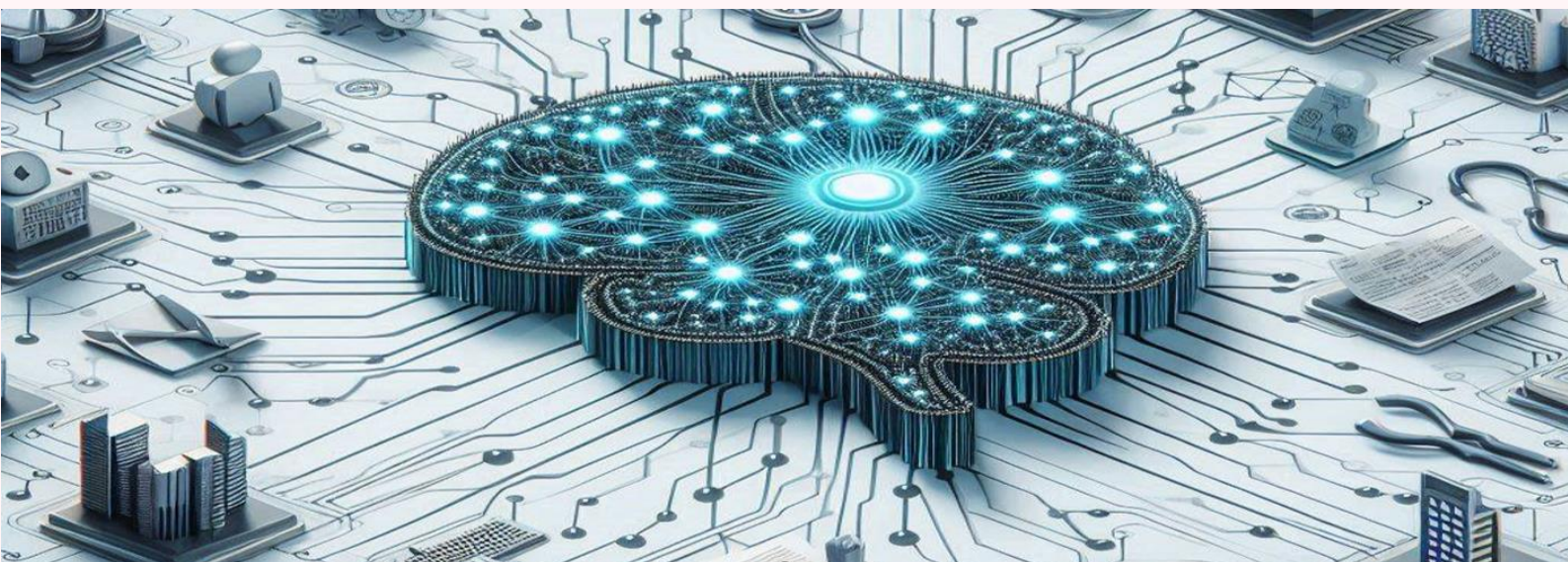


Bruk av kunstig intelligens i staten

Dokument 3:18 (2023–2024)



Til Stortinget

Riksrevisjonen legger med dette fram Dokument 3:18 (2023–2024)
Bruk av kunstig intelligens i staten.

Dokumentet har følgende inndeling:

- Riksrevisjonens konklusjoner, utdyping av konklusjoner, anbefalinger, statsrådets svar og Riksrevisjonens uttalelse til statsrådets svar
- Vedlegg 1: Riksrevisjonens brev til statsråden
- Vedlegg 2: Statsrådets svar
- Vedlegg 3: Forvaltningsrevisjonsrapport med vurderinger

Riksrevisjonen, 2. september 2024

For riksrevisorkollegiet

Karl Eirik Schjøtt-Pedersen
riksrevisor

Innhold

1	Innledning	6
2	Konklusjoner	9
3	Overordnet vurdering	10
4	Utdyping av konklusjoner	10
4.1	Statlige virksomheter utnytter mulighetene med kunstig intelligens ulikt, og kunstig intelligens er fortsatt lite i bruk.....	10
4.2	Viktige forutsetninger for å ta i bruk kunstig intelligens i større skala er fortsatt ikke på plass	14
4.2.1	Stort behov for å avklare juridiske spørsmål i bruken av kunstig intelligens	14
4.2.2	Mangelfull infrastruktur og tilgang på data av god kvalitet.....	17
4.2.3	Stort behov for kompetanse	19
4.2.4	Viktig med språkressurser på norsk.....	19
4.3	De etiske prinsippene for ansvarlig bruk av kunstig intelligens ivaretas i ulik grad. Kontrollmekanismer som sikrer ansvarlig bruk av kunstig intelligens, må være på plass.	20
4.4	Samordningen av arbeidet med kunstig intelligens i offentlig sektor er mangelfull, og den samlede innsatsen er for svak gitt ambisjonen om at Norge skal ha en infrastruktur for kunstig intelligens i verdensklasse	24
5	Anbefalinger	27
6	Statsrådets svar	27
7	Riksrevisjonens uttalelse til statsrådets svar	28
	Vedlegg	29
	Vedlegg 1 Riksrevisjonens brev til statsråden i Digitaliserings- og forvaltningsdepartementet	
	Vedlegg 2 Statsrådets svar	
	Vedlegg 3 Forvaltningsrevisjonsrapport med vurderinger	

Figuroversikt

Figur 1	Antall KI-systemer rapportert, fordelt på fagdepartementer, 2018 – vår 2023.....	12
Figur 2	Opplevde utfordringer med å utvikle og bruke kunstig intelligens. Prosent	14
Figur 3	Bruk av skriftlige retningslinjer eller policyer ved utvikling av KI-systemer	22

Faktaboksoversikt

Faktaboks 1	Eksempler på bruk av KI i staten	13
-------------	--	----

Riksrevisjonen kan gi kritikk etter disse tre alvorlighetsgradene:

1. **Sterkt kritikkverdig** er Riksrevisjonens sterkeste kritikk. Vi bruker dette kritikknivået når vi finner alvorlige svakheter, feil og mangler. Ofte vil disse kunne få svært store konsekvenser for enkeltmennesker eller samfunnet.
2. **Kritikkverdig** bruker vi når vi finner betydelige svakheter, feil og mangler som ofte vil kunne få moderate til store konsekvenser for enkeltmennesker eller samfunnet.
3. **Ikke tilfredsstillende** bruker vi når vi finner svakheter, feil og mangler, men som i mindre grad får direkte konsekvenser for enkeltmennesker eller samfunnet.

