnlaget. Ettersom det er fagpersoner som har tolkningsansvaret for frtl. § 8-7 a, har ikke juristen jeg intervjuet tatt konkret stilling til om de opplysningene som brukes er nødvendige.

204 Sml. avsnitt 4.2.2.1.

205 Se nærmere om dette i 4.4.5.

206 Se avsnitt 3.3.2.

207 Se mer om dette i avsnitt 3.3.2.

Juristen forklarte at behovet for hvert enkelt informasjonselement likevel må dokumenteres og begrunnes, slik at det er mulig å si sikkert at de er nødvendige. Utover dette forutsetter juristene at teamet har gjort faglig forsvarlige utvalg av informasjonselementer. Hun trakk imidlertid fram enkelte informasjonselementer som kan ha mindre åpenbar tilknytning til det rettslige grunnlaget, som for eksempel bruk av informasjon om sykmelders tidligere praksis og lokasjon.<sup>208</sup>

Både i intervjuet med juristen og data scientisten fremkom det at teamet ønsker å utvide antall datakilder, for eksempel til Folkeregisteret og Aa-registeret. Data scientistene ønsker å lage en best mulig modell for prediksjon, og flere datakilder kan potensielt gi bedre prediksjoner. Samtidig er det ikke alltid gitt at en bedre modell er forholdsmessig med inngrepet i andre rettigheter, for eksempel personvern. Dette er et spesielt viktig å ha med seg ved utvikling av løsninger i NAV, som sitter på svært mange personopplysninger om mennesker i særlig utsatte situasjoner. Juristene er derfor tydelige på behovet for å begrense seg til de informasjonselementene som kan utledes fra lovhjemmelen, gitt lovgivers intensjon. Enn så lenge benytter teamet dermed kun datakilder som inneholder informasjon tilsvarende den informasjonen veiledere allerede bruker til å predikere sykefraværslengde i dag.<sup>209</sup> Når teamet velger ut datakilder kan det uansett virke som det er et spenningsforhold mellom data scientister/teknologer og jurister, hvilket kan begrunne tydelige og eksplisitte rammer for bruk av personopplysninger i denne typen prosjekter.

#### **5.4.1.3 Eventuelle tilleggskrav i PVF artikkel 22 og NAV-loven § 4 a**

NAV har brukt en del tid på å vurdere hvorvidt profileringsystemet er omfattet av bestemmelsene i personvernforordningen artikkel 22. Det forutsetter at behandlingen (1) er helautomatisert og (2) «har rettsvirkning for eller på tilsvarende måte i betydelig grad påvirker» de NAV behandler personopplysninger om. Når en behandling er omfattet av personvernforordningen artikkel 22, omfattes den generelt også av NAV-loven § 4 a. Denne bestemmelsen gir supplerende rettslig grunnlag for helautomatisert behandling av personopplysninger, herunder helseopplysninger, til saksbehandling. Bestemmelsen presiserer imidlertid at den ikke kan brukes dersom beslutningen innebærer skjønsmessige vurderinger. Dersom behandlingen omfattes av personvernforordningen artikkel 22, og derfor også NAV-loven § 4 a, forventer NAV dermed at behandlingen ikke er lovlig uten et nytt rettslig grunnlag.

---

208 Se imidlertid avsnitt 3.1 for en gjennomgang av ftrl. § 8-7 a.

209 Se avsnitt 3.1.

NAV har enn så lenge vurdert at det er mulig å tolke personvernforordningen artikkel 22 både veldig snevert og veldig bredt, og er foreløpige usikre på hva som er den korrekte tolkningen. Jeg går kort gjennom NAVs foreløpige vurderinger knyttet til bestemmelsen under.

Den første spørsmålet er hvorvidt vilkåret om at behandlingen er helautomatisert dreier seg om profileringssystemet i seg selv, som er helautomatisert, eller om det dreier seg om hele saksbehandlingskonteksten, som ikke er helautomatisert. Inntrykket mitt er at NAV foreløpig mener det blir urimelig å la bestemmelsen omfatte en løsning som skal brukes som beslutningsstøtte.<sup>210</sup> Dersom beslutningsstøtte skal tolkes innenfor personvernforordningen artikkel 22, blir det en veldig streng tolkning av bestemmelsen. Så lenge en veileder vurderer prediksjonene i forkant av en vurdering av behovet for dialogmøte 2, er med andre ord beslutningen ikke omfattet av artikkel 22.

Slik jeg forstod juristen, er et mer utfordrende å tolke vilkåret om at beslutningen må ha rettsvirkning eller på annen måte i betydelig grad påvirke den sykmeldte. Juristene diskuterer for øyeblikket om det er mulig å trekke en parallell til beslutninger som regnes som enkeltvedtak etter forvaltningsloven § 2 bokstav b eller om personvernforordningen artikkel 22 favner bredere. Juristen påpekte at dette er noe de diskuterer med Datatilsynet, særlig etter at dette forholdet ikke ble behandlet i særlig ustrakt grad i NOU til ny forvaltningslov.

En vurdering av at profileringssystemet omfattes av personvernforordningen artikkel 22 og/eller NAV-loven § 4 a har tilsvarende virkning for teamet som et manglende rettslig grunnlag etter personvernforordningen artikkel 6 nr. 1 og 9 nr. 2. Da må arbeidet stanses, inntil NAV eventuelt har fått tilstrekkelig lov-hjemmel.

#### **5.4.2 Krav til forklarbarhet og rettferdighet**

NAV må behandle personopplysninger på en åpen og rettferdig måte, jf. personvernforordningen artikkel 5 nr. 1 bokstav a. Bestemmelsen må også forstås i lys av kravene i forvaltningsloven. Generelt brukes gjerne begrepet «forklarbarhet» om åpenhet i forbindelse med maskinlæring, og jeg bruker derfor også dette begrepet i det følgende. NAVs arbeid med både forklarbarhet og rettferdighet i maskinlæringsalgoritmen er fortsatt i startgropa, men slik jeg har forstått det har arbeidet med forklarbarhet kommet noe lenger enn arbeidet med rettferdighet.

---

210 Sml. også avsnitt 3.3.2.

Basert på intervjuene, virker det som at både forklarbarhet og rettferdighet primært forstås som mer praktiske spørsmål, der det handler om å finne en akseptabel løsning innenfor mer generelle rettslige rammer. Det betyr i praksis at det er sykefraværsteamet og enkelte andre teknologer/data scientister som jobber med disse vurderingene enn så lenge. Data scientisten forklarte også at det stort sett er teamet som igangsetter de rettslige diskusjonene rundt disse utfordringene. I den grad jurister er koblet på arbeidet, fremstår det som at forklarbarhet og rettferdighet istedenfor inngår i de helhetlige vurderingene juristene tar rundt løsningens lovlighet. Det behandles imidlertid etter min forståelse ikke som rettslige spørsmål i seg selv.

