Skatteforsk
Centre for Tax Research

March 2025 / Note 06-2025b

Fra visjon til virkelighet

Hvordan offentlig sektor kan utnytte egne data bedre

Annette Alstadsæter

Hector Ulloa

Margunn Aanestad

Skatteforsk
Centre for Tax Research

Skatteforsk – Centre for Tax Research er et uavhengig forskningssenter ved Handelshøyskolen ved Norges miljø- og biovitenskapelige universitet (NMBU). Vårt mål er å bygge bro mellom banebrytende akademisk forskning og praktisk skattepolitikk.

Denne rapporten og workshopene bygger på ulike forskningsprosjekter finansiert av Norges forskningsråd:

- Skatteforsk – Senter for skatte- og atferdsforskning (prosjekt nr. 341289, 2023–2027)
- Lærende kontrollvirksomhet for å sikre riktig refusjon fra helserefusjonsordningene (Innovasjonsprosjekt i offentlig sektor, prosjekt nr. 321044, 2020–2024)
- AI4Users: Responsible Use of Artificial Intelligence through Design for Accountability and Intelligibility (prosjekt nr. 311680, 2020–2025)

Deltakende forskere på dette prosjektet:

- Annette Alstadsæter, professor ved NMBU, leder Skatteforsk
- Hector Ulloa, direktør for utadrettet virksomhet, Skatteforsk
- Margunn Aanestad, professor UiO (tilknyttet UiA og NMBU)
- Polyxeni Vassilakopoulou, professor UiA (leder av prosjektet AI4Users)
- Elena Parmiggiani, førsteamanuensis, NTNU (kun første workshop)
- Charlotte H.Grøder, PhD-student, NTNU
- Johannes Dønnem, Masterstudent UiA (kun første workshop)
- Magnus Lonebu, Masterstudent, UiA (kun første workshop)

Deltakende institusjoner:

- Skatteetaten
- NAV
- Helsedirektoratet
- HELFO
- Kripos
- Finansdepartementet
- Datatilsynet

For more publications visit: https://skatteforsk.no/

Contact us at: skatteforsk@nmbu.no

Skatteforsk's Notes Series is financed through support from The Research Council of Norway (Grant number 341289).

Etter rapporten ble publisert, ble det avdekket at et pågående prosjekt i Skatteetaten er fremstilt på feil måte. For å sikre at innholdet i rapporten gjenspeiler gjeldende status og omfang, er den opprinnelige teksten derfor justert. Endringen er innarbeidet i denne nye versjonen av rapporten.

Opprinnelig tekst: *«Skatteetaten presenterte et prosjekt der de benyttet kunstig intelligens i innkrevingsarbeidet for å segmentere befolkningsgrupper basert på deres adferd. Målet var å kombinere data fra ulike «silo systemer» og verdikjeder, slik at man kunne utvikle nye analyseprodukter på tvers av etatens ansvarsområder, som blant annet innkreving, fastsetting og folkeregisteret.»*

Oppdatert tekst: *«Skatteetaten presenterte et prosjekt hvor man forsøkte å kombinere data fra ulike «silo systemer» og verdikjeder, slik at man kunne utvikle nye analyseprodukter på tvers av etatens ansvarsområder, som blant annet innkreving, fastsetting og folkeregisteret.»*

Sammendrag

Offentlig sektor har et økende behov for å bruke data mer effektivt til analyse, kunstig intelligens (KI) og tjenesteutvikling. Samtidig møter aktørene juridiske og organisatoriske barrierer, særlig knyttet til bruk av data. For å realisere potensialet i offentlig sektors data må juridiske rammer og teknologiske løsninger avklares, samtidig som balansen mellom personvern og effektiv bruk av data ivaretas.

For å adressere disse utfordringene arrangerte *Skatteforsk – Centre for Tax Research* ved NMBU sammen med forskere fra *UiO, UiA*, og *NTNU* i 2024 en workshop-serie med representanter fra *Skatteetaten, NAV, Kripos, Helsedirektoratet, Helfo* og *Datatilsynet*. Målet var å kartlegge sentrale utfordringer, dele erfaringer og identifisere løsninger for mer effektiv og ansvarlig bruk av data i offentlig sektor.

Diskusjonene avdekket flere felles barrierer, som juridisk usikkerhet knyttet til GDPR, taushetsplikt og hjemler for KI-utvikling, at ulike etater jobber med de samme utfordringene hver for seg uten felles retningslinjer, samt mangelfull intern organisering av data som hindrer analyse og deling.

Eksempler fra Skatteetaten og NAV viser hvordan etablering av interne dataplattformer kan gi gevinster, men også at det krever klare retningslinjer og tett samarbeid mellom teknologer, jurister og saksbehandlere. Kripos delte erfaringer med bruk av KI i arbeidet med å identifisere barn utsatt for seksuelle overgrep, etter en lang prosess med organisatoriske og juridiske avklaringer om KI kan tas i bruk i etterforskning av alvorlig kriminalitet. Samtidig ble det løftet frem hvordan fragmenterte hjemler, manglende felles begrepsapparat og ulike etiske vurderinger kan gjøre det utfordrende å utvikle og implementere datadrevne løsninger.

Erfaringene fra workshopene viser at bedre utnyttelse av data i offentlig sektor krever både teknologiske løsninger, juridiske avklaringer og en kultur for aktiv utforskning av handlingsrommet. Det vil også være nyttig med økt deling og samarbeid på tvers av aktører.

For å styrke databruken i offentlig sektor, oppsummerer vi innsiktene fra disse workshopene som fire konkrete anbefalinger til Digitaliseringsministeren:

Anbefaling 1: Styrk intern informasjonsforvaltning og kompetanse. Offentlige aktører må organisere og forvalte egne data bedre for å sikre at de er egnet til analyser, KI-utvikling og deling. Dette krever økte ressurser, bedre kompetanse og etablering av felles plattformer for deling av erfaringer, juridiske vurderinger og tekniske løsninger. Digitaliseringsdirektoratet bør styrke rådgivningstjenester og tilgjengeliggjøre eksisterende verktøy.

Anbefaling 2: Styrk samarbeidet mellom jurister og teknologer – få på plass 'tech-jurister'. Ansett jurister med teknologiforståelse som kan samarbeide tett med utviklere og brukere for å sikre bedre juridiske avklaringer. Samtidig bør også jus-utdanningene styrkes med mer teknologikunnskap.

Anbefaling 3: Frem en kultur for ansvarlig risikohåndtering. Offentlige aktører bør utforske sitt handlingsrom mer aktivt og unngå overdreven risikoaversjon, og samtidig vurdere alternativkostnader ved ikke å nyttiggjøre seg tilgjengelige data. Myndighetene bør tydeliggjøre vurderingskriterier for databruk og KI-testing.

Anbefaling 4: Ta nasjonalt ansvar for avklaringer for teknologi-jus. Regjeringen må sikre klare hjemler for databruk, testing av KI og deling på tvers av sektorer. Felles retningslinjer bør utarbeides for KI i saksbehandling, datadeling og grensegangen mellom utforskende og formålsbestemt databruk.

Innhold

1. Introduksjon

Det er økende fokus på å effektivisere offentlig sektor gjennom økt bruk av data, datadrevet innovasjon og kunstig intelligens[1]. De siste årene har det vært betydelig satsing på å tilgjengeliggjøre og dele offentlige data, blant annet gjennom initiativer som «Orden i eget hus», datadelingsplattformen *data.norge.no* og etableringen av Nasjonalt ressurssenter for deling og bruk av data.

Disse tiltakene er imidlertid primært rettet mot deling av data med andre offentlige aktører, innbyggere og næringslivet. De adresserer i mindre grad hvordan offentlige organisasjoner kan forbedre sin egen bruk av data. Dette er en utfordring forskergruppen bak dette prosjektet har erfart gjennom langvarige prosjektsamarbeid med ulike etater. Manglende tilrettelegging for intern bruk av data kan hindre nødvendig omstilling og effektivisering. Som både Personvernkommisjonens rapport og Riksrevisjonens undersøkelse om deling og gjenbruk av offentlige data viser, har dette bidratt til at arbeidet med datadrevet innovasjon går sakte[2].

En sentral utfordring som hittil har fått for lite oppmerksomhet, er usikkerheten rundt etterlevelse av personvernregelverket i offentlig sektor. Denne usikkerheten fører ofte til en defensiv tilnærming og til at jurister i ulike deler av organisasjonen i for stor grad vender tommelen ned for prosjekter rundt kobling og bruk av egne data. Dette skaper et vedvarende gap mellom ambisjonene om en mer effektiv offentlig sektor og den praktiske implementeringen av virkemidler og politikk. Men personvernregelverket er ikke nødvendigvis en innovasjonsbrems – det åpner for fleksibilitet dersom samfunnsnytten klart overstiger eventuelle personvernulemper. Det er derfor avgjørende at offentlig sektor etablerer en enhetlig praksis for risikovurdering, slik at handlingsrommet utnyttes uten at personvernet kompromitteres.

For å belyse og redusere dette gapet, arrangerte **Skatteforsk – Centre for Tax Research** i 2024 en workshop-serie hvor forskere og representanter fra offentlig sektor diskuterte utfordringer knyttet til økt bruk av egne data. Forskerne, med bakgrunn innen økonomi, jus og informatikk, kom fra Norges miljø- og biovitenskapelige universitet (NMBU), Universitetet i Oslo (UiO), Universitetet i Agder (UiA) og Norges Tekniske-Naturvitenskapelige Universitet (NTNU). De offentlige aktørene som deltok inkluderte Skatteetaten, Helsedirektoratet, Helfo, Kripos, NAV og Datatilsynet. Formålet var å skape en arena for erfaringsdeling, identifisere barrierer, og løfte frem løsninger som kan styrke offentlig sektors evne til å utnytte data på en trygg og effektiv måte.