1 Innledning

Norge står overfor flere store utfordringer i tiden som kommer, blant annet vesentlig færre yrkesaktive per pensjonist, større behov for arbeidskraft i helse- og omsorgssektoren, oljeinntekter som vil bety mindre enn tidligere, og høyere utgifter til alderspensjoner og helserelaterte ytelser. For at disse utfordringene skal kunne håndteres, vil det være helt avgjørende at offentlig sektor effektiviseres. Kunstig intelligens har potensial til å gjøre staten betydelig mer produktiv, og bidra til å endre statlig forvaltning på helt grunnleggende måter.¹

Potensialet og behovet for å ta i bruk kunstig intelligens i offentlig sektor er tatt opp i en rekke stortingsmeldinger, blant annet Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet* og Meld. St. 30 (2019–2020) *En innovativ offentlig sektor — Kultur, ledelse og kompetanse*, og i de to statlige strategiene – *Nasjonal strategi for kunstig intelligens (2020)* og *Én digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025*.² Stortingsmeldingene og strategidokumentene legger vekt på at offentlig sektor skal effektiviseres gjennom digitalisering. Stortingsmeldingene og strategidokumentene legger vekt på at offentlig sektor skal effektiviseres gjennom digitalisering. I behandlingen av Meld. St. 30 (2019–2020) viste kommunal- og forvaltningskomiteen til at kunstig intelligens er et eksempel på en teknologi i rask utvikling. Komiteen var svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, i form av digitaliseringsvennlig regelverk, gode språkressurser, raske og robuste kommunikasjonsnett og tilstrekkelig regnekraft.

Digitaliserings- og forvaltningsdepartementet har samordningsansvaret for regjeringens IT-politikk, jf. Prop 1 S (2023–2024). Departementet ble opprettet 1. januar 2024, og tok da over ansvaret for digitaliseringsarbeidet i offentlig sektor fra Kommunal- og distriktsdepartementet. Digitalisering og arbeid med kunstig intelligens i offentlig sektor følger ellers sektorprinsippet, der hvert departement er ansvarlig for sin sektor.

Våren 2024 uttalte digitaliserings- og forvaltningsministeren at det er et mål at 80 prosent av offentlig sektor skal ta i bruk kunstig intelligens innen 2025. Dette målet skal bidra til å få en bedre og fornyet offentlig sektor, og alle offentlige virksomheter blir oppfordret til å ta i bruk denne teknologien i enda større grad.³

I Prop. 1 S (2023–2024) for Kommunal- og distriktsdepartementet går det frem at formålet med IT-politikken er å utvikle rammebetingelser som støtter digitaliseringen i de ulike sektorene, på tvers av sektorene og i samfunnet som helhet. Digitaliserings- og forvaltningsdepartementets

¹ Se for eksempel Menon notat: *KI: BETYDNING FOR ARBEIDSSTYRKEN En analyse av potensialet for kunstig intelligens-drevet effektivisering i norsk næringsliv* (november 2023), National Audit Office «Use of Artificial Intelligence in Government» <https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government/> og AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions, Alan Turing Institute, mars 2024

² Kommunal- og moderniseringsdepartementet (2019) *Én digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025* og Nasjonal Strategi for kunstig intelligens (2020).

³ Muntlig spørretime på Stortinget, 8. mai 2024: [Sak nr. 1 \[10:01:35\] - stortinget.no](#). Hentet 13. juni 2024.

samordningsansvar innebærer blant annet å identifisere sektorovergrepene utfordringer og å initiere, koordinere og følge opp tverrgående tiltak. Digitaliseringen av samfunnet skal skje på en bærekraftig og inkluderende måte og bidra til forenkling og effektivisering i både offentlig og privat sektor, men samtidig hindre at den bidrar til en utvikling som sentraliserer oppgaver og arbeidsprosesser, svekker personvern eller forbrukervern, gjør samfunnet mer sårbart for digitale angrep, svekker konkurransen i de digitale markedene eller forsterker digitalt utenforskap.

I Kommunal- og distriktsdepartementets budsjettproposisjoner for 2021 og 2022 vises det til at regjeringen la frem en nasjonal strategi for kunstig intelligens i 2020. I budsjettproposisjonene vises det til at strategien angir retningen for Norges satsing på kunstig intelligens. Strategien inkluderer et sett etiske prinsipper for utvikling og bruk av kunstig intelligens. I strategien vises det videre til et potensial for å oppnå betydelige gevinster gjennom å ta i bruk kunstig intelligens, blant annet bedre og mer persontilpassede tjenester, større effektivitet og bedre planlegging.⁴ Det følger av strategien at offentlige virksomheter aktivt bør utforske potensialet i kunstig intelligens, og at grunnlaget skal være «ansvarlig og pålitelig KI». Målet med undersøkelsen har vært å vurdere om statlige virksomheter og heleide selskaper⁵ utnytter mulighetene som kunstig intelligens gir, og om de utvikler og bruker kunstig intelligens på en ansvarlig måte i samsvar med Stortingets vedtak og forutsetninger.

Undersøkelsen har hatt følgende tre problemstillinger:

1. Bruker statlige virksomheter KI til å effektivisere og forbedre tjenester?
2. Tar statlige virksomheter hensyn til etiske prinsipper i utvikling og bruk av kunstig intelligens?
3. Tilrettelegger Digitaliserings- og forvaltningsdepartementets samordning for ansvarlig bruk av KI i statlige virksomheter?

Begrepet kunstig intelligens kan dekke en rekke ulike systemer i statlig forvaltning. I denne revisjonen brukes den samme definisjonen som er brukt i regjeringens nasjonale strategi for kunstig intelligens (KI-strategien): *Kunstig intelligente systemer utfører handlinger basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål.*⁶ Tolkning av data kan skje på ulike måter, for eksempel fullstendig regelbasert i en programvarerobot eller basert på maskinlæring. Denne definisjonen av kunstig intelligens kan dermed inneholde enkle algoritmer som ikke faller under KI-definisjonen i for eksempel EUs KI-forordning⁷.

Kunstig intelligens er en teknologi i rask utvikling. Siden slutten av 2022 har verktøy basert på generativ kunstig intelligens⁸ blitt lett tilgjengelige, og de brukes i dag av mange. Verktøyene, som for eksempel ChatGPT⁹ og M365

⁴ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*.

⁵ I det følgende er heleide selskaper inkludert i begrepet statlige virksomheter.

⁶ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*.

⁷ European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act), [P9_TA\(2024\)0138](#), fortales punkt (12)

⁸ Generativ KI betegner KI-systemer utviklet for å produsere innhold som tekst, bilder eller kode.