Kravene til forklarbarhet kommer til å påvirke systeminnretningen, ved at behovet for ytterligere forklarbarhet kan være en utslagsgivende faktor for at teamet må bytte algoritme.<sup>211</sup> *Gradient boosting* er blant de mindre forklarbare algoritmene, og dersom kravene som stilles til forklarbarhet overgår det teamet greier å få til med algoritmen, må den erstattes. Dette er imidlertid under arbeid, og diskuteres både med eksterne tekniske partnere (Big Insight), forklarbarhetspartnere (AI4Users) og jurister (Datatilsynet, og etter hvert også Det juridiske fakultet i Oslo).<sup>212</sup>

Generelt er de rettslige kravene til forklarbarhet i personvernforordningen et hett tema i juridisk litteratur.<sup>213</sup> Ettersom NAV uansett må tolke kravene i forordningen i lys av øvrige krav i forvaltningsloven, herunder også prinsipper om forsvarlig saksbehandling, er omfanget av den retningsgivende rettslige litteraturen sannsynligvis mye mer begrenset. Samtidig påpeker jeg at EUs foreløpige forslag om regulering av kunstig intelligens innebærer enkelte mer konkrete krav knyttet til sluttbrukers (altså veileders) forståelse av utgangsverdiene fra maskinlæringsmodellen, for å kunne bruke disse riktig, jf. AIA art. 13 nr. 1. Det er også mulig at AIA art. 52 nr. 1 kan gjøre seg gjeldende for profileringssystemet, og det må i så fall fremgå tydelig for veileder at vedkommende interagerer med et maskinlæringsystem.

Av det jeg har sett, er arbeidet med å vurdere omfanget av kravet til rettferdighet bare så vidt i gang hos NAV. Resultatet av disse vurderingene kommer imidlertid garantert til å påvirke algoritmene, fordi det innebærer konkrete valg av såkalte «rettferdighetsmetriker». Jeg har ikke behandlet dette i gjennomgangen av den tekniske innretningen, fordi dette aspektet ved maskinlæringsalgoritmen har kommet veldig kort. Data scientisten påpekte imidlertid at noen av

---

211 Sml. avsnitt 4.3.2.

212 Se avsnitt 4.2.2.3.

213 Se f.eks. Bibal (2021).

vurderingene teamet jobber med, er hvordan man skal håndtere at et forsøk på å ta høyde for skjevheter innebærer at prediksjonskraften til modellen går ned – hva blir viktigst, prediksjonskraft for de fleste eller rettferdighet for enkeltgrupper? En fordel er uansett at det går an å måle modellen etter disse rettferdighetskriteriene, og på denne måten synliggjøre (u)rettferdige vurderinger.

Data scientisten forklarte at det er utfordrende for teamet å diskutere rettferdighetsvurderinger med øvrige personer i NAV uten teknisk kompetanse, uten å ta utgangspunkt i *konkrete* rettferdighetsvurderinger. Dermed er planen at teamet velger ut noen måter å håndtere rettferdighet, for å kunne bruke dette som basis for å forklare utfordringer med rettferdighet til ikke-teknologer. I praksis mente også data scientisten at rettferdighet basert på like typer feil/mengde på tvers av grupper, istedenfor rettferdighet basert på like mengder treningsdata på tvers av grupper, var en mer fruktbar måte å måle rettferdighet.<sup>214</sup> Dette skyldtes særlig at rettferdighet basert på feil/mengde på tvers av grupper er lettere å ta utgangspunkt i ved diskusjoner med personer som ikke har teknisk/statistisk bakgrunn.

Det må antas at enkelte krav knyttet til rettferdighet i teorien kan utledes fra norsk lov, som for eksempel likestillings- og diskrimineringsloven. Slik det er tilfelle for forklarbarhet, er også vurderinger rundt rettferdighet i litteraturen spredt og omdiskutert.<sup>215</sup> Et interessant poeng i forslaget til AIA, er at forordningen åpner for bruk av særlige kategorier av personopplysninger<sup>216</sup> for å kunne teste at en maskinlæringsmodell ikke predikerer dårligere for enkelte beskyttede grupper enn andre. Hvordan dette eventuelt skal kunne utnyttes i praksis av NAV er mindre klart. Det er vanskelig å se for seg hvordan NAV skal samle et datasett om for eksempel rasemessig eller etnisk opprinnelse, og det må antas at NAV neppe sitter på et fullstendig datasett med oversikt over denne typen opplysninger fra før.

### 5.4.3 Krav til dataminimering

Dataminimering innebærer at teamet kun skal bruke de personopplysningene som er nødvendige for å oppnå formålet med behandlingen.<sup>217</sup> Det har på flere tidspunkter, i alt fra intervjuer til de interne dokumenter, blitt trukket fram at data scientistene ikke vet hvilke data de trenger før de er ferdig med å lage algoritmer. Dermed forutsetter arbeidet med dataminimering at teamet tar fortløpende vurderinger rundt behovet for enkelt-informasjontyper i løsningen. Det kan dermed virke som at «nødvendighet» er knyttet tett opp mot prediksjonskraft. Dersom visse informasjonstyper konsekvent har ubetydelig predik-

214 Sml. også Hellman (2020).

215 Se en oversikt i Suresh (2019) s. 3–4.

216 Jf. personvernforordningen artikkel 9 nr. 1.

217 Jf. personvernforordningen artikkel 5 nr. 1 bokstav c.

sjonskraft, kan de eventuelt kuttes fra arbeidet. I så fall trenger NAV bare å sette en grense for prediksjonskraften til opplysningstypene, og deretter fjerne disse fra dataflyten sin. Hvorvidt denne grensen skal settes over 0, altså at det skal være strengere krav til prediksjonskraften enn at det ikke påvirker prediksjonene i det hele tatt, vet jeg ikke.

I tillegg til at dataminimering påvirker teamet på grunn av kravet om å vurdere behovet for den enkelte informasjonstype, er det enkelte andre måter dataminimering har påvirket arbeidet. For det første fremgår det av interne dokumenter at teamet jobber med pseudonymiserte personopplysninger, hvilket kan begrunnes i prinsippet om dataminimering.<sup>218</sup> Jeg påpeker også at arbeidet med å velge ut aktuelle datasett innebærer å kun behandle data om sykefraværsløp som har vart lenger enn 17 uker, og kun dataene som er tilgjengelige om sykefraværene i uke 17.<sup>219</sup> Dette kan også knyttes til dataminimering, uten at det har blitt gjort eksplisitt i intervjuene jeg gjennomførte.

I tilknytning til dataminimering er lagringsbegrensningsprinsippet veldig relevant.<sup>220</sup> Prinsippet ble ikke nevnt i intervjuene, men jeg spurte heller ikke eksplisitt om det. Så vidt jeg forstår på interne dokumenter, håndteres generell lagringsbegrensning av dataene som brukes til å trene modellen (altså sykmeldingene) av egne organisatoriske enheter som er ansvarlige for sykmeldingsmottak. Hvorvidt det er tatt vurderinger av dataminimering/lagringsbegrensning i denne sammenheng, vet jeg ikke. Teamet selv har etter min forståelse fastsatt forhåndsdefinerte slettetidspunkter for prediksjonene.

#### **5.4.4 Dokumentasjon**

Det er NAVs ansvar å kunne påvise at behandlingen av personopplysninger som skjer i forbindelse med trening og bruk av prediksjonene av sykefraværslengde foregår i samsvar med personvernforordningen.<sup>221</sup> Dette innebærer ingen eksplisitt forpliktelse til å dokumentere alle beslutninger og regelverksvurderinger NAV har gjort,<sup>222</sup> men implisitt vil det være nødvendig å ha en viss oversikt over hvorfor NAV for eksempel mener at de har et rettslig grunnlag for behandlingen.<sup>223</sup> Dette blir spesielt viktig i en setting der NAV jobber med et system som utfordrer en del av de rettslige prinsippene som allerede finnes.