En viktig innsikt fra workshopene var at ulike etater står overfor de samme juridiske og organisatoriske utfordringene, men ofte forsøker å løse dem hver for seg. Dette fører til unødvendig ressursbruk, forsinkelser og frustrasjon, samtidig som det bremser innovasjon. En annen erfaring var at deltakerne verdsatte muligheten til å dele praktiske løsninger som fungerer i praksis og etterlyste flere arenaer for konkret løsningsfokus på felles utfordringer.

Denne rapporten oppsummerer innsiktene fra workshopene. Først presenterer vi bakgrunnen for initiativet, og deretter beskrives organiseringen av workshopserien. I delkapittel fire beskrives temaer som ble belyst i løpet av workshopene, både konkrete eksempler på initiativer og utfordringer fra aktørene, og mer prinsipielle spørsmål som ble diskutert. Rapporten avsluttes med en grundig drøfting av mulige løsninger og tiltak som vil kunne styrke dette arbeidet.

[1] Meld. St. 22 (2020–2021) Data som ressurs – Datadrevet økonomi og innovasjon
Digitaliserings- og forvaltningsdepartementet (2024). Fremtidens digitale Norge. Nasjonal digitaliseringsstrategi 2024–2030

[2] NOU 2022:11. Ditt personvern-vårt felles ansvar, Riksrevisjonen: Dokument 3:8 (2023–2024)

2. Bakgrunn

Erfaringer fra forskningsprosjektene "Lærende Kontroll" og "AI4Users" viser at avansert dataanalyse kan styrke kontroll og etterlevelse, men at tillit og åpenhet er avgjørende for aksept av KI-systemer. Samtidig skaper regelverket, særlig GDPR, usikkerhet om hvordan data kan brukes og deles. GDPRs krav til risikobasert vurdering og formålsbegrensning fører ofte til en restriktiv praksis, som kan hindre innovasjon. Fremover vil EUs nye AI Act stille ytterligere krav til utvikling og bruk av kunstig intelligens, særlig for systemer med høy risiko. For å balansere personvern, innovasjon og effektivisering, er det nødvendig å avklare handlingsrommet innenfor regelverket og etablere en mer enhetlig praksis for risikovurdering.

2.1 Innsikter fra forskningsprosjektene «Lærende Kontroll» og «AI4Users»

To sentrale forskningsprosjekter danner et viktig grunnlag for denne prosjektet, begge finansiert av Norges Forskningsråd.

Lærende kontroll (2021-2024) utforsket hvordan avansert dataanalyse og automatiserte verktøy kan forbedre kontroll og regelverksetterlevelse i helsesektoren. Prosjektet viste at selv begrensede kontrolltiltak kan gi store adferdseffekter og at systematisk bruk av data kan bidra til en mer treffsikker og kostnadseffektiv offentlig forvaltning.

Samtidig har *AI4Users* (2020-2025) utforsket hvordan kunstig intelligens (KI) kan gjøres mer forståelig og ansvarlig i offentlig tjenesteyting. Prosjektet viste at både innbyggere og ansatte i offentlig sektor foretrekker systemer der mennesker har en rolle i beslutningsprosessene og at åpenhet om hvordan KI-systemer fungerer er avgjørende for å sikre tillit.

Erfaringene fra disse prosjektene understreker at offentlig sektor ikke bare trenger bedre tilgang til data, men også gode mekanismer for å bruke dem på en ansvarlig måte. Dette forutsetter både teknologiske løsninger og en juridisk forståelse som gir aktørene trygghet i hvordan data kan deles og anvendes.

2.2 Juridiske utfordringer for datadeling: GDPR

En av de største utfordringene for databruk i offentlig sektor er usikkerheten knyttet til regelverket. *General Data Protection Regulation* (GDPR) har siden 2018 vært det sentrale rammeverket for personvern i EU og Norge. Den setter klare krav til hvordan personopplysninger kan behandles og deler opp ansvaret mellom dataansvarlige og databehandlere. Sentralt i GDPR er prinsippet om risikobasert tilnærming: målet er ikke å eliminere all risiko, men å identifisere, vurdere og minimere risikoen på en ansvarlig måte. For å ivareta dette krever GDPR at offentlige virksomheter gjennomfører en *personvernkonsekvensvurdering (DPIA – Data Protection Impact Assessment)* når databehandlingen kan utgjøre en høy risiko for individers rettigheter.

Blant de sentrale prinsippene i GDPR, som også må vurderes i en DPIA, er:

- ***Lovlighet, rettferdighet og åpenhet***, som betyr at data må behandles på et gyldig rettslig grunnlag og på en åpen måte.
- ***Formålsbegrensning,*** som betyr at data kan kun brukes til spesifikke, legitime formål.
- ***Dataminimering,*** som betyr at kun nødvendige data skal samles inn og lagres.
- ***Sikkerhet og ansvarlighet,*** som betyr at data må beskyttes mot misbruk, og virksomheter må kunne dokumentere etterlevelse.

En DPIA handler ikke om å eliminere risiko, men om å identifisere potensielle utfordringer, vurdere nødvendigheten av behandlingen og iverksette tiltak for å redusere risiko der det er mulig.

2.3 Nye utfordringer framover: Fra GDPR til AI Act

AI Act (KI-forordningen) er EUs kommende regelverk for kunstig intelligens, vedtatt i 2024. I motsetning til GDPR, som regulerer all behandling av personopplysninger, fokuserer KI-forordningen på trygg og ansvarlig utvikling og bruk av KI-systemer. Systemer basert på kunstig intelligens vil bli klassifisert i ulike risiko-kategorier. Ut fra hvordan de blir klassifisert, vil systemene måtte oppfylle ulike forpliktelser om f.eks. kvalitet, åpenhet, menneskelig tilsyn og sikkerhet. For noen KI-systemer (i kategorien høy-risiko) vil man for eksempel måtte fremlegge en FRIA (Fundamental Rights Impacts Assessment).

Selv om GDPR (og nå framover KI-forordningen) skal sikre ansvarlig bruk av data og KI, kan regelverket også skape praktiske utfordringer for datadeling i offentlig sektor. GDPRs krav til rettslig grunnlag, formålsbegrensning og risikohåndtering fører ofte til juridisk usikkerhet, noe som gjør at mange jurister inntar en forsiktig tilnærming og sier nei til bruk og deling av data av frykt for å bryte regelverket.

For å unngå at dette fører til unødvendige hindringer for innovasjon, er det avgjørende å avklare handlingsrommet innenfor GDPR og etablere en felles praksis for risikovurdering. GDPR krever ikke at all risiko elimineres, men at den håndteres på en ansvarlig måte.

Ved å bygge en solid forståelse av disse prinsippene nå, kan offentlig sektor stå bedre rustet til å møte de kommende kravene under KI-forordningen og sikre en balanse mellom personvern, innovasjon og effektivisering.

3. Organisering

For å undersøke hvordan offentlig sektor kan utnytte egne data bedre, arrangerte forskere fra Skatteforsk – Centre for Tax Research, UiO, UiA, og NTNU to workshoper i 2024 for deltagere fra Skatteetaten, Helsedirektoratet, Helfo, Kripos, NAV, Finansdepartementet og Datatilsynet. Første workshop fokuserte på å identifisere sentrale utfordringer knyttet til intern databruk, inkludert dataminimering, sammenstilling av data og balansen mellom automatisering og beslutningsstøtte. Den andre workshopen diskuterte mulige løsninger, med særlig fokus på etablering av data-infrastruktur, utvikling av intern kompetanse og deling av data mellom aktører. En arbeidsgruppe bestående av forskere og deltagere forberedte den andre workshopen basert på innsikter fra den første. I etterkant har forfatterne av denne rapporten jobbet med å skrive ut erfaringene. De deltagende etatenes representanter har i løpet av februar 2025 blitt forelagt rapportutkastet og innspillene innarbeidet.

3.1. Workshop 1 – Identifisering av utfordringer

Første workshop, avholdt ved NMBU 19. april 2024, hadde som mål å kartlegge sentrale utfordringer knyttet til offentlige etaters bruk av egne data. Forskere, PhD- og Masterstudenter fra NMBU, UiO, UiA og NTNU med tilhørighet til forskningsprosjektene nevnt ovenfor deltok og fasiliterte gruppediskusjonene. Fra offentlig sektor deltok representanter fra Skatteetaten, Helsedirektoratet, Helfo, Kripos, NAV og Datatilsynet.

Etter innledende foredrag av forskere og deltagere ble tre problemstillingene A-C under diskutert i mindre grupper med medlemmer fra forskerteamet som referenter.

A. Utfordringer knyttet til håndteringa av prinsippet om dataminimering.
- del 4 i denne rapporten

B. Utfordringer knyttet til kontekstualisering og sammenstilling av data.
- del 5 i denne rapporten

C. Utfordringer knyttet til automatisering versus beslutningsstøtte.
- del 6 i denne rapporten

En arbeidsgruppe bestående av forskere og representanter for de ulike etatene ble opprettet for å arbeide med problemstillingene fram mot neste workshop. Arbeidsgruppen møttes to ganger, i juni og august 2024. En PhD-student gjennomførte intervjuer med deltakerne i etterkant av workshopen.