⁹ ChatGPT er en chatbot av OpenAI basert på en stor språkmodell som kan benyttes som virtuell assistent, se <https://openai.com/index/chatgpt/>

Copilot¹⁰, kan blant annet benyttes til å generere tekster, bistå med å lage presentasjoner og med å skrive kode. Denne undersøkelsen omfatter bruk av hyllevare som for eksempel ChatGPT *kun* dersom KI-systemet er videreutviklet eller tilpasset av virksomhetene, eller bygd inn i et eget KI-system.

Uttrykket *ansvarlig bruk av kunstig intelligens* brukes i denne undersøkelsen for utvikling, anskaffelse og bruk av KI-systemer som ivaretar grunnleggende etiske prinsipper gjennom å ta hensyn til personvern, likebehandling, transparens, teknisk sikkerhet og robusthet, og gjennom en styring som sikrer at disse prinsippene er ivaretatt i hele KI-systemets livssyklus.

For å belyse problemstillingene har vi gjennomført en kartleggingsundersøkelse av omtrent 200 statlige virksomheter. Vi har også gjennomført en spørreundersøkelse med over 100 respondenter i 66 statlige virksomheter som ga detaljert informasjon om i underkant av 200 KI-systemer. Videre har vi gjennomført caseundersøkelser av fire utvalgte KI-prosjekter henholdsvis i skatteetaten, i Arbeidstilsynet, i Statens Pensjonskasse og på St. Olavs Hospital. Casene ble valgt ut på bakgrunn av blant annet modellenes kompleksitet og bruk av personopplysninger. Videre har vi analysert dokumenter som blant annet tildelingsbrev, årsrapporter og forskningsrapporter. Vi har i tillegg intervjuet blant annet flere departementer og statlige etater.

Undersøkelsen omfatter perioden 2018–2023.

Undersøkelsen har blant annet tatt utgangspunkt i følgende vedtak og forutsetninger fra Stortinget:

- Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet*, jf. Innst. 84 S (2016–2017)
- Meld. St. 30 (2019–2020) *En innovativ offentlig sektor – Kultur, ledelse og kompetanse*, jf. Innst. 191 S (2020–2021)
- Meld. St. 22 (2020–2021) *Data som ressurs – Datadrevet økonomi og innovasjon*, jf. Innst. 568 S (2021–2022)
- Kommunal- og distriktsdepartementets budsjettproposisjoner for perioden 2021–2023 (jf. Prop. 1 S 2020–2021 og 2022–2023), som viser til den nasjonale strategien for kunstig intelligens fra 2020, og at strategien angir retningen for Norges satsing på kunstig intelligens. / Innst. 16 S (2020–2021) Innstilling fra kommunal- og forvaltningskomiteen om bevilgninger på statsbudsjettet for 2021.

Rapporten ble forelagt Digitaliserings- og forvaltningsdepartementet, Arbeids- og inkluderingsdepartementet, Finansdepartementet, Helse- og omsorgsdepartementet, Justis- og beredskapsdepartementet og Kultur- og likestillingsdepartementet ved brev 24. april 2024. Departementene har i brev gitt kommentarer til rapporten.¹¹ Kommentarene er i hovedsak innarbeidet i rapporten og i dette dokumentet.

¹⁰ M365 Copilot er et KI-system fra Microsoft basert på en stor språkmodell og bygd inn i Microsoft 365, som kan brukes som virtuell assistent med tilgang til interne data.

¹¹ Brev fra Helse- og omsorgsdepartementet av 22. mai, brev fra Digitaliserings- og forvaltningsdepartementet 27. mai og brev fra de øvrige departementene 24. mai 2024.

2 Konklusjoner



- Statlige virksomheter utnytter mulighetene med kunstig intelligens ulikt, og kunstig intelligens er fortsatt lite i bruk.
- Viktige forutsetninger for å ta i bruk kunstig intelligens i større skala er fortsatt ikke på plass.
 - stort behov for å avklare juridiske spørsmål i bruken av kunstig intelligens
 - mangelfull infrastruktur og tilgang på data av god kvalitet
 - stort behov for kompetanse
 - viktig med språkressurser på norsk
- De etiske prinsippene for ansvarlig bruk av kunstig intelligens ivaretas i ulik grad. Kontrollmekanismer som sikrer ansvarlig bruk av kunstig intelligens, må være på plass.
- Samordningen av arbeidet med kunstig intelligens i offentlig sektor er mangelfull, og den samlede innsatsen er for svak gitt ambisjonen om at Norge skal ha en infrastruktur for kunstig intelligens i verdensklasse.

3 Overordnet vurdering

Ikke tilfredsstillende



Det er ikke tilfredsstillende at Digitaliserings- og forvaltningsdepartementet gjennom styring og i samarbeid med de andre departementene ikke har lagt bedre til rette for at offentlig sektor kan utnytte mulighetene med kunstig intelligens og ta i bruk kunstig intelligens på en ansvarlig måte. Den samlede innsatsen er for svak gitt ambisjonen om at Norge skal ha en infrastruktur for kunstig intelligens i verdensklasse.

4 Utdyping av konklusjoner

4.1 Statlige virksomheter utnytter mulighetene med kunstig intelligens ulikt, og kunstig intelligens er fortsatt lite i bruk

Kunstig intelligens er et viktig verktøy for utvikling av en bærekraftig, effektiv og brukerrettet offentlig forvaltning, jf. blant annet Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet*, Meld. St. 30 (2019–2020) *En innovativ offentlig sektor – Kultur, ledelse og kompetanse* og nasjonal strategi for kunstig intelligens.¹² Stortinget viser samtidig til betydningen av at utviklingen og bruk av kunstig intelligens må skje innenfor ansvarlige rammer der blant annet personvernet ivaretas.¹³

Potensialet for effektivisering av offentlig sektor med bruk av kunstig intelligens er stort. I Norge anslår for eksempel det forskningsbaserte analyse- og rådgivingselskapet Menon at den potensielle årlige verdiskapningen med kunstig intelligens i privat og offentlig sektor samlet sett er ca. 500–600 milliarder kroner. Det er særlig stort effektiviseringspotensial innen saksbehandling i offentlig sektor. Ved full utnyttelse av dagens teknologi for kunstig intelligens, kan offentlig sektor effektivisere arbeidsoppgaver som årlig vil tilsvare 155 000 årsverk, ifølge rapporten.¹⁴ Internasjonalt er det også flere studier som peker i samme retning. Ifølge en rapport fra Alan Turing Institute kan over 80 prosent av

¹² Kommunal- og moderniseringsdepartementet (2019) *En digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025*

¹³ Innst. 191 S (2020–2021).