---

218 Sml. avsnitt 3.3.2.

219 Se avsnitt 4.3.1.2.

220 Se avsnitt 3.3.2.

221 Jf. personvernforordningen artikkel 24 nr. 1.

222 Se imidlertid enkelte, konkrete krav til dokumentering i personvernforordningen artikkel 30.

223 Schartum (2020) s. 296.

Juristen jeg intervjuet forklarte at dokumentasjon av vurderingene, med tydelige begrunnelser for de valgene som er tatt, er helt sentralt for at hun skal føle at de kan stå stødig i beslutningene sine. Til tross for at det fremgikk av intervju med juristen at dokumentasjon anses som sentralt, virker det som at det er utfordrende å vite *hvor omfattende* begrunnelsene av rettslige vurderinger må være. Tilsvarende forklarte data scientisten at et grunnleggende problem med dokumentasjon av løsningen, er at det ikke finnes noen klare retningslinjer for hva som er god nok dokumentasjon. Slik jeg nevnte i forbindelse med vurderinger av rettslig grunnlag over, kan dermed selve vurderingsprosessen bli veldig langtekkelig.

Et av de mest interessante dokumentene er oversikten over informasjonselementer. Det illustrerer hvordan uklarheter i rettsreglene gjør det utfordrende for NAV å vite hvordan dokumentasjon skal skje i praksis. Det kan virke som at oversikten over informasjonselementer er et slags kompromiss mellom jurister og teamet, som følge av at arbeidet med personvernkonsekvensvurderingene forutsetter en fullstendig oversikt over de opplysningene teamet skal benytte i analysene sine. Ettersom teamet er avhengig av å prøve og feile med ulike opplysningstyper, kan ikke teamet lage en slik fullstendig oversikt på forhånd.<sup>224</sup> Resultatet blir da at teamet lager veldig brede, vage kategorier av informasjonselementer, som de deretter kan trekke ut konkrete informasjonstyper fra. Et viktig poeng er at oversikten over de konkrete informasjonstypene da bare er tilgjengelige gjennom kildekoden, som det altså kun er teamet selv som kan forstå.<sup>225</sup> Det er med andre ord bare informasjonselementene som i praksis er tilgjengelige for intern og ekstern granskning av andre enn teamet som jobber med prediksjonene.

Spørsmålet blir da om listen over informasjonselementer oppfyller formålet med å dokumentere hvilken informasjon teamet bruker i maskinlæringsmodellen sin. Av Datatilsynets veileder for personvernkonsekvensvurderinger fremgår det at det er tilstrekkelig å ha oversikt over de brede kategoriene av personopplysninger.<sup>226</sup> Jeg mener imidlertid at slike brede kategorier av informasjonselementer, som i praksis kan utlede 300 informasjonstyper basert på 13 kategorier,<sup>227</sup> er lite hensiktsmessige å bruke som utgangspunkt for drøftelse av innholdet i løsningen. Grunnen til at det er viktig å kunne drøfte enkelt-informasjonstyper, både for jurister og andre eksterne, er fordi informasjonstypene formes av data scientistene.<sup>228</sup> Jo viktigere informasjonstypene er for prediksjon-

224 Se kort om dette i avsnitt 4.3.1.2.

225 Sml. avsnitt 4.2.3.

226 Datatilsynet (u.å).

227 Jf. avsnitt 4.3.1.2.

228 Se avsnitt 4.3.1.



nene, og jo mer utfordrende det har vært for teamet å forme informasjonstypen, jo større grunn er det til å dokumentere avgrensningene som er gjort tydelig. Det kan for eksempel virke som at avgrensning av sykefraværsløp har vært såpass utfordrende, og samtidig er såpass sentral for det endelige resultatet, at det bør defineres tydelig.<sup>229</sup>

Ettersom teamet ikke kan vite på forhånd akkurat hvilke personopplysninger de trenger, kan en mulighet være å beholde oversikten over informasjonselementer, *samtidig* som en oversikt over de mer sentrale informasjonstypene også dokumenteres. Det innebærer med andre ord at definering og avgrensning av konkrete informasjonstyper i løsningen må hentes ut fra kildekoden, og dokumenteres eksplisitt i et eget dokument. Det kan også være et poeng å definere en egen standard for når slik dokumentasjon av informasjonstyper er nødvendig, slik at teamet kun bruker tid på de informasjonstypene som det er nødvendig å synliggjøre for interne jurister, og eventuelt eksterne virksomheter og interesseorganisasjoner.

Data scientisten trakk for øvrig fram at teamet har fått noen halvveis konkrete føringer som gjelder krav til arkivering og reproduserbarhet av resultatene fra arkiv-avdelingen i NAV. Det er imidlertid noe uklart hvordan disse kravene kommer til anvendelse for bruk av maskinlæringsalgoritmer, og teamet har dermed definert disse grensene litt selv. Dette aspektet var jeg ikke innom med juristen jeg intervjuet.

Det nye forslaget til regulering av kunstig intelligens, AIA, inneholder enkelte nye krav til dokumentasjon av maskinlæringsystemer. Det gjelder for eksempel hvilke metrikker som brukes til å måle maskinlæringsmodellen, hvordan data ble valgt ut, samt metodikk og fremgangsmåte for utviklingsarbeidet, jf. AIA artikkel 11, jf. Annex IV og artikkel 18. Til tross for at regelverket ikke gjelder enda, kan det være grunn til å tenke på forordningen som retningsgivende for dokumentasjonen teamet utarbeider allerede nå.

Det er også enkelte forslag om nye krav til dokumentasjon av beslutnings(støtte)systemer i forslag til ny forvaltningslov og ny arkivlov. I ny forvaltningslov fremgår det at innholdet i et beslutnings(støtte)system må dokumenteres.<sup>230</sup> Utvalget påpeker at slik dokumentasjon både skal muliggjøre legalitetskontroll av systemet, og at saksbehandler vil ha behov for tilgang til dokumentasjon av det rettslige innholdet i systemet for å kunne begrunne vedtak til parten. I forslag til

---

229 Jeg påpeker for øvrig at dette er dokumentert som et informasjonselement i listen. Dokumentasjonen gir imidlertid (foreløpig) liten innsikt i hvordan elementet faktisk er avgrenset.

230 NOU 2019: 5 s. 265.

ny arkivlov er det forslag om en ny plikt både til å dokumentere datatyper,<sup>231</sup> datakilder, og de behandlingsreglene som er utledet fra rettsregler i beslutnings(støtte)systemer.<sup>232</sup> Det er også et forslag om en egen plikt til å dokumentere avgjørelser i forvaltningen, og hva som har ledet fram til disse avgjørelsene. I hvilken grad dette er begrenset til forvaltningsvedtak, eller gjelder bredere, fremgår ikke eksplisitt. Ettersom bestemmelsen peker til «avgjørelser» istedenfor «enkeltvedtak», er det imidlertid grunn til å tenke at det favner noe bredere.

#### 5.4.5 Rettslige uklårheter

Rettsregler påvirker enkelte aspekter ved modellen direkte, slik jeg har gjennomgått over. I tillegg er det imidlertid enkelte påvirkningsforhold som skyldes det uklare regelverket, som jeg skal skissere kort.