3.2 Aktørenes utgangspunkt – Deling av erfaringer

Under første workshop delte etatene erfaringer knyttet til egne initiativer for å bruke data mer effektivt.

Datatilsynet delte erfaringer fra prosjekter som hadde vært gjennomført i den «regulatoriske sandkassen». Data-relatert tematikk kom opp i flere, og noen gjennomgående problemstillinger var de følgende: dataminimering og formålsprinsippet, intern dataforvaltning («orden i eget hus»), hvordan sikre data også hos avtaker/brukere av data, hvorvidt hjemlene til å bruke data også kunne omfatte trening av KI-modeller eller om det behøves eksplisitte hjemler til dette formålet.

Helfo Kontroll beskrev erfaringene med å etablere et internt datavarehus («Nøkkelinfo») for data som brukes til kontrollformål. Dataene eies av Helsedirektoratet og noe ligger av historiske årsaker hos NAV. Dette har vært krevende teknisk og forvaltningsmessig, dataene brukes på andre måter med andre krav, feiltoleranse osv. Det har vært en lang prosess hvor man måtte jobbe på ulike fronter, og man ønsker også å koble på andre typer data for å få et mer helhetlig bilde

Kripos beskrev utfordringer knyttet til mye ustrukturerte data, data med varierende kvalitet, og mye teknisk gjeld og utdatert teknologi. Ulike prosjekter støter på utfordringer dersom man vil koble data fra flere datakilder, og det mangler hjemmel for bruk av politiets egne data i utvikling av KI. Det arbeides med både språk-, objekts- og ansiktsdeteksjon. Det rettslige rammeverket for politiets bruk av passregistret i etterforskning av alvorlig kriminalitet har vært på plass siden 2013. Først i 2023 kunne politiet ta i bruk automatiserte søk mot passregistret i arbeidet med å identifisere barn utsatt for seksuelle overgrep.

Skatteetaten har automatisert mange prosesser, også innen datadeling (A-ordningen som eksempel). Automatisering bidrar til å effektivisere, men åpner også for feil og kriminalitet. Jobber man med avdekking av økonomisk kriminalitet, kan det være vanskelig å finne hjemmel til å dele data for å varsle andre aktører. Med andre ord er en av utfordringene å sette slik varsling inn i en prosess eller arbeidsflyt, samt finne løsninger for å kunne varsle og samtidig unngå å avsløre for mye om kontrollsettinger.

NAV beskrev hvordan de flere hundre autonome teamene deler datasett på datamarkedsplassen. Andre team som vil bruke data i produktutvikling, må søke om og begrunne behov for tilgang, og får tidsbegrenset tilgang. Her deles også behandlingsdefinisjoner, og ulike behandlingsgrunnlag (slik at teamene kan gjøre en vurdering av om de ulike grunnlagene passer), slik at dette blir en støtte til etterlevelse. De får en oversikt over lovkrav og kode, de kan se hvordan vurderingene er gjort, og selv vurdere om dette treffer løsningen de selv utvikler. Teamene kan også støtte seg på data scientist og

legal coaching. Operative data og funksjonell statistikk er ikke åpent tilgjengelig på grunn av usikkerhet om rettslig handlingsrom.

Disse erfaringene dannet grunnlaget for videre diskusjoner om hvordan offentlig sektor kan håndtere utfordringene knyttet til økt bruk av egne data på en ansvarlig og effektiv måte.

3.3. Workshop 2 – Identifisering av løsninger

Den avsluttende workshopen ble avholdt 7. november 2024. Forskere fra NMBU, UiO, UiA og NTNU deltok og fasiliterte gruppediskusjonene. Fra offentlig sektor deltok representanter fra Skatteetaten, Helfo, Kripos, Finansdepartementet og Datatilsynet.

Basert på innspill fra første workshopen og diskusjonene i arbeidsgruppen ble det diskutert mulige løsninger på de følgende utfordringene:

D. Muligheter ved etablering av data-infrastruktur, intern kompetanse og kapasitet.
- del 7 i denne rapporten

E. Muligheter for bruk av egne data på nye måter.
- del 8 i denne rapporten

F. Muligheter for deling og mottak av data fra eksterne samarbeidspartnere.
- del 9 i denne rapporten

For å sikre erfaringsutveksling og målrettede diskusjoner, ble deltakerne delt inn i grupper der hver gruppe inkluderte noen som sto overfor bestemte utfordringer og andre som allerede hadde funnet løsninger på lignende problemstillinger. Her forsøkte arrangørene å sette sammen aktører som eide problemet med andre aktører som hadde funnet en løsning eller framgangsmåte for å håndtere utfordringen.

4. Utfordringer knyttet til håndteringa av prinsippet om dataminimering.

Bakgrunn: *Det trenges ofte store datamengder for å utvikle f.eks. gode prediktive modeller. Dette kan være i strid med prinsippet om dataminimering, som innebærer at man skal begrense mengden innsamlede personopplysninger til det som er nødvendig for å realisere formålet med innsamlingen. Hvordan kan dette dilemmaet håndteres?*

Oppsummering av innspill: *Diskusjonene viste at utfordringene knyttet til dataminimering først og fremst handler om å balansere behovet for data med personvernhensyn og risikostyring. Mange mente at risikoen for å samle inn for mye data er størst i tidlige faser av prosjekter, og at en iterativ tilnærming – der datagrunnlaget evalueres og justeres underveis – kan bidra til å begrense dette. Flere etterlyste en mer fleksibel prosess, hvor jurister og teknologer samarbeider for å utforske handlingsrommet fremfor å avvise muligheter. Bruk av pseudo-anonymiserte eller syntetiske data ble løftet frem som en løsning i tidlige faser, samtidig som utfordringer knyttet til tilgangsstyring og sletting av data i datavarehus ble påpekt. Generelt ble det understreket at selve dataminimeringsprinsippet sjelden er hovedutfordringen, men at manglende samspill mellom tekniske og juridiske perspektiver kan føre til unødvendig forsiktighet og hemme innovasjon.*

Det overordnede temaet som gikk igjen i diskusjonene, var hvordan man på en forsvarlig måte kan balansere behovet for data mot risiko- og personvernhensyn, samtidig som man ivaretar handlingsrommet for innovasjon og utvikling. En gjennomgående erkjennelse var at faren for å inkludere et for omfattende datagrunnlag ofte er størst i de innledende fasene av et prosjekt, når man ikke alltid har fullstendig oversikt over hva man faktisk trenger. Etter hvert som man får mer kunnskap, kan det være hensiktsmessig å redusere eller «skrelle vekk» deler av datagrunnlaget. Dermed kan man lettere sikre at man ikke sitter igjen med unødvendig mye informasjon. Den opprinnelige vurderingen, der man kanskje har en tendens til å samle inn mer data enn strengt nødvendig, vil derfor sjelden være den endelige. Prosjekter bør ideelt sett ha en fleksibel tilnærming med kontinuerlige vurderinger av hva som er relevant og proporsjonalt, slik at man kan justere kursen og databehovenes omfang underveis.

I denne sammenhengen ble det også pekt på betydningen av å ha en såkalt «ja da»-prosess, der man i stedet for å avvise potensielle bruksområder eller analysemetoder tidlig, heller undersøker hvordan de kan gjennomføres på en måte som er innenfor de gjeldende juridiske og etiske rammene. Dette krever et tett samspill med jurister, hvor man både må ta hensyn til lovverk og retningslinjer, men også unngå at juridiske forhold blir et hinder for reelle muligheter. Likevel understrekes det at enhver avgjørelse må være fundert i en forholdsmessighetsvurdering som viser at innsamling og bruk av data står i rimelig forhold til formålet. Ved å argumentere godt for nødvendigheten av hvert element i datagrunnlaget, kan man oppnå en større trygghet rundt bruken av informasjonen.

Et annet gjennomgående tema var dataminimeringsprinsippet. Flere poengterte at dette sjelden er den største barrieren eller den mest presserende utfordringen i praksis, så lenge man har gode metoder for å vurdere og begrense databehovet. Bruk av pseudo-anonymiserte eller syntetiske data ble løftet frem som et nyttig verktøy, særlig i tidlige utviklingsfaser. Ved å benytte slike løsninger kan man teste og utforske modeller uten å utsette reelle personopplysninger for unødvendig risiko. Samtidig kan man organisere arbeidet som en slags «trakt»-prosess, der man begynner med smalere eller mindre komplekse datasett, og så øker tilgangen eller detaljnivået dersom det blir tydelig at det er nødvendig. På den måten unngår man å bli sittende med mer informasjon enn man faktisk har behov for.

Flere deltakere bemerket at jurister i enkelte sammenhenger kan velge en mer forsiktig linje og svare nei dersom de opplever at prosjektet vil bruke et for stort eller uoversiktlig datasett. Dersom det ikke tydelig fremgår hvorfor et omfattende datagrunnlag er nødvendig, kan det virke tryggere å begrense eller avvise forespørsler i stedet for å utnytte handlingsrommet fullt ut. Dette fikk flere til å påpeke at man trenger jurister som også er villige til å utforske mulighetene, og ikke bare se begrensningene. Et slikt samarbeid mellom tekniske, etiske og juridiske perspektiver kan være avgjørende for å komme fram til gode løsninger der personvernet ivaretas, samtidig som man får utnyttet data på en hensiktsmessig måte.