¹⁴ Menon Economics (2023), *KI: Betydning for arbeidsstyrken. En analyse av potensialet for kunstig intelligens-drevet effektivisering i norsk næringsliv*

komplekse repetitive handlinger i britisk statsforvaltning automatiseres ved hjelp av kunstig intelligens.¹⁵ McKinsey & Company anslår at det er et betydelig potensial for produktivitetsøkning med generativ kunstig intelligens i offentlig sektor, inklusive helsetjenester og utdanning. Konsulentfirmaet anslår at 40 prosent av arbeidstimene på tvers av bransjer kan påvirkes av KI-systemer som oppsummerer og analyserer tekst, og at kunderettete tjenester har et automatiseringspotensial på 60 prosent innen fem til ti år gjennom teknologi som for eksempel chatboter.¹⁶

Imidlertid er kunstig intelligens i rask utvikling, og studier av produktivetsgevinster er oftest basert på et fåtall tidlige brukere. For at teknologien skal ha langvarig effekt på samfunnsnivå, må KI tas i bruk i stor utstrekning over tid og integreres på en god måte i organisatoriske prosesser.¹⁷

Undersøkelsen viser at det er stor variasjon i hvilken grad statlige virksomheter har tatt i bruk kunstig intelligens, og at kunstig intelligens i dag fortsatt er lite i bruk. Blant de rundt 200 undersøkte statlige virksomhetene¹⁸ er det under 50 prosent som har noe erfaring med utvikling og/eller bruk av kunstig intelligens. I tillegg har i underkant av 20 prosent av statlige virksomheter uten erfaring med KI planer om å ta i bruk kunstig intelligens i fremtiden. I denne undersøkelsen har Riksrevisjonen lagt til grunn en relativt bred definisjon av kunstig intelligens, tilsvarende som den nasjonale strategien for kunstig intelligens. Mange prosjekter og systemer kan derfor rubiseres som kunstig intelligens uten at det er noe annet enn en enkel programvarerobot. Over halvparten av KI-systemene i undersøkelsen bruker for eksempel kun enkle modeller for å tolke data. Med en strengere definisjon av hva som kan betegnes som et KI-system – med krav til for eksempel mer avanserte modeller – ville antallet oppgitte KI-systemer i staten ha vært vesentlig lavere. Som omtalt innledningsvis inngår imidlertid ikke enklere bruk av hyllevarer, som for eksempel ChatGPT i det som er kartlagt i undersøkelsen.

Blant virksomhetene som har erfaring med kunstig intelligens, viser undersøkelsen at et flertall av disse (56 prosent) har erfaring med kun ett til tre KI-prosjekter. I overkant av 10 prosent av virksomhetene har erfaring med ti eller flere KI-prosjekter.

Uavhengig av hvor avanserte KI-systemene er, viser undersøkelsen at virksomhetenes mål med å utvikle og ta i bruk KI-systemer i stor grad er å effektivisere og forbedre egen oppgaveløsning, inkludert å frigjøre tid. Undersøkelsen viser videre at de fleste virksomhetene mener de i stor grad lykkes med å nå målene de har satt seg med bruken av kunstig intelligens.

¹⁵ Alan Turing Institute. (2024). *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions.*

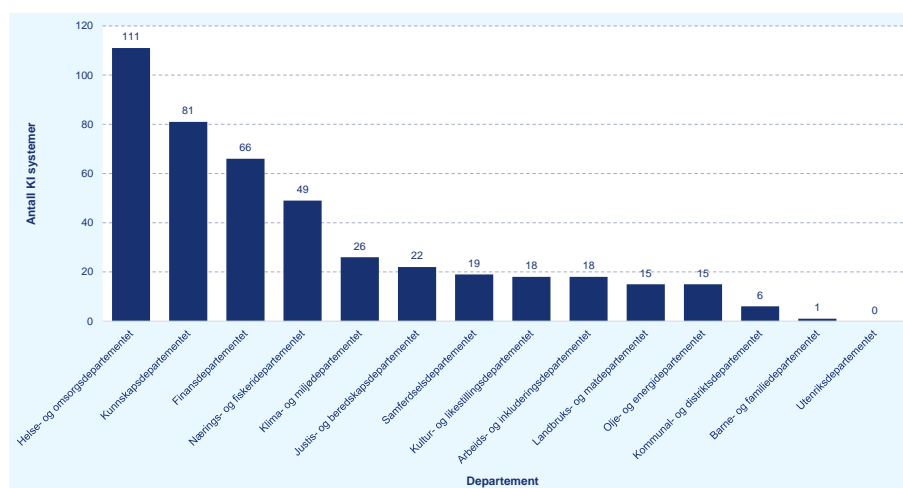
¹⁶ McKinsey & Company (2023), *Unlocking the potential of generative AI: Three key questions for government agencies*, [nettartikkel](#)

¹⁷ OECD (2024), *The impact of artificial intelligence on productivity, distribution and growth. Key mechanisms, initial evidence and policy challenges. OECD artificial intelligence papers No. 15*

¹⁸ Dette utgjør nært alle statlige virksomheter, inklusive heleide selskaper. Unntaket er Stortinget og Sametinget, forsvarssektoren, alle fagdepartementene, statsforvalterne og statlige fond

Figur 1 viser antall KI-systemer rapportert etter departement. Helseforetakene har klart flest KI-prosjekter. Det er også helseforetakene som i størst grad bruker mer avanserte modeller. Etter Helse- og omsorgsdepartementet følger kunnskapssektoren og noen av virksomhetene under Finansdepartementet med flest KI-prosjekter. Under Nærings- og fiskeridepartementet er det også flere KI-prosjekter blant både underliggende virksomheter og heleide selskaper. Nærmere 70 prosent av samtlige innrapporterte KI-prosjekter ligger under disse fire departementene.

Figur 1 Antall KI-systemer rapportert, fordelt på fagdepartementer, 2018 – vår 2023¹⁹



Kilde: Svar på brev til statlige virksomheter og heleide statlige selskaper (n = 130)

Noen av ulikhetene mellom sektorene kan etter Riksrevisjonens vurdering blant annet tilskrives at sektorene har ulike behov og grunnleggende forutsetninger for å utvikle og bruke kunstig intelligens. Overordnet styring og graden av tilrettelegging kan også være medvirkende faktorer.