Det første er at utydelige rammer for bruk av maskinlæring generelt gir jurister og team et uklart forhold til hva som er aktuelle endelige mål. Når ikke lovverket tar tydelig stilling til hvordan maskinlæring kan brukes i forvaltningen, utforsker forvaltningen selv mulighetsrommene. For øyeblikket forsøker NAV for eksempel å delautomatisere saksbehandlingsprosessen. Dette utfordrer grunnleggende forvaltningsprinsipper i mindre grad enn ved fullstendig automatisering, fordi en menneskelig saksbehandler har kontroll over beslutningen.<sup>233</sup> Både juristen og data scientisten forklarte imidlertid at sykefraværsteamet på sikt ønsker å helautomatisere deler av sykefraværsoppfølgingen. Da må juristen forholde seg til to rettslige vurderinger parallelt; både den som gjelder det teamet utvikler per dags dato (beslutningsstøtte med ganske begrenset funksjonalitet), og det teamet ønsker å utvikle på sikt (helautomatisert beslutningssystem). Dette gjør at det blir enda flere vurderinger å holde styr på, i et allerede uoversiktlig rettskildelbilde.

Når regelverket er uklart, overlates en del prinsipielle vurderinger rundt lovligheten av maskinlæringsalgoritmer i saksbehandling til enkeltjurister i forvaltningen, som sitter på omfattende og sensitive data som helseopplysninger.<sup>234</sup> Juristen jeg intervjuet forklarte at hun var skuffet over at forslaget til ny forvaltningslov ikke i større grad har tatt høyde for prinsipielle utfordringer i møte med denne teknologien. Hun forklarte videre at hun kan diskutere problemstillinger hun møter på med andre legal coacher, fagjurister i produktområdet og med juridisk seksjon for øvrig. I praksis får imidlertid legal coachene varierende grad av tilbakemeldinger på sine konklusjoner i seksjonen – det er rett og slett mange av de rettslige spørsmålene det ikke er så lett å ha så mange meninger om

231 Jf. mitt begrep «informasjonstyper».

232 NOU 2019: 9 s. 20.

233 Sml. Boe (2021) s. 709.

234 Jf. personvernforordningen artikkel 9 nr. 1.

uten å ta dypdykk i litteraturen. Det ender i praksis med at legal coachene tar mange foreløpige vurderinger, som igjen gir uklare rammer for utviklingsarbeidet. Dette kan også i praksis innebære at utviklingsarbeidet stopper opp.

Det er ikke dermed sagt at innføring av maskinlæring i saksbehandlingen norsk offentlig sektor bør gå så mye raskere; bruk av maskinlæring i saksbehandling forutsetter en rekke rettslige diskusjoner, som kommer til å ta tid. Det er imidlertid problematisk at det virker som at det er uklarheter i regelverket som gjør at arbeidet tar lang tid, istedenfor at det er gitt klare retningslinjer fra lovgiver om at utviklingen bør skje eller ikke bør skje. Her går det langt på vei an å se hen til paralleller rundt digitaliseringsvennlig regelverk: det er ikke gitt at alle regelverk bør være digitaliseringsvennlige, men dersom regelverket ikke er digitaliseringsvennlig, bør det være med vilje.<sup>235</sup>

Uten tydeligere signaler fra lovgiver, blir det (1) svært vanskelig for enkeltjurister å kunne uttale seg retningsgivende for arbeider som dette, og (2) opp til nybrottsarbeider som dette å vurdere hva rettferdighet, forklarbarhet, og lovlig maskinlæring skal innebære i forvaltningen. Det er ikke dermed sagt at teamene som jobber med slike nybrottsarbeidene finner dårligere løsninger enn det lovgiver ville funnet, eller kommer til å tolke regelverket feil, men det betyr i så fall at mye makt legges til forvaltningen selv. Ettersom forvaltningens arbeid i liten grad prøves i domstolene i Norge,<sup>236</sup> er det grunn til å tenke at det vil ta lang tid før eventuelle tolkningsfeil kommer til syne.

## 5.5 Samlet drøftelse

Forskningsspørsmålene mine dreide seg om hvordan profilering basert på maskinlæring kan forankres i rettsregler. For å besvare dette spørsmålet, har jeg undersøkt utviklingsarbeidet til NAV. Der har jeg sett på organiseringen av utviklingsarbeidet, den tekniske innretningen i utviklingsarbeidet, og hvordan rettsregler har påvirket utviklingsarbeidet. I kapittel 4 har jeg så langt vært innom mange forskjellige funn, med mange forskjellige implikasjoner for den rettslige forankringen til profileringssystemet til NAV. Jeg skal nå trekke noen overordnede linjer mellom funnene, for å understreke utfordringer jeg har kartlagt.

Profileringssystemet NAV utvikler baserer seg på data som er formet av teamet som bygger modellen. Maskinlæringsmodellen baserer seg både på mønstrene

---

235 Sml. Schartum (2018) s. 313.

236 NOU 2020: 11 s. 67.

som finnes i datasettene sykefraværsteamet har laget, og på kunnskaper om sammenhenger fra fageksperter i NAV. Organisatorisk er det primært teknologer, data scientister og veiledere som jobber med profileringssystemet fra dag til dag. Det kommer derfor til å preges spesielt mye av dette kompetansegrunnlaget. Innenfor teamet tas beslutninger fortløpende og uformelt, hvilket innebærer at en del av beslutningene ikke nødvendigvis dokumenteres. Også enkelte eksterne virksomheter påvirker utviklingsarbeidet, direkte eller indirekte, der Datatilsynet og Big Insight kanskje er de to mest fremtredende for øyeblikket.

Når det kommer til juridisk påvirkning på utviklingsarbeidet, er det enkelte utfordringer. NAV har en organiseringsstruktur der jurister jobber innenfor et større produktområde, der de veileder flere produktteam. Legal coach-strukturen er laget for smidig utvikling, og fungerer sikkert godt når teamet befinner seg innenfor mer «tradisjonelle» utviklingsprosjekter, der det finnes noen relativt klare svar.<sup>237</sup> Jeg stiller meg imidlertid mer tvilende til at en slik legal coach-struktur egner seg i et nybrottsarbeid som det sykefraværsteamet jobber med. Som nevnt påpeker både jurist og teamet at det er mangel på tid, hvilket gjør at juristenes involvering i arbeidet er utsatt.<sup>238</sup> I tillegg reiser bruk av maskinlæring i saksbehandling en del nye, rettslige spørsmål, som ikke nødvendigvis favnes av de standardiserte avsjekkingslistene NAV allerede har. Det gjør (1) at det sannsynligvis er enda mer utfordrende for teamet å vite når jurister bør involveres i diskusjoner, og (2) at juristene egentlig må bruke mye tid for å komme fram til rettslig holdbare konklusjoner.

Det fremgår også av avsnitt 4.4 over at det er store utfordringer med tolkning av regelverket for utvikling av beslutningsstøttesystemer som baserer seg på maskinlæring til bruk i saksbehandling i offentlig sektor. Flere prinsipielle problemstillinger knyttet til profileringssystemets lovlighet gjør at det ikke nødvendigvis er tilstrekkelig at NAV (eventuelt Arbeids- og sosialdepartementet) allokterer mer ressurser til juristene. I tillegg er det istedenfor nødvendig å vurdere hvor involvert lovgiver skal være, og hvor mye som skal være opp til NAV å løse selv. For de internasjonale rammeverkene, som for eksempel personvernforordningen, avhenger lovgivers involvering i tolkningsspørsmål og –presiseringer av det nasjonale handlingsrommet som gis i rammeverkene. For de norske rammeverkene er det istedenfor mer åpent.