Datavarehus ble trukket frem som en mulig, men ofte utfordrende løsning, fordi man da samler store mengder data på ett sted. Dette kan gi bedre oversikt og tilrettelegge for analyser, men det fører også med seg en risiko for at man mister kontroll over hvem som får tilgang til hvilken informasjon. Samtidig blir sletting av data mer komplekst: man må forsikre seg om at slettingen skjer overalt der dataene har blitt lagret eller brukt, ikke bare i én database. Det ble påpekt at dette kan være krevende, særlig i større organisasjoner med mange ulike systemer og rutiner.

Risikoen ved automatiserte beslutninger sammenlignet med menneskelige feilvurderinger fikk også en del oppmerksomhet. Noen pekte på at alternativkostnadene ved å si nei til datadrevne løsninger kan være høye, spesielt dersom dette hindrer innovasjon, effektivisering eller bedre innsikt i store datamengder. Samtidig er det en utbredt oppfatning at «blackbox»-modeller kan føre til usikkerhet: når man ikke forstår hvordan algoritmene kommer fram til sine konklusjoner, blir det lettere å avvise

dem for sikkerhets skyld. Dette forsterker behovet for å øke kompetansen om slike modeller og kunne dokumentere hvordan de virker, slik at avgjørelser fattes på et velinformert grunnlag.

Til tross for utfordringene som ble beskrevet, var det en enighet om at dataminimering i seg selv ofte er håndterbart gjennom ansvarlige metoder, iterativ tilnærming og tett dialog med jurister. Det ble snarere fremhevet at det er manglende samspill mellom tekniske og juridiske roller som kan føre til unødvendig skepsis eller strenge begrensninger, særlig i situasjoner der muligheten for å veie nytten mot risikoen ikke er tilstrekkelig utforsket. Dersom man klarer å etablere en felles forståelse av både tekniske og juridiske aspekter, og samtidig anerkjenner at risikobildet endres gjennom prosjektets levetid, vil man i større grad kunne realisere potensialet i datadrevet utvikling. Dette fordrer dog en bevissthet rundt hvilken informasjon som trengs på hvilket tidspunkt, og en klar plan for hvordan man skal dokumentere behovet for og bruken av data underveis.

Alt i alt tegner diskusjonene et bilde av prosesser som er dynamiske og krevende, men også fulle av muligheter for å finne en god balanse mellom innovasjon og personvern. Ved å bygge bro mellom tekniske og juridiske perspektiver, være villig til å justere prosjektets omfang etter hvert som man får mer innsikt, og ta i bruk løsninger som pseudo-anonymisering eller syntetiske datasett, kan man etter beste evne unngå å samle inn mer data enn nødvendig. Videre ser mange det som sentralt at man styrker forståelsen av hvor grensen går mellom rettslige krav og mulighetsrommet for praktiske løsninger, slik at jurister og fagpersoner sammen kan bidra til mer effektive, men også mer bevisste, datadrevne beslutninger.

5. Utfordringer knyttet til kontekstualisering og sammenstilling av data.

Bakgrunn: *Dersom man skal bruke data fra ulike kilder, eller bruke data på nye måter, kan det kreves arbeid med å forstå og organisere dataene. I industri-kontekst kalles dette ofte data-kontekstualisering og består i at man definerer relasjonen mellom ulike datatyper, legge til ekstra informasjon m.m., noe som er ressurskrevende. Hva er erfaringene med dette?*

Oppsummering innspill: *Diskusjonene viste at hovedutfordringen med kontekstualisering og sammenstilling av data er manglende felles begrepsapparat, varierende datakvalitet og usikkerhet rundt eksterne data. Ulike definisjoner av sentrale begreper skaper utfordringer for harmonisering på tvers av etater. Mange aktører jobber fortsatt med å strukturere og dokumentere egne data, noe som er en nødvendig forutsetning før avansert analyse og KI-utnyttelse. Kontekstualisering ble trukket frem som avgjørende for å sikre korrekt bruk av data, og etterprøvbarhet ble fremhevet som essensielt for sporbarhet og kvalitetssikring. Det var enighet om at tydelig dokumentasjon og felles definisjoner er nødvendig for videre digital utvikling og samarbeid.*

Diskusjonene kretset i stor grad rundt de praktiske utfordringene med å samkjøre og gjenbruke data når ulike systemer eller virksomheter opererer med ulike formater og begrepsapparat. Betydningen av entydige definisjoner og god dokumentasjon ble trukket fram som et stadig mer presserende behov, særlig fordi samme begrep kan romme ulike meninger hos ulike etater. Inntektsbegrepet kan for eksempel ha én definisjon i Skatteetaten og en annen i NAV, noe som gjør det tidkrevende å harmonisere data og sikre at de faktisk kan brukes på tvers. Det kreves ofte både en gjennomgang felt for felt og et kontinuerlig arbeid med begrepskataloger, metadata og strukturerte beskrivelser. Selv om

en slik manuell prosess kan oppleves ressurskrevende, har den også en viktig funksjon i å begrense datamengden underveis, ettersom man da sorterer vekk informasjon som ikke er relevant.

Flere uttrykte skepsis til både selvregistrerte data og informasjon fra tredjepartsleverandører. Det er ikke gitt at en bruker gir den informasjonen man trenger, enten på grunn av manglende forståelse eller fordi vedkommende velger å holde noe tilbake, og det er heller ikke sikkert at tredjepartsdata er kvalitetssikret på en måte som passer bruksbehovene i en bestemt virksomhet. Noen beskrev at mange organisasjoner befinner seg i en fase der de fremdeles er i ferd med å strukturere og organisere eget datagrunnlag, og at de dermed ikke føler seg klare for mer avanserte kunstig intelligens-prosjekter. Arbeidet med å få kontroll på kildene og skape et felles begrepsgrunnlag ses på som en forutsetning for å kunne gå videre til mer automatiserte og prediktive analyser.

Behovet for kontekstualisering ble fremhevet gjentatte ganger. Uten et felles vokabular blir det vanskelig å avgjøre i hvilke kontekster bestemte data bør eller kan brukes, og hvem som i praksis har kompetanse til å foreta en slik vurdering. Enkelte pekte på usikkerheten som oppstår når data kommer fra utenlandske aktører, der rettslige og kulturelle rammer kan avvike betydelig fra det man er vant til, eller når data stammer fra mer uklare kilder – for eksempel lekkasjer. Det ble også problematisert at data fra ulike domener kan være krevende å kombinere, ettersom meningsinnholdet lett forvitrer når man «rører sammen» datasettene uten tilstrekkelig innsikt i opprinnelig kontekst. Derfor mente flere at det er nødvendig å kryssevaluere opplysninger som kommer utenfra, samt etablere datasett som tydelige og avgrensede «produkter» med dokumentert formål, kvalitet og definisjoner.

Et annet tilbakevendende poeng var manglende samordning av begreper i offentlig sektor, der man ofte opplever at det ikke finnes felles koder eller presise definisjoner av sentrale felter. Dette forsterkes av at data brukes på ulike måter i ulike systemer: man kan stille strengere krav til nøyaktighet i et system som gir umiddelbare vedtaksresultater, enn i et system som først og fremst benyttes til statistiske analyser. Siden data er «levende» og kan ha flere bruksområder, blir kontekstualisering og dokumentasjon avgjørende for å unngå misbruk eller misforståelser. Mange påpekte at når man skal trekke ut data, enten fra et system eller et datavarehus, må den som gjør selve uttrekket forstå hensikten grundig og ha nok kompetanse til å velge riktige variabler og presisjonsnivå. Etterprøvbarhet i denne prosessen ble beskrevet som særlig viktig, ikke minst for at man i etterkant skal kunne spore hvordan, hvorfor og av hvem et gitt datasett ble opprettet eller endret.

For mange vil det derfor være svært nyttig å ha mer konkret informasjon om hva som faktisk finnes av data, ettersom dette legger til rette for en fruktbar dialog mellom de som bestiller eller trenger data, og de som faktisk forvalter eller henter dem ut. Det handler om å forstå både hvorfor man trenger visse typer informasjon, og hva slags nøyaktighetsnivå som er tilstrekkelig. Flere la vekt på at det ofte er avgjørende å «snakke samme språk» som den man samarbeider med, særlig i større organisasjoner hvor ansvaret for data er fordelt på ulike enheter. Å finne rett person å kontakte og faktisk ha et språk man begge kan enes om, kan i seg selv være halve jobben.

Samlet sett ga diskusjonene et tydelig bilde av at et felles begrepsapparat, klar dokumentasjon og en bevisst holdning til kontekst og bruksområder er fundamentet for all videre arbeid med data, særlig i en tid der mulighetene innen analyse og automatisering er raskt voksende. Selv om mange aktører fremdeles ikke ser seg helt klare for avanserte KI-prosjekter, blir det også understreket at nettopp en solid og gjennomsiktig fundamentering i grunnleggende datastrukturer og definisjoner er en nødvendig forutsetning for å kunne ta steget videre.