I helsesektoren er det eksplisitt uttalt at kunstig intelligens kan bidra til mer effektiv ressursbruk og til et bærekraftig helsevesen, jf. for eksempel Meld. St. 7 (2019–2020) *Nasjonal helse- og sykehusplan 2020–2023*. I oppdragsbrevet for 2024 til de regionale helseforetakene har Helse- og omsorgsdepartementet videre gitt føringer om å ta et aktivt grep om å utvikle og bruke kunstig intelligens i egen sektor. Dette gjelder også tildelingsbrevene til Helsedirektoratet for perioden 2020–2024, der det nasjonale koordineringsprosjektet «Bedre bruk av kunstig intelligens» er en del av arbeidet med å følge opp *Nasjonal helse- og sykehusplan 2020–2023*. Helsektoren har i utgangspunktet i tillegg noen grunnleggende forutsetninger for å kunne utvikle og å ta i bruk kunstig intelligens. Selv om det kan være utfordringer med manglende digitalisering og kobling av ulike datakilder i helsesektoren, finnes det store mengder data, blant annet gjennom helseregistrene, som kan utnyttes til å trene og utarbeide modeller. I tillegg finnes det større etablerte forskningsmiljøer i helseforetakene.

¹⁹ Dokument 3:8 (2023–2024) Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen, vedlegg om «Kartlegging av utvikling og bruk av kunstig intelligens i statsforvaltningen og statseide selskaper».

Blant de andre undersøkte sektorene kan det i noen tilfeller ha vært gitt føringer til underliggende virksomheter om kunstig intelligens, men gjennomgående er det mest opp til den enkelte virksomhet å eventuelt ta i bruk kunstig intelligens. De intervjuede departementene viser i den sammenhengen til prinsippene om mål- og resultatstyring, og følgelig at virksomhetene selv må velge og bruke de virkemidlene og den teknologien som gir de beste løsningene for å sikre en effektiv oppgaveløsning.

Undersøkelsen viser videre at de undersøkte departementene kan ha ulik tilnærming til å utarbeide hjemler som tydeliggjør om underliggende virksomheter har et klart juridisk grunnlag for å behandle personopplysninger. Dette omtales i detalj under kapittel 4.3.

Samlet sett er omfanget av utvikling og bruk av kunstig intelligens i staten så langt lite. Enkelte sektorer og virksomheter begynte imidlertid å utforske mulighetene for mer enn ti år siden, og noen har satt mer avanserte systemer i produksjon. Blant helseforetakene og for eksempel i Skatteetaten og Statens pensjonskasse finnes det eksempler på utvikling og bruk av kunstig intelligens som andre virksomheter kan lære av.

Faktaboks 1 Eksempler på bruk av KI i staten

To eksempler på KI-systemer brukt i offentlige virksomheter

- Skatteetatens *fradragmodell* for risikobasert kontrollutplukk predikerer sannsynlighet for feil i oppgitte skattefradrag. Systemet brukes til beslutningsstøtte ved utplukk av skattemeldinger for å kontrollere fradrag. KI-systemet ble utviklet av skatteetaten selv og har vært i produksjon siden 2014. Systemet har bidratt til betydelig effektivisering av skatteetatens kontrollarbeid og til mer lønnsomme kontroller.
- St. Olavs Hospital HF bruker et *fluorescensmikroskop* med et innebygd KI-system for klassifisering av celler som del av en analyse som brukes i diagnostisering av kreft. KI-systemet er utviklet av en leverandør, og St. Olavs hospital HF har i samarbeid med leverandøren tilpasset det til pasientgrunnlaget. KI-systemets resultater blir verifisert av fagpersoner. Systemet gjør at sykehuset sparer tid på diagnostisering, og det har bidratt til høyere kvalitet på analysene og bedre kvalitetssikring.

Kilde: Riksrevisjonens caseundersøkelser

Undersøkelsen viser på den annen side at det finnes store statlige etater som har arbeidet med å utvikle egne modeller for kunstig intelligens over flere år, men ennå ikke har noen produksjonssatte løsninger basert på egenutviklet kunstig intelligens. Alternativet er å kjøpe ferdige KI-systemer som hylleware. Utover de generative modellene kan det imidlertid være vanskelig å finne ferdige modeller som passer til virksomhetens egenart. Ved et eventuelt innkjøp av avanserte modeller vil det også kunne kreve omfattende testing, tilpasninger og kvalitetssikring for å sikre at modellen fungerer som intendert. I tillegg vil det ofte være et fåtall tilbydere av spesialiserte systemer.

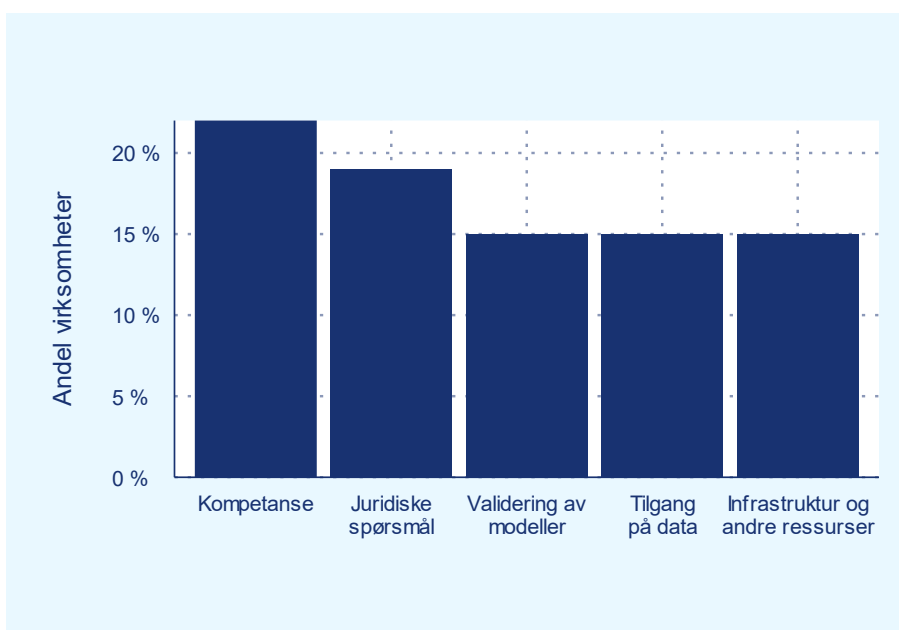
Ettersom det totalt sett er få eksempler på at det har blitt utviklet og tatt i bruk KI-systemer, er det få eksempler hvor kunstig intelligens i stor grad har bidratt til effektivisering og en mer produktiv offentlig sektor.