Av gjennomgangen over, mener jeg det er enkelte av problemstillingene NAV jobber med som lovgiver må ta stilling til. Det gjelder for eksempel hvorvidt

---

237 Se Schartum (2018) for de mange rettslige utfordringene et team står overfor, bare i «vanlige» utviklingsarbeider.

238 Se mer om dette i avsnitt 4.2.2.2.

manglende eksplisitt angivelse av hvem som skal gjennomføre en vurdering, innebærer at vurderingen kan gjøres ved hjelp av maskinlæring eller ikke. Lovgiver bør også vurdere å angi aktuelle datakilder eksplisitt i rettskilddegrunnlaget, ettersom dette spørsmålet kan innebære spenninger mellom jurister og data scientister. Tydelige føringer for når dokumentasjon av det tekniske innholdet er nødvendig er et annet moment som bør behandles av lovgiver, fordi slik dokumentasjon gjerne nedprioriteres til fordel for å ferdigstille løsningen. Aktuell dokumentasjon kan for eksempel gjelde forholdet mellom mønstre i dataene og ekspertkunnskaper fra veiledere, og de mest sentrale og/eller utfordrende informasjonstypene. Hvordan forvaltningsorganer skal håndtere prinsippet om dataminimering ved bruk av maskinlæringsmodeller, er også et aktuelt spørsmål.

De forvaltningspolitiske rammene oppfordrer til at forvaltningsorganer bør utnytte maskinlæring, også i saksbehandling.<sup>239</sup> Hvis ikke lovgiver tar stilling til de prinsipielle spørsmål som oppstår ved bruk av maskinlæring, risikerer vi at den retningen forvaltningen tar ved bruk av maskinlæring primært begrunnes i forvaltningens prioriteringer, ikke bredere demokratiske føringer. Et sterkt fokus på ansvarlig og etisk kunstig intelligens, slik NAV har, er viktig, men ikke tilstrekkelig for å sikre rettslig forankring.<sup>240</sup> Spørsmålet blir da om disse prinsipielle vurderingene skal tas i generelle lover, som forvaltningsloven, eller spesielle lover, som NAV-loven eller folketrygdloven.<sup>241</sup>

Det finnes allerede enkelte konkrete forslag om regulering av kunstig intelligens i særlovgivning,<sup>242</sup> spesifikt innenfor helsesektoren. Her gis det hjemmel til utvikling av maskinlæringsmodeller. Helsedepartementet påpeker samtidig at en del av de større, prinsipielle utfordringene som knytter seg til bruk av maskinlæring bør håndteres for seg, på grunn av behovet for et omfattende kunnskapsgrunnlag. Det henvises til et samarbeid mellom Helsedirektoratet, Direktoratet for e-helse, og Statens legemiddelverk, som skal gå gjennom det gjeldende regelverket for å kunne legge til rette for bruk av kunstig intelligens. På grunn av sektorinndelingen i offentlig sektor kan det være vanskelig for NAV å utnytte dette initiativet direkte. Det kan likevel bidra til noen generelle forståelser av behovet for regelverksutvikling, og dermed vært verdt å følge med på. Det er uansett mulig å se for seg at NAV får en tilsvarende lovhjemmel fra Arbeids- og sosialdepartementet.

---

239 Se avsnitt 3.2.

240 Black (2019) s. 16.

241 Sml. Backer (2013) s. 27 flg.

242 Prop. 112 L (2020–2021) s. 45–46.

Ettersom spørsmålene rundt bruk av maskinlæring er av såpass prinsipiell og generell karakter, er det uansett god grunn til å vurdere om et utvalg skal se på det i forbindelse med oppdatering eller nyskriving av en generell lov.<sup>243</sup> Dessverre valgte utvalget som skrev NOU for ny forvaltningslov i 2019 å ikke vurdere særlige bestemmelser for bruk av maskinlæring, fordi utvalget mente det var for tidlig å regulere bruk av kunstig intelligens i offentlig forvaltning. Jeg påpeker i den sammenheng at det allerede i 2021, to år etter utredningen om ny forvaltningslov, har kommet forslag til regulering av kunstig intelligens fra Europakommisjonen.<sup>244</sup> NOU-utvalget understreket imidlertid «behovet for å være oppmerksomme på problemstillinger knyttet til kunstig intelligens i tiden som kommer».<sup>245</sup> Det betyr altså at de nødvendige, prinsipielle spørsmålene foreløpig ikke er håndtert generelt av norsk lovgiver. Flere aktører påpekte denne mangelen i hørings svar til NOUen,<sup>246</sup> og forhåpentligvis gjøres det et større arbeid med disse problemstillingene før ny forvaltningslov trer i kraft.

Det kan være grunn til å vurdere lovendringer både i generelle og særlige lover.<sup>247</sup> Et framtidig regulatorisk initiativ i Norge må uansett forholde seg til den nye forordningen for kunstig intelligens som er under arbeid i EU. Forordningen kan innebære enkelte tydeligere rettslige føringer for NAV, som for eksempel krav til dokumentasjon av utviklingsarbeidet og forklarbarhet. Det er imidlertid fortsatt for tidlig å si noe om hvilke faktiske konsekvenser forordningen kommer til å ha for velferdssektoren i Norge.

---

243 Backer (2013) s. 27.

244 COM (2021) 206.

245 NOU 2019: 5 s. 268.

246 Se Schartum (2019), NAV (2019a) s. 9 flg., og Boe (2019).

247 Sml. Boe (2020) s. 703.

## 6 Avsluttende bemerkninger

NAV holder på med et veldig interessant arbeid, og er også tidlig ute blant forvaltningsvirksomheter i Norge. Det er gode sjanser for at de erfaringene NAV gjør seg i dette arbeidet, kommer til å være svært verdifulle også for andre virksomheter. Jeg mener det er positivt at NAV har valgt å utvikle et beslutningsstøtteverktøy til et formål som ikke primært innebærer kontroll av individet. Det at NAV er såpass involvert i offentlige fora, samt brukes som case i flere forskningsprosjekter, er også bra.

Det viktigste funnet i denne rapporten, er at det virker å være svært stor usikkerhet om behandlingens lovlighet. Juristene som jobber med rettslige vurdering av profileringssystemet begrenses av lite tid. Det er imidlertid ikke bare et spørsmål om tid: Lovgiver har i stor grad overlatt til det enkelte forvaltningsorgan, og dermed i praksis til enkeltjurister innenfor disse organene, å vurdere store prinsipielle spørsmål i vår tid. Det blir åpenbart svært utfordrende, både for utviklingsteamene, juristene, og samfunnet for øvrig. Det bør derfor gjøres et større rettspolitisk arbeid, for å trekke noen klarere linjer for lovlig og ulovlig bruk av maskinlæring i norsk offentlig forvaltning. Dette arbeidet må ta høyde for kommende bestemmelser i ny forvaltningslov, arkivlov, og forordning om kunstig intelligens.

Jeg påpeker avslutningsvis at min oppfordring om å ta prinsipielle vurderinger på lovgivernivå, eller eventuelt fastslå autoritative fortolkninger av gjeldende lovbestemmelser, ikke bør forstås som en oppfordring om at regelverket skal tilpasse seg de tekniske behovene som oppstår i forbindelse med bruk av maskinlæring.<sup>248</sup> Regelverket bør imidlertid ta tydelig stilling til at forvaltningsorganer så smått har begynt å utvikle systemer som utnytter maskinlæring. Det innebærer å legge føringer for hvordan maskinlæring kan og bør anvendes i forvaltningen. Til syvende og sist er det lovgivers ansvar å sikre at de grunnleggende rettighetene til borgerne ivaretas i møte med forvaltningen.