6. Utfordringer knyttet til automatisering versus beslutningsstøtte.

Bakgrunn: *Det kan være ulike måter å bruke for eksempel et resultat av en prediksjonsmodell (en skår) eller fra en risikomodell (sannsynlighet for avvik), og det er forskjell på en helautomatisert beslutning og der resultatet inngår i beslutninger på andre måter. Hvordan bruker etatene slike (del-)resultater i saksbehandlingen? Hvilke etiske momenter og juridiske regler bør man være oppmerksom på i disse situasjonene?*

Oppsummering innspill: *Diskusjonene viste et tydelig skille mellom helautomatiserte beslutninger og bruk av KI som beslutningsstøtte i offentlig sektor. De fleste etater benytter i dag KI til å identifisere mønstre, analysere store datamengder og gi saksbehandlere støtte, men ikke til å fatte endelige vedtak. Flere understreket at beslutningsstøtte krever andre juridiske vurderinger enn helautomatisering, siden mennesker fortsatt har ansvar for den endelige avgjørelsen. Usikkerhet knyttet til KI-modellers forklarbarhet og feilmarginer gjør at helautomatiserte løsninger sjelden er aktuelle, særlig i høyrisikoområder som myndighetsutøvelse. Likevel ble det pekt på at både helautomatiserte og delvis automatiserte prosesser må ha tydelige rammer for hvordan KI-resultater skal tolkes og brukes, slik at teknologien blir et verktøy for bedre beslutninger, ikke en ukritisk erstatning for menneskelig vurdering.*

Diskusjonene pekte på at anvendelse av kunstig intelligens kan være særlig krevende i domener med høy risiko, slik som ved myndighetsutøvelse eller andre oppgaver der konsekvensene av feil kan være store. Samtidig ble det understreket at ikke all intern virksomhet nødvendigvis faller inn under denne kategorien. For noen vil KI først og fremst kunne brukes til å identifisere mulige spor eller indikatorer, mens en manuell oppfølging uansett er påkrevet for å sikre riktige vurderinger. Dermed oppstår spørsmålet om hvilke rettslige konsekvenser KI-bruk kan ha, og hvordan man bør avveie nytten mot både kostnader og mulig risiko. Flere stilte seg også spørsmålet om hvorvidt man aksepterer større feilmarginer fra mennesker enn fra automatiserte systemer, og om KI kan ende opp med å vektlegge bestemte variabler mer enn det som faktisk gjenspeiler virkeligheten. Derfor fremsto det viktig å skille mellom ulike bruksområder: et verktøy som kun gir veiledning, krever en annen form for risikohåndtering enn et system som skal fatte bindende vedtak.

Det kom også frem at man i mange organisasjoner fremdeles ikke har kommet til et modenhetsnivå hvor hel-automatisert saksbehandling er aktuelt. I stedet blir KI oftest benyttet til beslutningsstøtte, for eksempel ved at saksbehandleren får støtte i å forstå kundens henvendelse eller hjelp til å se sammenhenger i store datamengder. Noen beskrev «dulting» som en form for myk påvirkning, men samtidig la flere vekt på at KI som sådan ofte fungerer best som et verktøy som identifiserer tendenser, mens mennesker gjør de endelige vurderingene. Dette reflekteres særlig i situasjoner der man skal forklare avgjørelser i etterkant. Enkelte typer algoritmer eller modeller gjør det svært vanskelig å gi en intuitiv forklaring på hvorfor systemet har kommet fram til et bestemt resultat, og denne mangelen på transparens gjør at man i praksis nøler med å gi KI mer formell beslutningsmyndighet. Samtidig mente noen at mer automatisering kunne være til nytte innen styringsprosesser og HR-prosesser, men at den nåværende utnyttelsen av KI i stor grad er begrenset av at man ikke har tilstrekkelig teknisk eller organisatorisk modenhet.

Et annet sentralt tema var hvordan man over tid kan opparbeide større tillit til KI. Det ble pekt på at teknologien i utgangspunktet kan prosessere langt mer informasjon enn det mennesker klarer, noe som igjen kan gi mer treffsikre eller effektive resultater, men bare dersom man har tillit til dataenes

kvalitet og algoritmenes robusthet. Denne tilliten fordrer at man vet hvilke konsekvenser verktøyene kan få, og at man har jobbet med å definere rammer som håndterer blant annet bias og feilmarginer. Mye av usikkerheten som ble beskrevet, var knyttet til at dagens juridiske rammer ikke alltid er tilpasset en virkelighet der KI brukes til ulike typer støtte i forvaltningsbeslutninger. Flere påpekte at lovverket kan være utydelig på hvilke vilkår som må oppfylles for at automatisering skal være lovlig og forsvarlig, og at dette i praksis fører til nøling med å ta i bruk teknologien i stor skala.

I forlengelsen av dette kom det fram at veien mot å bli en mer data- og KI-drevet organisasjon ofte er smertefull. Man må investere betydelige ressurser i å få oversikt over eksisterende datasett, kvalitetssikre dem og skape rutiner og systemer som er i stand til å håndtere store datamengder. Dersom man ikke har tilstrekkelig kontroll på dataene, er man heller ikke reelt «datadrevet». Noen stilte spørsmål ved om bruken av KI i beslutningsstøtte-sammenheng kan føre til samme form for skjevheter som en helautomatisert løsning ville hatt, hvis man ikke er bevisst på hvordan anbefalingene påvirker menneskers vurderingsevne. Dette illustrerer et behov for ikke bare å bygge tekniske modeller, men også å etablere klare retningslinjer, prosedyrer og kompetanse for hvordan forslag fra KI skal forstås og brukes i praksis. Siden det ofte er de samme underliggende mekanismene som skaper statistiske skjevheter, er det avgjørende at bruken av KI – enten det dreier seg om støtte eller en mer automasjonsorientert prosess – får et grundig regelverk og skikkelige kontrollmekanismer.

Samlet sett vitner diskusjonene om et landskap der KI fremdeles i hovedsak tas i bruk som et analyse- og beslutningsstøtteverktøy. Mange er interesserte i muligheten for å automatisere flere oppgaver, men enn så lenge blir slike ambisjoner dempet av usikkerhet rundt juridiske føringer, mangel på teknisk og organisatorisk modenhet og spørsmålet om hvor stor feilmargin man kan tillate. Når man likevel ser at KI kan gi gevinst i form av tidssparing, bedre innsikt og potensielt mer kvalitetssikrede vurderinger, oppstår det et åpenbart behov for klarere retningslinjer og en felles forståelse av hva som er formålstjenlig bruk. Diskusjonene illustrerte derfor en overgangsfase, der mange erkjenner at de ikke er klare for helautomatiserte prosesser, men samtidig innser at en vellykket overgang til mer avansert KI-bruk krever både robust teknologisk infrastruktur, juridisk avklaring og en plan for å bygge tillit internt og eksternt.

7. Muligheter ved etablering av data-infrastruktur, intern kompetanse og kapasitet.

Bakgrunn: *Å etablere en god datainfrastruktur og dataforvaltning er en krevende oppgave. Hva har de ulike etatene gjort som kan være verdifullt å dele med de andre.*

Oppsummering innspill: *Diskusjonene viste at etablering av en robust datainfrastruktur i offentlig sektor er en kompleks, men avgjørende oppgave. Erfaringene fra Helfo og Skatteetaten illustrerte både utfordringer og løsninger knyttet til datavarehus og skybaserte plattformer. Helfo fremhevet viktigheten av å bygge intern kompetanse for å unngå avhengighet av eksterne konsulenter, mens Skatteetaten understreket behovet for tydelige eierskapsstrukturer og juridiske avklaringer ved overgang til skybaserte løsninger. Flere deltakere påpekte at utfordringer som tilgangsstyring, sporbarhet og ansvarliggjøring går igjen i hele offentlig sektor, og etterlyste felles rutiner og praksis for å spare tid og ressurser på tvers av etater. En sentral innsikt var at tett samarbeid mellom teknologer og jurister, samt deling av erfaringer og risikovurderinger, kan bidra til mer effektiv etablering og forvaltning av data-infrastruktur i offentlig sektor.*

Case 1: Helfo

Helfo fortalte om hvordan de har etablert et internt datavarehus (kalt «Nøkkelinfo») og utviklet ulike analyseverktøy, som Power BI-rapporter, standardspørringer og Excel-modeller. Arbeidet har strukket seg over lang tid, særlig fordi det tok omfattende ressurser å avklare GDPR-etterlevelse. De har også møtt utfordringer fordi noen av databasene eies av andre aktører, noe som skaper problemer når det er behov for feilrettinger. Helfo kan i slike tilfeller ha et annet innsyns- eller kvalitetssikringsbehov enn eieren av databasen.

Prosjektet var delvis finansiert av Norges Forskningsråd gjennom innovasjonsprosjektet «Lærende Kontroll», og hovedsakelig utført av interne krefter. I løpet av prosessen ble det nødvendig å engasjere eksterne jurister, som ikke var kjent med verken systemene eller praksisen i Helfo. Dette skapte et behov for mange runder med avklaringer om formål, databruk og sikringsmekanismer. Når brukerne samtidig slet med å konkretisere sine behov, ble det viktig å etablere en felles forståelse av både systemene og begrepsbruken.

Erfaringene i etterkant har vist at det gir store gevinster å ha en jurist tett på analytikerne. Gjennom samarbeidet kan juridiske krav og faglige behov raskt oversettes og avstemmes, samtidig som man enklere kan avgjøre hvilke problemstillinger som bør håndteres lokalt, og hvilke som bør løftes opp i organisasjonen. Et annet positivt resultat av prosjektet er at Helfo har bygd opp intern kompetanse og dermed unngått å bli avhengig av eksterne konsulenter. De kan derfor «tweake» systemet fortløpende, ofte på kort varsel, og på den måten videreutvikle løsningene i takt med endrede behov.

Helfo påpeker at det er flere interessante aspekter å diskutere med tanke på prosjektet. For det første kan man vurdere fordeler og ulemper ved standardløsninger versus mer egenutviklede løsninger. I tillegg reiser datavarehus-strukturen spørsmål om hvordan man skal styre ansattes tilgang til informasjon, samt hvordan man best kan spore bruk og sikre kontroll i et miljø der store datamengder er samlet på ett sted.