Etter Riksrevisjonens vurdering er det et stort, uutnyttet potensial i offentlig sektor til å effektivisere gjennom digitalisering og bruk av kunstig intelligens.

4.2 Viktige forutsetninger for å ta i bruk kunstig intelligens i større skala er fortsatt ikke på plass

Undersøkelsen viser at det er flere barrierer for utvikling og bruk av kunstig intelligens i staten. Noen av de mest sentrale barrierene er vist i Figur 2. I tillegg trekker flere statlige virksomheter frem andre faktorer som også kan hindre en effektiv utvikling av kunstig intelligens, herunder utvikling av språkmodeller.

Figur 2 Opplevde utfordringer med å utvikle og bruke kunstig intelligens. Prosent



Kilde: Svar på brev til statlige virksomheter og heleide statlige selskaper (n = 130)

4.2.1 Stort behov for å avklare juridiske spørsmål i bruken av kunstig intelligens

I Norge har vi generelt et teknologinøytralt regelverk. Offentlig sektor står dermed fritt til å velge hvilken teknologi som passer best til sin oppgaveløsning, og regelverket skal ikke være til hinder for bruk av teknologi. I Norge finnes det heller ingen spesifikk regulering av kunstig intelligens. Men en rekke lover er relevante i utviklingen og bruken av kunstig intelligens. Det gjelder for eksempel personopplysningsloven og likestillings- og diskrimineringsloven. I tillegg er KI-forordningen, som ble vedtatt i EU parlamentet i mars 2024, vurdert som EØS-relevant, og Digitaliserings- og forvaltningsdepartementet har begynt forberedelsen for innlemming i norsk lov.

Undersøkelsen viser at både mangel på tydelige lovhjemler i sektorlovgivning og usikkerhet rundt tolkning av lover og annet regelverk generelt, er faktorer som ser ut til å hindre at statlige virksomheter i større grad utvikler og tar i bruk kunstig intelligens. For å kartlegge blant annet utfordringer med å utvikle og bruke kunstig intelligens sendte Riksrevisjonen en spørreundersøkelse til statlige virksomheter som har erfaring med å utvikle og/eller bruke kunstig intelligens. Spørreundersøkelsen ble sendt til 143 kontaktpersoner og 78 prosent svarte. Rundt 40 prosent av respondentene svarte at det er vanskelig å ivareta personvern hensyn i arbeidet med å utvikle systemer for kunstig intelligens. Det er mye usikkerhet om hva som er et adekvat juridisk grunnlag til å behandle personopplysninger i både utvikling og bruk av KI-systemer.

Noen av de departementene som inngår i undersøkelsen, som Finansdepartementet og Helse- og omsorgsdepartementet, har gjennom sektorlovverket arbeidet med å gi underliggende virksomheter et tydeligere juridisk grunnlag til generelt å kunne behandle personopplysninger. Som omtalt over viser undersøkelsen at det også er innenfor disse sektorene det er flest KI-prosjekter, men det har ikke nødvendigvis en direkte sammenheng. Imidlertid har eksempelvis Nav og Statens pensjonskasse, som begge er underlagt Arbeids- og inkluderingsdepartementet, ulike hjemmelsgrunnlag til å behandle personopplysninger i sine respektive lover. Statens pensjonskasse har et tydelig rettslig grunnlag til å behandle personopplysninger. For Navs del er dette imidlertid fortsatt uavklart selv om Datatilsynet pekte på dette problemet overfor Nav i januar 2022.

Undersøkelsen viser at uavhengig av om virksomhetene har et generelt hjemmelsgrunnlag til å behandle personopplysninger, synes virksomhetene ofte det er vanskelig å håndtere spørsmål om personvern i det enkelte KI-prosjekt. Det er særlig fordi bruken av personopplysninger må vurderes spesifikt i det enkelte prosjektet, og fordi personvern må være ivaretatt i alle deler av et KI-prosjekt, også i utviklingsfasen. Noen av de sentrale spørsmål som må håndteres, er om bruken av personopplysninger er forenlig med grunnlaget for innhenting av opplysningene, og om behandlingen er nødvendig og proporsjonal. Her er det mye usikkerhet blant statlige virksomheter. Det er positivt at noen virksomheter velger å håndtere slik usikkerhet ved å prøve ut mindre prosjekter for å bygge erfaring. Etter Riksrevisjonens vurdering er det likevel stor risiko for at den vedvarende usikkerheten rundt flere juridiske spørsmål vil forsinke og hindre statlige virksomheter i å ta i bruk de mulighetene som kunstig intelligens gir.

Departementene og virksomhetene har selv ansvaret for å vurdere eget hjemmelsgrunnlag og behov for eventuelle endringer i eget regelverk for å tydeliggjøre grunnlaget for å behandle personopplysninger. Samtidig er det en rekke juridiske problemstillinger som er tilnærmedesvis like for mange virksomheter. Virksomhetene kan for eksempel i stor grad ha like spørsmål rundt hjemmelsgrunnlag og tolkningen av regelverket, og da særlig spørsmål om behandlingen av personopplysninger. For å bistå virksomhetene i slike spørsmål har Digitaliserings- og forvaltningsdepartementet arbeidet med veiledning, særlig gjennom Datatilsynet, men også gjennom Digitaliseringsdirektoratet. Digitaliseringsdirektoratet opplyser i intervju at direktoratet ser behov for en mer samordnet tolkning av regelverk for

utvikling og bruk av kunstig intelligens. Etter Riksrevisjonens vurdering tyder de utfordringene som et stort antall virksomheter møter i arbeidet med personopplysninger, på at det er et særlig stort behov for veiledning på dette området.