---

248 Sml. Boe (2021) s. 700.

# Kilder

## Bibliografi

AICAST. «Robindra Prabhu og Cathrine Pihl Lyngstad, NAV – Anvendt brukersentrisk AI». Frank Vevle og Michael Løiten i samtale med Robindra Prabhu og Cathrine Pihl Lyngstad om NAVs arbeid med kunstig intelligens. *AICAST – Intervjuer med menneskene bak AI*. [Podkast], 22.10.2019. <https://www.buzzsprout.com/346424/1910975-robindra-prabhu-og-cathrine-pihl-lyngstad-nav-anvendt-brukersentrisk-ai?play=true>.

AI4Users. «Project Aim» (u.å.1) <https://ai4users.uia.no/research/>, hentet 27.03.2021.

AI4Users. «Research Team» (u.å.2) <https://ai4users.uia.no/about/>, hentet 21.05.2021.

Askland, Andrew. «Introduction: Why Law and Ethics Need to Keep Pace with Emerging Technologies» i *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight – The pacing Problem*. Cutter, Anthony Mark, Bert Gordijn, Gary E. Marchant, Alain Pampidou red., Dordrecht: Springer, 2011, s. xii-xxvii.

AlgorithmWatch. *Automating Society Report 2020*. Berlin: Oktober, 2020. <https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/10/Automating-Society-Report-2020.pdf>.

Backer, Inge Lorange. *Loven – hvordan blir den til?* Oslo: Universitetsforlaget, 2013.

Bench-Capon, Trevor. «Legal Cases: Argumentation versus ML». 2018. (<https://intranet.csc.liv.ac.uk/~tbc/publications/commaWorkshop18.pdf>)

Berk, Richard. *Machine Learning Risk Assessments in Criminal Justice Settings*. Cham: Springer Nature Switzerland AG, 2019.

Bibal, Adrien, Michael Lognoul, Alexandre de Streele og Benoît Frénay. «Legal requirements on explainability in machine learning.» *Artificial Intelligence and Law* vol. 29 nr. 2 (2021) s. 149–169.



Big Insight (u.å.) <https://www.biginsight.no/new-page>, hentet 27.05.2021.

Black, Julia, Andrew D. Murray. «Regulating AI and Machine Learning: Setting the Regulatory Agenda». *European Journal of Law and Technology* årg. 10, nr. 3 (2019).

Boe, Erik Magnus. *Høring – NOU 2019:5 Ny forvaltningslov [høringssvar]* 2019.

Boe, Erik Magnus. *Grunnleggende juridisk metode – En introduksjon til rett og rettstenkning*, 4. utg., Oslo: Universitetsforlaget, 2020.

Boe, Erik Magnus. *Rettskildelære under debatt*. 2. utg., Oslo: Universitetsforlaget AS, 2021.

boyd, danah, Kate Crawford. «Critical Questions for Big Data. Provocations for a cultural, technological and scholarly phenomenon.» *Information, Communication & Society* vol. 15, nr. 5 (2012) s. 552–679.

Broomfield, Heather og Lisa Reutter. *Kunstig Intelligens/Data Science: En kartlegging av status, utfordringer og behov i norsk offentlig sektor – Første resultater*. (2019) <https://ntnuopen.ntnu.no/ntnu-xmli/bitstream/handle/11250/2634733/KI%252C%2bdata%2bscience%2b-%2bKartlegging%2bav%2bstatus%252C%2butfordringer%2bog%2bbehov%2bi%2bnorsk%2boffentlig%2bsektor.pdf?sequence=1&isAllowed=y>

Brønnøysundregistrene. «Om Enhetsregisteret» (2020) <https://www.brreg.no/om-oss/oppgavene-vare/alle-registrene-vare/om-enhetsregisteret/>, hentet 15.05.2021.

Brønnøysundregistrene. «Åpne data – Enhetsregisteret: API-dokumentasjon – Ordbok» (2021) <https://data.brreg.no/enhetsregisteret/api/docs/index.html>, hentet 25.05.2021.

Büchi, Moritz, Eduard Fosch-Villaronga, Christoph Lutz, Aurelia Tamò-Larrieux, Shruthi Velidi, og Salome Viljoen. «The chilling effects of algorithmic profiling: Mapping the issues.» *Computer Law & Security Review* vol. 36 (2020).

Datakatalogen. «Begrep – etterlevelseskrav» (u.å) <https://data.nav.no/begrep/BEGREP-1873>, hentet 05.05.2021.

Datatilsynet. *Kunstig intelligens og personvern*. Rapport, januar 2018. Oslo: 2018. <https://www.datatilsynet.no/globalassets/global/dokumenter-pdf-er-skje-ma-ol/rettigheter-og-plikter/rapporter/rapport-om-ki-og-personvern.pdf>

Datatilsynet. *Personvernpodden – #1 Sandkasten – episode 1*. Arild Opheim i samtale med Lars Sutterud, Cathrine Pihl Lyngstad, Robindra Prabhu og Dag Grytli om rettferdighet i kunstig intelligens. [Podkast], 27.05.2021. <https://shows.acast.com/personvernpodden/episodes/1-sandkasten>

Datatilsynet. «Sjekkliste for vurdering av personvernkonsekvenser (DPIA).» (u.å.) <https://www.datatilsynet.no/globalassets/global/dokumenter-pdf-er-skje-ma-ol/regelverk/veiledere/dpia-veileder/sjekkliste-for-dpiafaser.pdf>, hentet 21.05.2021.

Datatilsynet. «Prosjektplan» (u.å.1) <https://www.datatilsynet.no/contentassets/ea0683afe9d04ff1a9e87be9747380c9/prosjektplan-for-nav---sladdet.pdf>, hentet 29.05.2021.

Digitaliseringsdirektoratet. «Produkteier» (17.10.2019) <https://www.prosjektveiviseren.no/roller/produkteier>, hentet 21.05.2021.

Digitaliseringsdirektoratet. *Fagforum for Kunstig intelligens i offentlig sektor: Prinsipper for ansvarlig KI v/Cathrine Pihl Lyngstad, NAV*. [Videoklipp], (28.02.2021) <https://www.digdir.no/fagforum-kunstig-intelligens-i-offentlig-sektor/2334>.

Digitaliseringsdirektoratets Seksjon for Informasjonsforvaltning. «[EPOS] Begrepets livssyklus #153». (20.03.2019) <https://github.com/Informasjonsforvaltning/behov/issues/153?fbclid=IwAR0VcOnu15iXiXpUAlTnj-H2X-uSUhz8xZLwQ5RkvAZ-XWx-5C6lZMAQpXiQ>, hentet 20.05.2021.

Delacroix, Sylvie. «From Agency-Enhancement Intentions to Profile-Based Optimisation Tools: What is Lost in Translation» i *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen*. Bayamlioglu, Emre, Irina Baraliuc, Liisa Janssens, Mireille Hildebrandt red., Amsterdam University Press, 2018, s. 16–19.

Eubanks, Virginia. *Automating inequality: how high-tech tools profile, police and punish the poor*. New York: St. Martin's Press, 2018.

Europakommisjonen. «White Paper. On Artificial Intelligence – A European approach to excellence and trust.» 19.02.2020.

Favaretto, Maddalena, Eva De Clercq, og Bernice Simone Elger. «Big Data and discrimination: perils, promises and solutions. A systematic review.» *Big Data* vol. 6 nr. 12 (2019).

Feenberg, Andrew. «Ten paradoxes of technology». *Techné: Research in Philosophy and Technology* vol. 14, nr. 1 (2010), s. 3–15.

Felles datakatalog. «sykmelder» (16.09.2020) <https://data.norge.no/concepts/b016d803-34de-4da9-b83f-65a6c4dc268d> hentet 16.04.2021.