Slike problemstillinger er noe Skatteetaten også har lang erfaring med, og de ble derfor invitert til å kommentere Helfos case i lys av egne prosesser. Skatteetaten fortalte at de har hatt datavarehusløsninger i 20 år, men at de nå er i ferd med å ta i bruk en ny, skybasert plattform. Denne overgangen har krevd en grundig prosess med risikovurderinger, og prosjektet har involvert en stor grad av konsulentbruk. Nå går det over i en ren forvaltningsfase, der hovedoppgaven blir å etablere god og effektiv bruk av løsningen, samt løfte kompetansenivået blant de ansatte som skal håndtere mer avanserte analyser.

Case 1: Diskusjon

Bakgrunnen for overgangen til en skybasert løsning i Skatteetaten ble forklart med behovet for bedre sikkerhet, sporbarhet og fleksibilitet. Ved å gå over til skyen ønsker etaten å utvikle såkalte «dataprodukter» som gir tydelig definerte utsnitt og strukturerte datasett. Dette forutsetter god intern orden, klare eierskapsroller og juridiske avklaringer om hvordan ulike dataprodukter kan og bør brukes.

I den påfølgende samtalen oppstod enighet om at mange av disse vurderingene vil være allmenngyldige for hele offentlig sektor. Flere stilte spørsmålet om det ikke burde være mulig å enes om felles rutiner og praksis på tvers av etater og virksomheter, siden mange opplever de samme hindringene med ansvarliggjøring, tilgangsstyring og lovmessige forankringer. Tanken er at den omfattende kartleggingen og risikovurderingen som for eksempel Skatteetaten har gjennomført, kan fungere som en rettesnor for andre som vurderer tilsvarende skykonsepter. Dermed kan organisasjoner spare tid og krefter på å løse de samme problemstillingene hver for seg, og heller bygge videre på eksisterende erfaringer og praksis.

8. Muligheter for bruk av egne data på nye måter.

Bakgrunn: *Å bruke data på nye måter kan medføre behov for juridiske avklaringer, og være krevende rent praktisk, med tanke på å bygge pålitelig innsikt gjennom nye metoder. Hva har de ulike etatene gjort som kan være verdifullt å dele med de andre?*

Oppsummering innspill: *Diskusjonene viste at bruk av egne data på nye måter gir store muligheter, men også utfordringer knyttet til juridiske rammer, datakvalitet og risiko. Skatteetaten demonstrerte hvordan de kombinerer data fra ulike systemer for å utvikle nye analyseverktøy, men påpekte at fragmenterte lovverk og taushetspliktbestemmelser gjør dette krevende. Helfo delte erfaringer med et «analyserom» for avidentifiserte data, men fant at klassiske regelbaserte modeller ofte fungerte bedre enn mer avansert KI. Kripos presenterte et KI-prosjekt for identifisering av overgrepsutsatte barn, der juridiske avklaringer tok fire år, noe som understreket behovet for en raskere «fast track»-prosess for KI-bruk i offentlig sektor. Flere aktører etterlyste bedre juridiske rammer for KI-trening og utvikling, samt økt deling av erfaringer og verktøy, som NAVs modell for systematisering av behandlingsgrunnlag. En sentral problemstilling var hvordan man kan bruke data for å identifisere systematiske variasjoner i atferd uten å skape diskriminering eller stigmatisering. Diskusjonene viste at dagens manuelle saksbehandling også kan innebære bias, men uten samme grad av transparens som man etterlyser for KI-baserte løsninger. Det ble derfor understreket at klare retningslinjer og en bevisst vurdering av risiko er nødvendig for å sikre ansvarlig bruk av data i offentlig sektor.*

Case 2: Skatteetaten

Skatteetaten presenterte et prosjekt hvor man forsøkte å kombinere data fra ulike «silo systemer» og verdikjeder, slik at man kunne utvikle nye analyseprodukter på tvers av etatens ansvarsområder, som blant annet innkreving, fastsetting og folkeregisteret. Selv om disse datasettene nå formelt befinner seg under samme etat, har de historisk sett hatt ulike eiere, formål og lovbestemmelser. Dette gjør kombinasjonen av data komplekst. I tillegg er taushetspliktsreglene fragmentert på tvers av lovverk, slik at etaten må vurdere ulike formålsbegrensninger og unntak hver for seg. Prosessen med KI-utvikling blir dessuten mer krevende av at man ikke alltid vet hvilke data som er nødvendige før man starter. Å lage detaljerte planer på forhånd kan være lite hensiktsmessig, siden formålet med KI-metoder ofte er å oppdage mønstre man ikke kjente til. Det ble derfor argumentert for behovet for mer fleksible rutiner, der man kan ta noen utforskende runder uten å få permanent eller ubegrenset tilgang til hele datasettet.

Case 2: Diskusjon

I forlengelsen av Skatteetatens presentasjon ble det understreket hvor lett det er å tenke at man «bare» kan samle inn alt av data når man skal arbeide analytisk. Samtidig er det ikke slik at man alltid har tilgang på alt som teoretisk kunne vært nyttig. I noen sammenhenger har aktørene insentiver til å rapportere flest mulig opplysninger (som ved refusjonskrav til Helfo), mens i andre sammenhenger er insentivet til å rapportere svakt eller til og med negativt (som ved skatt og merverdiavgift). Resultatet blir at man ofte sitter igjen med et skjevt bilde av situasjonen hvis man kun baserer seg på tilgjengelige data. De som diskuterte disse forholdene var enige om at man derfor må være bevisst hvor hullene i datagrunnlaget kan oppstå.

Helfo delte sine erfaringer med et eget «analyserom», der de arbeider med avidentifiserte data og kan legge inn visse «forstyrrelser» slik at det ikke skal være mulig å spore enkeltpersoner. De så i utgangspunktet potensial for KI-baserte analyser, men erfarte at datasettet ofte ikke ga god nok prediksjonsverdi. De kom derfor tilbake til mer klassiske, regelbaserte modeller for enkelte formål. Samtidig mente de at uklarhet i regelverk og hjemmelsgrunnlag skapte usikkerhet, blant annet fordi lover og regler ble oppfattet som litt «gammeldagse». For å finne ut hvor mye rom man faktisk har, har de sett til Skatteetatens praksis og erfaringer innen kontroll.

Case 3: Kripos

Kripos presenterte et KI-prosjekt hvor de brukte søke- og filtreringsmetoder i arbeidet med identifisering av barn utsatt for seksuelle overgrep.[3] Arbeidet viste seg å være omfattende, særlig fordi juridiske avklaringer rundt automatisert teknologi i politiet er komplekse og ofte tidkrevende. Det tok fire år fra prosjektet startet til det var tilstrekkelig avklart. Gjeldende passlov åpner for utlevering av opplysninger ved forbrytelser med en strafferamme på over seks måneder, og Kripos valgte å avgrense bruken til nettopp slike saker. KI-delen består i å filtrere eller fremheve enkelte treff, men den endelige identifiseringen foregår fremdeles manuelt. På denne måten begrenser Kripos omfanget av personer som kan bli uberettiget undersøkt, og gjør samtidig etterforskningsprosessen langt mer effektiv.

Prosjektet avdekket også hvordan ulike faggrupper i politiet har forskjellige perspektiver på risiko. Noen er mest opptatt av personvern eller informasjonssikkerhet, andre av mulige omdømmetap dersom politiet bruker verktøyet feil. For å håndtere utfordringene har Kripos – og politiet for øvrig – begynt å ansette «tech-jurister». Disse skal være tettere på teknologiutviklingen og kunne vurdere handlingsrommet med et bredere forståelsesgrunnlag enn det man får gjennom et rent personvernfokus.

Case 3: Diskusjon

I diskusjonen rundt Kripos-prosjektet påpekte flere at lignende situasjoner oppstår i andre sektorer, og at man ofte savner et raskere, mer standardisert løp for juridiske avklaringer. Noen trakk frem NAVs arbeid med å definere behandlingsgrunnlag og -formål i en «data-markedsplass» som et eksempel på en systematisk tilnærming. Skatteetaten nevnte at de har noe lignende, men i enklere form. De etterlyste muligheter for å dele denne type verktøy, for eksempel i form av en «DPIA-bank» eller en felles «juridisk fast track»-prosess på tvers av etater.

Et konkret problem Kripos står overfor, er at politiregisterloven per i dag forbyr bruk av skarpe data til test og utvikling. Dermed kan man i utgangspunktet ikke trene KI-modeller på autentiske skarpe opplysninger, samtidig som EU-forpliktelser krever grensekontroll med ansiktsgjenkjenning. Flere pekte på at dette er et eksempel på at nasjonale regler kan være «for snevre», og ikke oppdaterte i lys av nye teknologiske muligheter og internasjonale krav. Det ble også fremhevet at man generelt bør bli flinkere til å beskrive hva man risikerer å tape om man sier nei til KI-initiativer, ikke bare hva man risikerer å tape ved å si ja. Digitale ressurssentre som DigDirs veiledningstjeneste for etisk bruk av KI ble omtalt positivt, men flere mente at slike enheter burde styrkes ytterligere, både med hensyn til teknologisk og juridisk kompetanse.