Opprettelsen av den regulatoriske sandkassen hos Datatilsynet i 2020 var et sentralt tiltak for å bistå blant annet offentlige virksomheter i å øke forståelsen av de regulatoriske kravene som stilles på personvernområdet. Bedre kunnskap vil kunne korte ned tiden fra utvikling og testing til faktisk utrulling av KI-løsninger. Løsningene som settes i drift etter å ha deltatt i sandkassen skulle fungere som foregangseksempler. På overordnet nivå synes sandkassen å være et godt tiltak. Det gir læring både for virksomhetene som deltar, og for Datatilsynet. Undersøkelsen viser imidlertid at siden sandkassen ble opprettet i november 2020, har kun fem prosjekter i statlig regi deltatt. Det har vært relativt få søkere til ordningen, og Datatilsynet har heller ikke kapasitet til å håndtere flere prosjekter enn de gjør i dag. I intervju med to av virksomhetene som har deltatt i sandkassen ser det videre ut til at erfaringene med å delta spriker. En av disse opplevde prosessen som mindre positiv fordi det ble ikke gitt noe svar på det prosjektet ønsket å avklare. Det andre prosjektet opplevde deltakelsen som positiv. Samtidig oppsto det større, prinsipielle spørsmål som ikke kunne svares på innenfor prosjektets rammer. Selv om deltakelse kan gi læring for enkeltprosjektene, synes det også å være en utfordring for sandkassen å formidle viktig læring videre og sikre at den blir brukt av andre KI-prosjekter til videre utvikling og avklaring av juridiske spørsmål. Etter Riksrevisjonens vurdering er det også vanskelig å se at prosjektene som har deltatt i sandkassen har fungert som foregangseksempler for andre prosjekter og kortet ned tiden fra uttesting til at KI-prosjekter blir satt i full drift.

I behandlingen av Meld. St. 30 (2019–2020) viste kommunal- og forvaltningskomiteen til at den var svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, herunder et digitaliseringsvennlig regelverk. I den nasjonale strategien for kunstig intelligens vises det også til at regjeringen skal gjennomgå og vurdere regelverk som er til hinder for hensiktsmessig og ønsket bruk av kunstig intelligens i offentlig og privat sektor. Undersøkelsen viser at en slik gjennomgang ikke har blitt gjort. Digitaliserings- og forvaltningsdepartementet legger til grunn at det er departementene og virksomhetene som må vurdere behovet for dette. Digitaliserings- og forvaltningsdepartementet mener at strategien ikke legger opp til en bred gjennomgang og vurdering av regelverket. Som forberedelse til KI-forordningen satte regjeringen sommeren 2023 ned en hurtigarbeidende arbeidsgruppe under ledelse av Justis- og beredskapsdepartementet. Arbeidsgruppen skulle utarbeide en plan for å innlemme EUs forordning om kunstig intelligens i norsk lov og vurdere nasjonale behov for regulering (sivil sektor) utover EUs forordning. Denne juridiske arbeidsgruppen er videreført for at den skal arbeide med andre aktuelle tverrgående problemstillinger. Tilpasninger til EUs forordning vil alene ikke være noen løsning på den påviselig store usikkerheten på tvers av sektorer og virksomheter rundt bruken av dagens samlede regelverk. Digitaliserings- og forvaltningsdepartementet har etter Riksrevisjonens vurdering ikke i tilstrekkelig grad fylt sin overordnede rolle ved å bidra til å utvikle



Den regulatoriske sandkassen

Datatilsynets regulatoriske sandkasse skal stimulere til personvernvennlig innovasjon og digitalisering. Sandkassen skal hjelpe enkeltaktører med å følge regelverket og utvikle løsninger med godt personvern. Sandkassen skal tilby gratis veiledning til utvalgte private og offentlige virksomheter, av forskjellige typer og størrelser, på tvers av ulike sektorer.

Kilde: Datatilsynet

rammebetingelser som støtter digitaliseringen i og på tvers av de ulike sektorene.

4.2.2 Mangelfull infrastruktur og tilgang på data av god kvalitet

Tilgang på store datamengder av god kvalitet er ofte en viktig forutsetning for å kunne utvikle effektive KI-systemer. Det kan være data både fra egen virksomhet og fra andre virksomheter. Dataene må være av en kvalitet og foreligge i et omfang som gjør at de kan brukes til å trene opp maskiner før et system kan settes inn i produksjon. Tilgang på data er også nødvendig for å validere innkjøpte KI-modeller og til å kalibrere og kontinuerlig overvåke eksisterende modeller. Potensialet for å effektivisere gjennom kunstig intelligens ligger nettopp i mulighetene til å utnytte både eksisterende og nye data.

Over flere år har prinsippet om «orden i eget hus» fremhevet at data av god kvalitet er særlig viktig i arbeidet med digitalisering i offentlig sektor, jf. blant annet Meld. St. 27 (2015–2016) *Digital Agenda for Norge*. Gjennom Digitaliseringsrundskrivet er statlige virksomheter også pålagt å sikre god datakvalitet, og dette er i tillegg løftet frem i den nasjonale strategien for kunstig intelligens.

Undersøkelsen viser at mangelfull tilgang på data av god kvalitet kan være til hinder for å utvikle og bruke kunstig intelligens. Rundt halvparten av respondentene i Riksrevisjonens spørreundersøkelse svarte at dårlig tilgang på data og lav datakvalitet har vært en utfordring i arbeidet med kunstig intelligens. At tilgang på data av god kvalitet er begrenset, kan skyldes flere forhold. Noen virksomheter mangler gode, interne data. Andre virksomheter har mye data, men de er ikke nødvendigvis digitale eller tilgjengelige i et felles system der dataene kan utnyttes til å utvikle KI-modeller. Hvorvidt tilgang på data er et problem, kan også avhenge av hvilket ambisjonsnivå en virksomhet har. Kompliserte modeller som skal bruke flere kilder som bilder og tekst, i tillegg til strukturerte registerdata, vil kreve et langt større datagrunnlag enn eksempelvis en enkel regnskapsrobot.

Videre har flere virksomheter ikke god nok infrastruktur til verken å lagre eller å håndtere store datamengder til analyse. Virksomhetene peker blant annet på betydningen av moderne datavarehus og analyseplattformer for å utvikle og bruke kunstig intelligens. Helseforetak trekker videre frem at manglende infrastruktur mellom systemer på sykehusene er en utfordring i arbeidet med kunstig intelligens. Undersøkelsen viser at mangelfull digital infrastruktur fører til forsinkelser i arbeidet med å utvikle KI-modeller, forringer kvaliteten på løsningene og gir høyere kostnader.

Særlig for små virksomheter kan det være kostbart å investere i ulike digitale løsninger som kreves for utvikling og bruk av kunstig intelligens. Flere statlige virksomheter etterlyser i den forbindelse felles digitale løsninger og muligheter for å få tilgang på skyløsninger. Riksrevisjonen merker seg at Skate peker på at det er store synergier å hente gjennom mer samhandling på tvers med fellesløsninger og mulighet for å spare ressurser.²⁰ Videre har Nasjonal sikkerhetsmyndighets konseptvalgutredning for nasjonal skytjeneste pekt på at det for noen datatyper og IT-systemer er sårbart å være avhengig av utenlandske skytjenester, og anbefaler overfor Justis- og beredskapsdepartementet etablering av en nasjonal skytjeneste for Norge.²¹ Riksrevisjonen merker seg at det imidlertid er uenighet om hva som er den beste fremgangsmåten for etablering av en nasjonal skyløsning, blant annet fordi flere statlige virksomheter allerede har investert i kommersielle skytjenester.