Feng, P. og A. Feenberg. «Thinking about design. Critical Theory of Technology and the Design Process» i *Philosophy and Design. From Engineering to Architecture*, Pieter E. Vermaas, Peter Kroes, Andrew Light og Steven A. Moore red., Dordrecht: Springer Netherlands, 2008, s. 105–118.

FN. *Report of the Special Rapporteur on extreme poverty and human rights*. [Rapport fra spesialrapportøren til FNs generalforsamling, 74. møte]. A/74/493. New York: 11. oktober 2019. <https://undocs.org/A/74/493>

Frischsenteret. «Forside» (u.å.) <https://www.frisch.uio.no/>, hentet 27.05.2021.

Gitelman, Lisa. «Introduction» i «*Raw Data*» is an oxymoron, Lisa Gitelman red., Cambridge: MIT press, 2013, s. 1–14.

González, Elena Gil, Paul de Hert. «Understanding the legal provisions that allow processing and profiling of personal data – an analysis of GDPR provisions and principles» *ERA Forum* 19 (2019) s. 597–621.

Graver, Hans Petter. *Alminnelig forvaltningsrett*. 5. utg., Oslo: Universitetsforlaget AS. 2019.

Grønmo, Sigmund. *Samfunnsvitenskapelige metoder*. Bergen: Fagbokforlaget. 2004.

Grønning, Tuva Bønke og Ole-Martin Gangnes. «Bygger et digitalt juridisk miljø i NAV.» *Juristen*, 05.07.2019, <https://www.ub.uio.no/skrive-publisere/referere/fag/rettsvitenskap.html>, hentet 19.05.2021.

Haggerty, Kevin D., Richard V. Ericson. «The surveillant assemblage.» *British Journal of Sociology* vol. 51, nr. 4 (2000), s. 605–622.

Hellman, Deborah. "Measuring Algorithmic Fairness." *Virginia Law Review* vol. 106, nr. 4 (2020) s. 811–866.

Hildebrandt, Mireille. «Defining Profiling: A New Type of Knowledge?» i *Profiling the European Citizen: Cross-Disciplinary Perspectives*. M. Hildebrandt og S. Gutwirth red., Dordrecht: Springer Netherlands, 2008, s. 17–45.

Hildebrandt, Mireille. «Threats to fundamental rights in the onlife world» i *Smart Technologies and End(s) of Law*. Hildebrandt, Mireille red., Cheltenham: Edward Elgar Publishing Limited, 2015, s. 77–103.

Introna, Lucas D. «Maintaining the Reversibility of Foldings: Making the Ethics (Politics) of Information Technology Visible.» *Ethics and Information Technology* vol. 9, nr. 1 (2007) s. 11–25.

Jacobsen, Dag Ingvar. *Hvordan gjennomføre undersøkelser? Innføring i samfunnsvitenskapelig metode*. 3. utg., Oslo: Cappelen Damm AS, 2015.

Karrierestart.no. «Medarbeidere i IA og sykefraværskontoret (19/19138). Arbeids og velferdsdirektoratet, NAV». 2020. <https://karrierestart.no/ledig-stilling/1433488>, hentet 29.05.2021.

Kaufmann, Mareile. «The co-construction of crime predictions: Dynamics between digital data, software and human beings.» i *Moral issues in intelligence-led policing*. Nicholas R. Fyfe, Helene Gundhus, og Kira Vrist Rønn red., London: Routledge, 2017, s. 143–160.

Kaufmann, Mareile, Nanna Bonde Thylstrup, J. Peter Burgess, Ann Rudinow Sætnan. «Data Criticality.» *STS Encounters – Research Papers from DASTS – Special Issue* vol. 11, nr. 1 (2020), s. 227–254.

Kitchin, Rob, Tracey P. Lauriault. «Towards critical data studies: Charting and unpacking data assemblages and their work.» *The Programmable City Working Paper 2*; pre-print versjon til publisering i *Geoweb and Big Data*, Eckert, J., Shears, A. and Thatcher, J., red. 2014.

Kommunal- og moderniseringsdepartementet. «Strategi for kunstig intelligens». 2020.

Kommunal- og moderniseringsdepartementet. *Digitaliseringsrundskrivet*, 05.01.2021. H-7/20 <https://www.regjeringen.no/no/dokumenter/digitaliseringsrundskrivet/id2826781/>

Leese, Matthias. «The new profiling: Algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union.» *Security Dialogue* vol. 45, nr. 5 (2014), s. 494–511.

Lum, Kristian, William Isaac. «To Predict and Serve?» *Significance* vol. 13, nr. 5 (2016), s. 14–19.

LØRN.TECH. «#0215: AI: Cathrine Pihl Lyngstad: Ansvarlig AI.» Silvija Seres i samtale med Cathrine Pihl Lyngstad om ansvarlig kunstig intelligens hos NAV. *LØRN.TECH.* [Podkast], 05.02.2019. <https://shows.acast.com/lrnstech/episodes/215-ai-cathrine-pihl-lyngstad-ansvarlig-ai>.

Mendoza, Isak, Lee A. Bygrave. «The Right Not to be Subject to Automated Decisions Based on Profiling» i *EU Internet Law: Regulation and Enforcement*. Synodinou, Tatiana-Eleni, Philippe Jougoux, Christiana Markou, Thalia Prastitou, red., Cham: Springer International Publishing AG, 2017, s. 77–98.

Muller, Michael, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q. Vera Liao, Casey Dugan og Thomas Erickson. «How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation.» *Systems Proceedings (CHI'19)*. 2019.

NAV – Arbeids- og velferdsetaten. *Rundskriv til ftrl kap 8 – Sykepenger*, 1. mai 1997. Dokumentnr. R08-00. [Lest i Lovdata Pro]

NAV. «Retningslinjer» (2015) <https://www.nav.no/no/nav-og-samfunn/om-nav/brukermedvirkning/retningslinjer>, hentet 27.05.2021.

NAV. «Oppfølging av sykmeldte arbeidstakere (2019) [https://www.nav.no/\\_/attachment/download/51baa808-7f98-40e9-a574-31de-ac242db9:657330d0bd089139506442121e624f8f8a81f416/Oppf.sykmeldte%202019%20BM%20digital.pdf](https://www.nav.no/_/attachment/download/51baa808-7f98-40e9-a574-31de-ac242db9:657330d0bd089139506442121e624f8f8a81f416/Oppf.sykmeldte%202019%20BM%20digital.pdf) hentet 19.04.2021.

NAV. *Høringssvar – NOU 2019:5 Ny forvaltningslov – Arbeids- og velferdsetaten*. 2019a. <https://www.regjeringen.no/no/dokumenter/horing---nou-2019-5-ny-forvaltningslov/id2652945/Download/?vedleggId=4030036b-ba0c-4d60-a47a-d1961ae0f9d2>

NAV. *Årsrapport 2019*. Oslo: 2020. [https://www.nav.no/\\_/attachment/download/acd4c93a-c905-4734-ad61-8fe1f7f680d6:6e6c2bb21f38f08dd-2202c0936a47f5d1777b70c/%C3%85rsrapport%202019%20NAV.pdf](https://www.nav.no/_/attachment/download/acd4c93a-c905-4734-ad61-8fe1f7f680d6:6e6c2bb21f38f08dd-2202c0936a47f5d1777b70c/%C3%85rsrapport%202019%20NAV.pdf)

NAV. «Medlemmer» (2020a)  
<https://www.nav.no/no/nav-og-samfunn/om-nav/brukermedvirkning/medlemmer>

NAV. «Spørsmål og svar om stenging av tilgang til CV-er» (2021)  
<https://www.nav.no/no/person/arbeid/nav-har-gjort-din-cv-tilgjengelig-for-arbeidsgivere-pa-arbeidsplassen.no-uten-at-vi-har-hatt-lov-til-det/sporsmal-og-svar-om-stenging-av-tilgang-til-cv-er> hentet 21.05.2021.