[3] https://www.nrk.no/nordland/omegle-saken_-kripos-brukte-eget-verktoy-for-a-identifisere-190-barn-i-oergrepssak-1.16901885

Case 4: Diskriminering og stigmatisering

Det siste caset tok utgangspunkt i spørsmålet om hvordan man kan undersøke systematiske variasjoner i atferd mellom ulike grupper, uten å havne i en situasjon der man diskriminerer eller stigmatiserer. Bakgrunnen for problemstillingen var blant annet at Skatteetaten hadde avdekket utbredt skatteunndragelse blant enkelte grupper fastleger.[4] Det kunne teoretisk være relevant også for refusjonsutbetalinger fra Helfo, men Helfo opplyste at de av flere grunner ikke har hatt noe særskilt søkelys på denne gruppen. Dette illustrerer imidlertid en mer generell utfordring: Kan man se på potensielle mønstre knyttet til eksempelvis etnisitet eller kjønn på en etisk forsvarlig måte? Og hvordan skiller man mellom å «kartlegge» slike forskjeller og å la en KI-modell «lære» at visse grupper er mer risikoutsatt?

Skatteetaten nevnte at de har reflektert over denne typen variabler i sin KI-policy, men at de ikke har noe eget organ for å diskutere slike spørsmål. Dermed blir håndteringen ofte case-basert, uten at det nødvendigvis finnes felles retningslinjer. Dette står i kontrast til ønsker om at man burde ha et mer strukturert rammeverk for etiske dilemmaer, særlig når potensielt sensitive kjennetegn kommer inn i bildet.

Case 4: Diskusjon

I den påfølgende diskusjonen ble det påpekt at jurister av og til tar med etiske vurderinger og omdømmerisiko lenger inn i den juridiske tolkningen enn det rent formelle krever. Noen mente at det kan være positivt at etaten tenker helhetlig, mens andre understreket at det er viktig å vite hvor jusen slutter og hvor etikken begynner. Flere trakk frem «VG-testen» eller «rødmetesten» som en måte å vurdere om det man gjør, kan forsvares offentlig. Samtidig ble det pekt på at passivitet også kan svekke tillit: publikum forventer at myndighetene gjør en effektiv jobb, og det kan slå negativt ut dersom man unngår tiltak som faktisk kan avdekke uønsket atferd.

Spørsmålet om «etnisk profilering» kom raskt opp. I politiet er denne praksisen eksplisitt forbudt, og flere anså det som lite sannsynlig at man ville innføre en ordning der man aktivt bruker etniske kjennetegn i en automatisk analyse. Samtidig kan dagens manuelle saksbehandling i praksis være utsatt for like stor risiko for bias, uten at det dokumenteres. Enkelte stilte spørsmål ved om en mer strukturert, transparent prosess faktisk kunne redusere faren for feilslutninger. Dette viser at diskusjonen om KI-etikken også henger tett sammen med dagens praksis, der uformelle «magefølelser» og lave dokumentasjonskrav kan føre til ubevisste skjevheter.

Mot slutten av diskusjonen ble det igjen poengtert at mye av lovverket – både i politiet og andre etater – i liten grad er tilpasset treningsbehovene KI-metoder krever. Flere opplevde hjemlene som «for korte», siden de åpner for innsamling og bruk av data, men ikke nødvendigvis for å trene opp KI-modeller. Dermed kan det ofte være nødvendig med lovendringer eller nye retningslinjer som gjenspeiler dagens teknologiske muligheter.

[4] https://www.dn.no/kriminalitet/erik-nilsen/skatteetaten/skattekrim/skatteetaten-fastleger-unnlot-a-oppgi-162-millioner-kroner-i-inntekter/2-1-1692043

9. Muligheter for deling og mottak av data fra eksterne samarbeidspartnere

Bakgrunn: *Å ta imot data fra eksterne kilder kan være utfordrende, for eksempel på grunn av ulike dataformat/begreper, utfordringer med kvalitet og sporbarhet, men også med mottakerens kapasitet og evne til å ta imot, prosessere og nyttiggjøre seg det som kommer. Hva har de ulike etatene gjort som kan være verdifullt å dele med de andre?*

Oppsummering innspill: *Diskusjonene viste at utfordringer med eksterne data ofte skyldes kapasitet, kvalitet og juridiske begrensninger. Bankenes rapportering av mistenkelig økonomisk aktivitet resulterer i store datamengder som Økokrim har begrenset kapasitet til å håndtere, særlig fordi mange meldinger er tekstbaserte og vanskelig å systematisere. Bedre merking og kategorisering ble foreslått som tiltak. Skatteetaten delte erfaringer fra CRS-ordningen, der vellykket datahåndtering har krevd grundig forarbeid og internasjonalt samarbeid, men taushetsplikt og regelverk begrenser fortsatt effektiv deling. Økokrims pilot viste at utfordringen ikke lå i innsendingen av data, men i ustrukturerte rapporteringsformater. Det ble derfor understreket at tekniske løsninger og rapporteringsformat bør utvikles i samarbeid med IT-eksperter for å sikre at dataene blir mer systematiske og maskinlesbare fra start.*

Siste diskusjonen omhandlet bankenes rapportering av mulig økonomisk kriminalitet. Årlig sender norske banker om lag 5000 meldinger om mistanke om manglende skattebetaling (såkalte MT-rapporter). Økokrim har begrenset kapasitet til å behandle dette enorme volumet, noe som fører til at mange rapporter ikke blir grundig fulgt opp. I et pågående prosjekt hos Skatteforsk («Tidlig intervensjon ved skatterelaterte mistenkelige transaksjoner») (I samarbeid med BNbank og NTNU) prøver forskere å se om målrettet veiledning (nudging) fra banken kan ha en forebyggende effekt. Siden Økokrim primært prioriterer de mest alvorlige meldingene, reiser dette spørsmålet om man kan øke kapasiteten gjennom bedre merking eller kategorisering av saker, slik at de enklere kan sorteres og analyseres. Per i dag inneholder mange meldinger lange tekstfiler, som er ressurskrevende å vurdere.

Et annet sentralt tema er at meldingene som sendes inn, i liten grad deles videre med andre myndigheter som Skatteetaten eller Arbeidstilsynet. Det ble stilt spørsmål ved om antihvitvaskingsloven gir rom for mer effektiv informasjonsdeling, eventuelt om loven bør tolkes eller revideres for å gjøre bankene til en mer aktiv førstelinje i bekjempelsen av økonomisk kriminalitet. Bankene investerer tusenvis av årsverk i å fremskaffe data, men mye blir verken analysert eller utnyttet tilstrekkelig. En annen problemstilling er at det ikke er lov å dele informasjon mellom banker, på grunn av faren for konkurransevridning. Konsekvensen er at aktører med uredelig hensikt kan «shoppe rundt» mellom ulike banker.

Diskusjon

Skatteetaten kommenterte sin erfaring med å håndtere store datamengder, for eksempel via den automatiske utvekslingsordningen CRS (Common Reporting Standard). Denne ordningen, innført av OECD, skal bekjempe skatteunndragelse og unngå dobbeltbeskatning ved at formuesopplysninger automatisk rapporteres på tvers av landegrenser. Skatteetaten har bygd opp både teknisk og faglig kapasitet for å håndtere slike data, og trekker frem at suksessen skyldes et omfattende forarbeid, politisk vilje og internasjonal koordinering. Likevel gir ikke dette noen enkel løsning på alle juridiske begrensninger, spesielt når det gjelder taushetsplikt og hvordan ulike nasjoners regelverk samspiller.

Det kom også opp at DSOP (Digital Samhandling Offentlig Privat) er et samarbeidsinitiativ hvor man utforsker deling av data mellom offentlig og privat sektor. Blant annet har Økokrim og Skatteetaten etablert et senter med kapasitet til å håndtere rundt 80 saker per år, men det er fremdeles et godt stykke unna å dekke hele rapportvolumet fra bankene. I tillegg ble det poengtert at mindre banker og enkelte eiendomsmeglere ofte sender inn rapporter med tekstlige beskrivelser, noe som er vanskelig å systematisere.

En pilot i Økokrim viste at man først trodde flaskehalsen lå i å forbedre selve innsendingsløsningen. Etter hvert innså man at utfordringen i større grad skyldtes de tekstlige formatene, der informasjonen ikke var lagt til rette for automatisert analyse eller enkel kategorisering. Fra teknisk hold ble det anført at man bør få IT-personell og skjema-utviklere med i konseptfasen, slik at man allerede ved utforming av rapporteringssystemene legger til rette for brukervennlig, gjenbrukbar og maskinlesbar data. Det ble konkludert med at hvis slike hensyn tas tidlig, kan man unngå at hovedtyngden av arbeidet blir manuelle vurderinger i etterkant.

10. Avsluttende diskusjon: Ønsker, innspill og prioriteringer fra deltakerne

I den avsluttende delen av workshopen fikk deltakerne anledning til å oppsummere erfaringer og identifisere viktige prioriteringer for videre arbeid med databruk i offentlig sektor, utover det som var dekket under tidligere tema. Diskusjonene resulterte i en rekke konkrete forslag, blant annet behovet for økt tverrfaglig samarbeid, bedre tilrettelegging for kodedeling, og mer systematisk deling av eksisterende vurderinger og strategier. Under oppsummerer deltakernes innspill i punktform, med fokus på tiltak som kan bidra til mer effektiv, sikker og bærekraftig databruk i offentlig sektor.

- **Krev tverrfaglighet i prosjekter**: Deltakerne oppfordret til å involvere både teknologer, jurister og framtidige brukere i hele utviklingsløpet.

- **Arena for kodedeling**: Det ble foreslått å etablere en felles plattform (eksempelvis GitLab) for deling og gjenbruk av kode på tvers av offentlige etater, samt å danne et aktivt nettverk for kunnskapsutveksling. Nettverket bør være på et lavt nok nivå i organisasjonene, for å treffe de som jobber med dette. For å stimulere til deling kan man vurdere en årlig pris eller lignende insentiver.

- **Gjør vurderinger og rapporter lettere tilgjengelig**: Deltakerne ønsket et grensesnitt som gir enklere innsyn i eksisterende rapporter, artikler og vurderinger, for eksempel fra Datatilsynets sandkasse. Videre foreslo de å bruke KI for å søke i og systematisere store mengder dokumentasjon, slik at erfaringer og praksis blir mer tilgjengelig og gjennomskinnelig.

- **Deling av nyttige strategier og kode**: Gruppen mente at det burde legges til rette for at eksempelvis NAVs etterlevelsesstrategi og annen åpen kode kan deles mer systematisk, slik at flere kan dra nytte av disse erfaringene.

- **Behov for flere tech-jurister og oppdatert juristutdanning**: Det ble påpekt at mangelen på juridisk kompetanse innen teknologi er et hinder i mange prosjekter. Utdanningen av nye jurister bør derfor i større grad tilpasses behovene i en mer digitalisert virkelighet.

- **Insentiver for data- og informasjonsdeling**: Det bør innføres konkrete tiltak og belønninger som oppmuntrer til å dele og tilgjengeliggjøre data mellom aktører.

- **Miljø- og bærekraftsperspektiv på KI**: Gruppen foreslo at man bør etablere rammer for når man faktisk skal sette inn KI-løsninger, gjerne med utgangspunkt i miljømål og bærekraftsvurderinger.

- **Felles prinsipiell avklaring**: Det etterlyses klarere retningslinjer og begrepsavklaringer, eksempelvis rundt bruk av skytjenester. Det ble fremhevet at politikkutforming og lovarbeid i større grad bør samkjøres, ettersom disse prosessene i dag er fordelt på flere departementer.

11. Anbefalinger til Digitaliseringsministeren

Formålet med disse workshopene har vært å avdekke konkrete utfordringer som hindrer offentlige etaters bruk av egne data samt å foreslå tiltak som ikke nødvendigvis løser alt, men som kan gjøre prosessene noe enklere. Forfatterne av denne rapporten har basert seg på diskusjonene og innsiktene fra de ulike workshopen og har formulert fire konkrete anbefalinger til Digitaliseringsminsteren som vi diskuterer i mer detalj under. Disse anbefalingene dekker konkrete tiltak som kan møte behovet for bedre intern dataforvaltning, utvikling av juridisk kompetanse innen teknologi, en mer balansert tilnærming til risiko, fjerning av juridiske barrierer, nasjonal samordning av teknologi-jus og etablering av felles plattformer for deling.

Anbefaling 1: Styrk intern informasjonsforvaltning og kompetanse.

For at offentlig sektor skal kunne utnytte data effektivt – enten til analyse, kunstig intelligens eller datadeling – må hver enkelt aktør ha god kontroll på egne data. Dette krever økte ressurser og kompetanse innen intern dataforvaltning.

Det bør også etableres fellesarenaer for erfarings- og kunnskapsdeling, der offentlige aktører kan utveksle beste praksis, juridiske vurderinger og tekniske løsninger. En digital plattform, for eksempel GitLab, kan brukes til å dele kode, etterlevelsesdokumentasjon, eksempler på DPIA-er (inkludert utforskende DPIA-er) og andre dataprosesseringsrammeverk.

I tillegg bør Digitaliseringsdirektoratets rammeverk og verktøy for «orden i eget hus» gjøres bredere kjent og mer brukt. Direktoratet tilbyr allerede faglige arenaer for datadeling og informasjonsforvaltning, samt et nasjonalt ressurssenter for deling og bruk av data. Det bør vurderes om disse tilbudene er tilstrekkelig tilgjengelige, og om de adresserer praktiske behov knyttet til analyse, KI og rapportering. En utvidet rådgivingstjeneste kan være et nyttig tiltak for å støtte aktørene i dette arbeidet.

Anbefaling 2: Styrk samarbeidet mellom jurister og teknologer – få på plass 'tech-jurister'.

For å sikre bedre databruk i offentlig sektor må jurister og teknologer samarbeide tettere. Det er økende behov for jurister med teknologisk innsikt – såkalte 'tech-jurister' – som ikke bare vurderer digitale løsninger gjennom et personvernfilter, men også ser mulighetene og det juridiske handlingsrommet for innovasjon. For å støtte denne utviklingen bør myndighetene jobbe for at juridiske utdanninger styrkes med mer teknologikunnskap, slik at fremtidens jurister er bedre rustet til å håndtere digitaliseringens krav og muligheter.

Samtidig bør dette samspillet gå begge veier: Dataanalytikere og utviklere bør også få grunnleggende juridisk forståelse, spesielt innen personvern og GDPR. Dette vil bidra til mer effektive prosesser, der teknologiske valg kan gjøres på en måte som ivaretar prinsipper som dataminimering og dermed forenkler de juridiske avklaringene.

Ledere i offentlig sektor har også en viktig rolle i å klargjøre ansvarsfordelingen – og skillet – mellom juridiske, tekniske og etiske vurderinger. Jurister bør fokusere på rettslige spørsmål, mens beslutninger om etikk og omdømmerisiko bør ligge hos ledelsen. Dette vil sikre tydeligere roller og mer målrettede vurderinger i utviklingen av digitale løsninger.

Anbefaling 3: Frem en kultur for ansvarlig risikohåndtering.

I offentlig sektor er det ofte en sterk tilbøyelighet til å velge den tryggeste løsningen – å si nei til nye måter å bruke data på. Dette fører til at verdifulle muligheter kan gå tapt. I stedet bør organisasjonene aktivt utforske sitt handlingsrom og ta en mer balansert tilnærming til risiko.

Et viktig element i dette er å vurdere alternativkostnader: Hva går vi glipp av når vi velger å ikke utnytte egne data bedre? To eksempler illustrerer dette tydelig: Politiet har ikke lov til å teste KI-systemer, noe som har ført til at de nå må bruke KI-løsninger pålagt av EU – uten å ha testet dem selv i forkant. På samme måte tok det Kripos fire år med juridiske og organisatoriske avklaringer før de kunne bruke ansiktsgjenkjenning i arbeidet med å identifisere barn utsatt for seksuelle overgrep. Hvor mange saker kunne ha blitt løst raskere om denne teknologien hadde vært tilgjengelig tidligere?

For å skape en mer nyansert risikovurdering må organisasjoner ikke bare spørre «hva kan gå galt?» – men også «hva går vi glipp av hvis vi lar være?» Å utfordre dagens risikoaversjon handler ikke om å fjerne sikkerhetsmekanismer, men om å sørge for at data brukes på en måte som både ivaretar personvern og samtidig gir samfunnsmessig verdi.

Anbefaling 4: Ta nasjonalt ansvar for avklaringer for teknologi-jus.

Digitaliserings- og Forvaltningsdepartementet bør ta ansvar for den nødvendige jus-utviklingen som digitaliserings-strategien krever. Dette innebærer både konkrete og mer overordnede aspekter.

Et sentralt problem er mangelen på tydelige hjemler for utvikling og testing av KI-modeller. Myndighetene bør avklare om dagens brede hjemler for databruk også dekker KI-utvikling, eller om det trengs en presisering. I tillegg bør det vurderes om dagens formålsbestemmelser for ulike datakilder er tilstrekkelige, eller om de må revideres – slik det ble gjort i revisjonen av helselovgivningen i 2022. Taushetspliktsbestemmelser i spesiallovgivningen kan også utgjøre en barriere for datadeling og bruk, både internt i organisasjoner og mellom aktører. Slike spørsmål må ikke behandles isolert, men ses i en større sammenheng for å sikre en helhetlig og balansert tilnærming.

Det ville vært nyttig om det ble formulert retningslinjer for:

- Hvordan skal aktørene forholde seg til spenningen mellom forhåndsdefinerte formål versus utforskende utvikling? Hvilke måter å arbeide på er akseptable?
- Hvordan kan man ta KI i bruk i saksbehandling på en ansvarlig måte?
- Hvordan skal man gjennomføre miljø- og bærekrafts-vurderinger av KI-løsninger?
- Hvordan kan man gjøre profilering av grupper på en ansvarlig måte?
- Utred problemstillinger omkring deling av data: hvem har oppfølgingsansvar om man finner noe, hvordan kan man dele data og varsle uten å avsløre kontrollsettinger, under hvilke forutsetninger kan man dele data fra tredjepart?

Et mer overordnet tiltak er å reforhandle den sosiale kontrakten i lys av muligheter og utfordringer som ny teknologi bringer inn. Ny teknologi skaper både muligheter og utfordringer, og på et overordnet plan er det viktig å ta en bred diskusjon om hvordan vi skal balansere ulike hensyn som personvern, transparens og effektivitet. For eksempel har innføringen av GDPR gjort det mer krevende å utvikle enkelte digitale løsninger, noe som illustreres av uttalelsen: «Vi ville nok ikke hatt digital skattemelding om vi ikke hadde laget den før GDPR kom.» For å unngå at strenge regler hindrer innovasjon som faktisk kan gagne både samfunnet og individet, bør det legges til rette for åpne debatter og offentlig diskusjon om hvordan regelverk og praksis kan tilpasses en digital virkelighet