Nordisk ministerråd. *Nordiske kommuners arbeid med kunstig intelligens*. Analyse 03/2019. København: 2019. <http://norden.diva-portal.org/smash/get/diva2:1365285/FULLTEXT01.pdf>

Panoptykon Foundation. *To track or not to track? Towards privacy-friendly and sustainable online advertising*. November 2019.  
[https://panoptykon.org/sites/default/files/publikacje/panoptykon\\_to\\_track\\_or\\_not\\_to\\_track\\_final.pdf](https://panoptykon.org/sites/default/files/publikacje/panoptykon_to_track_or_not_to_track_final.pdf)

Pasquale, Frank. *The black box society: the secret algorithms that control money and information*. Cambridge: Harvard University Press, 2015.

Redden, Joanna. «Democratic government in an age of datafication: Lessons from mapping government discourses and practices.» *Big Data & Society* vol. 5, nr. 2 (2018), s. 1-13.

Rudin, Cynthia. «Stop explaining black box machine learning models for high stakes decision making and use interpretable models instead.» *Nature Machine Intelligence* vol. 1 (2019), s. 206-215.

Russell, Stuart, Peter Norvig. *Artificial Intelligence – A Modern Approach*. 4. utg., utgivelsessted: utgiver, 2021.

Schartum, Dag Wiese og Lee Bygrave. *Personvern i informasjonssamfunnet – en innføring i vern av personopplysninger*. 3. utg., Bergen: Fagbokforlaget, 2016.

Schartum, Dag Wiese, Arild Jansen, Tommy Tranvik. *Digital forvaltning – en innføring*. Bergen: Fagbokforlaget, 2017.

Schartum, Dag Wiese. *Digitalisering av offentlig forvaltning – Fra lovtekst til programkode*. Bergen: Fagbokforlaget, 2018.

Schartum, Dag Wiese. *Høringsuttalelse vedrørende spørsmål om digital forvaltning i NOU 2019: 5*. 2019. <https://www.regjeringen.no/no/dokumenter/horing--nou-2019-5-ny-forvaltningslov/id2652945/Download/?vedleggId=-60b4ef70-4584-4330-aec8-16e45dd400e3>

Schartum, Dag Wiese. *Personvernforordningen – en lærebok*. Bergen: Fagbokforlaget, 2020.

Skullerud, Åste Marie Bergsens, Cecilie Rønnevik, Jørgen Skorstad og Marius Engh Pellerud. *Personvernforordningen (GDPR) Kommentartutgave*. 1. utg. Oslo: Forlag: Universitetsforlaget 2018.

Suresh, Harini, John V. Guttag. «A Framework for Understanding Unintended Consequences of Machine Learning.» *arXiv preprint arXiv:1901.10002* (2019).

Taylor, Linnet. «On the presumption of innocence in data-driven government» i *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen*. Bayamlıoğlu, Emre, Irina Baraliuc, Liisa Janssens, Mireille Hildebrandt red., Amsterdam University Press, 2018, s. 72-77.

Teknologirådet. *Rettferdige algoritmer*. [Videoklipp], (2021) <https://vimeo.com/534441761>.

Thylstrup, Nanna Bonde, Mikkel Flyverbom, Rasmus Helles. «Datafied knowledge production: Introduction to this special theme» *Big Data & Society*, (Juli, 2019).

Veale, Michael, Max Van Kleek, Reuben Binns. «Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making». *CHI '18: CHI Conference on Human Factors in Computing Systems*, New York: Association for Computing Machinery, 2018.

Veale, Michael og Irina Brass. «Administration by Algorithm? Public Management meets Public Sector Machine Learning» i *Algorithmic Regulation*. Karen Yeung og Martin Lodge red., Oxford: Oxford University Press, 2019.

Weiskopf, Richard. «Algorithmic Decision-Making, Spectrogenic Profiling, and Hyper-Facticity in the Age of Post-Truth.» *Le foucauldien* vol. 3, nr. 1 (2020), s. 1-37.

Yeung, Karen. «Algorithmic regulation: a critical interrogation.» *Regulation & Governance* vol. 12, nr. 4 (2018) s. 505-523.

Yin, Robert K. *Case Study Research – Design and Methods*. 5. utg., Thousand Oaks: SAGE Publications, Inc., 2014.

Yin, Robert K. *Case Study Research – Design and Methods*. 6. utg., Thousand Oaks: SAGE Publications, Inc., 2018.

Zuboff, Shoshana. *The Age of Surveillance Capitalism – The Fight for a Human Future at the New Frontier of Power*. London: Profile Books, 2019.

## Lov- og forarbeidsregister

- 1814                      Grunnlov 17. mai 1814 om Kongeriket Norge (Grunnloven) [Grl.]
- 1967                      Lov 10. februar 1967 om behandlingsmåten i forvaltningssaker (forvaltningsloven) [fvl.]
- 1997                      Lov 28. februar 1997 nr. 19 om folketrygd (folketrygdloven) [ftrl.]
- 2005                      Lov 17. juni 2005 nr. 62 om arbeidsmiljø, arbeidstid og stillingsvern mv. (arbeidsmiljøloven) [aml.]
- 2006                      Lov 19. mai 2006 nr. 16 om rett til innsyn i dokument i offentlig verksemd (offentleglova) [offl.]
- 2006                      Lov 16. juni 2006 nr. 20 om arbeids- og velferdsforvaltningen (arbeids- og velferdsforvaltningsloven) [NAV-loven]
- Prop.89 L (2010–2011)    *Endringer i arbeidsmiljøloven og folketrygdloven mv. (raskere oppfølging og sanksjonering av brudd på regelverket ved arbeidstakers sykdom)*
- Prop. 39 L. (2014–2015)    *Endringer i arbeidsmiljøloven og sosialtjenesteloven (adgang til midlertidig ansettelse mv. og vilkår om aktivitet for stønad til livsopphold)*
- Meld. St. 33 (2015–2016)    *NAV i en ny tid – for arbeid og aktivitet*
- For (EU) 2016/679            Europaparlamentets- 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] [PVF]

- 2017 Lov 16. juni 2017 om likestilling og forbud mot diskriminering (likestillings- og diskriminering-sloven).
- 2018 Lov 15. juni 2018 nr. 38 om behandling av personopplysninger (personopplysningsloven)
- NOU 2019: 5 *Ny forvaltningslov*
- NOU 2019: 9 *Fra kalveskinn til datasjø – Ny lov om samfunnsdokumentasjon og arkiver*
- NOU 2020: 11 *Den tredje statsmakt – Domstolene i endring*
- Prop. 1 S (2020–2021) *For budsjettåret 2021, Arbeids- og sosialdepartementet*
- Prop. 112 L (2020–2021) *Endringer i helsepersonelloven og pasientjournalloven (bruk av helseopplysninger for å lette samarbeid, læring og bruk av kunstig intelligens i helse- og omsorgstjenesten mv.)*
- COM (2021) 206 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 [AIA]

## Domsregister

- NCJM and Others v. State of the Netherlands* Rechtbank Den Haag, C/09/550982/HA ZA 18/388, 05.02.2020.