Digitaliserings- og forvaltningsdepartementet viser i intervju til Sigma2 som en felles tungregningsinfrastruktur for forskningsmiljøet. Departementet peker på at ambisjonen er at denne løsningen i større grad skal kunne brukes også av andre sektorer enn forsknings- og utdanningssektoren.

Etter Riksrevisjonens vurdering ser det ut til at mangel på fellesløsninger, som muligheter til å lagre og analysere store datamengder, kan svekke effektiviteten i arbeidet med utvikling av kunstig intelligens. Større tilgang på nødvendig infrastruktur kunne også ha styrket den teknologiske modenheten i hele offentlig sektor.

Manglende datadeling på tvers av virksomheter kan være en annen årsak til at tilgangen på data er mangelfull. Skal datadelingen fungere, må hver enkel offentlige virksomhet legge til rette egne data slik at andre kan bruke dataene på en hensiktsmessig måte. Digitaliserings- og forvaltningsdepartementet viser til at den enkelte virksomhet selv må ta ansvar for at egne data er av god kvalitet. Etableringen av Felles datakatalog og andre tiltak i regi av Digitaliseringsdirektoratet har bidratt positivt til mer deling av åpne data. Det er likevel anslagsvis bare kun en fjerdedel av de virksomhetene som er omfattet av digitaliseringsrundskrivet som oppfyller kravene om deling og videre bruk av informasjon på Felles datakatalog. Her vil vi også vise til utfordringer med deling av data i offentlig sektor og manglende registrering i Felles datakatalog, som ble påvist i Riksrevisjonens Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen*. I tillegg skal Felles datakatalog kun inneholde åpne data, og legge til rette for deling av lukkede datasett gjennom metadata om disse. Faktisk tilgang på lukkede data for eksempel med personopplysninger krever andre løsninger. Her vil Riksrevisjonen igjen vise til Dokument 3:8 (2023–2024). I rapporten pekes det blant annet på at offentlig sektor har mye data som ikke deles og gjenbrukes, og at mange virksomheter ikke har god nok oversikt over egne data. Rapporten viste også til at manglende og



Skate

Topplederorganet Skate - Styling og koordinering av tjenester i e-forvaltningen - er et strategisk samarbeidsråd og rådgivende organ til Digitaliseringsdirektoratet og digitaliserings- og forvaltningsministeren. Skate skal bidra til en samordnet digitalisering av offentlig sektor som gir gevinster for innbyggerne, næringslivet, frivillig sektor, og offentlige virksomheter.

Kilde:
Digitaliseringsdirektoratet

²⁰ Skates erfaringsrapport for 2023.

²¹ NSM (2023) *Konseptvalgutredning for nasjonal skytjeneste*. <https://nsm.no/regelverk-og-hjelp/rapporter/konseptvalgutredning-for-nasjonal-skytjeneste>

sene avklaringer av juridiske problemstillinger hindrer deling og gjenbruk av data mellom offentlige virksomheter.

Etter Riksrevisjonens vurdering skaper begrenset tilgang på data av god kvalitet, ofte koblet med mangelfull infrastruktur og svake systemer for å håndtere data, utfordringer i arbeidet med kunstig intelligens. Sektorene har et selvstendig ansvar for å sikre god datakvalitet og at åpne data er av en kvalitet som enkelt kan deles med andre sektorer. Gitt utfordringene med begrenset tilgang på data av god kvalitet synes det etter Riksrevisjonens vurdering ikke å være tilstrekkelig oppfølging av at kravene til deling av data og særlig til orden i eget hus blir etterlevd i virksomhetene. Dette får etter Riksrevisjonens vurdering konsekvenser for offentlig sektors muligheter til å utvikle og utnytte potensialet i kunstig intelligens.

4.2.3 Stort behov for kompetanse

Undersøkelsen viser at en av de største utfordringene med å utvikle og bruke kunstig intelligens er mangel på kompetanse. Det gjelder ofte IT-relatert kompetanse, men også juridisk kompetanse. Videre er det behov for å heve kompetansen på risikoen for diskriminering med kunstig intelligens, og de ulike kompetansemiljøene internt i en virksomhet må samhandle bedre. I noen store virksomheter som skatteetaten, Nav og helseforetakene, og også i mellomstore virksomheter som Statens pensjonskasse, har det blitt satset på å etablere større miljøer med kompetanse til å arbeide med kunstig intelligens. Flere virksomheter har imidlertid ikke kapasitet eller ressurser til å ansette flere med nødvendig digital kompetanse. Mange virksomheter er avhengig av å leie inn konsulenter for å kunne utvikle og bruke kunstig intelligens. Da vil det være viktig med bestillerkompetanse og at det legges opp til nødvendig overføring av kompetanse.

Når det gjelder utfordringene med fagkompetanse, merker Riksrevisjonen seg at det i Meld. St. 14 (2022–2023) *Utsyn over kompetansebehovet i Norge* vises til at myndighetene og universitetene og høyskolene har et delt ansvar for å dimensjonere utdanningstilbudet – dels ut fra særlige arbeidskraftsbehov. Behovet for IT-kompetanse er beregnet til å øke i årene frem til 2030.

Fagkompetanse, på tvers av fagområder som IT og jus, er imidlertid bare en del av utfordringen med manglende kompetanse. Et annet aspekt er manglende kompetanse blant ledere og beslutningstakere: For at KI-prosjekter skal lykkes og realisere gevinster, er det ofte nødvendig med endringer i virksomhetens prosesser og god endringsledelse.

Undersøkelsen viser at manglende forankring internt, urealistiske forventninger fra brukere eller ledelsen, samt motstand i organisasjonen for å endre eksisterende prosesser eller arbeidsstrukturer er utfordringer i arbeidet med å utvikle kunstig intelligens.

4.2.4 Viktig med språkressurser på norsk

For at norske innbyggere skal få mer avanserte KI-tjenester på sitt eget språk, er det avgjørende at det finnes gode språkressurser på bokmål, nynorsk og samisk. Offentlige organer har etter språkloven plikt til å bruke, utvikle og styrke bokmål, nynorsk og samiske språk, og dette gjelder også

for digital kommunikasjon ved hjelp av nye KI-løsninger.²² Med bruk av kunstig intelligens kan man for eksempel effektivisere arbeidet med å oversette og skrive tekster og med å analysere innholdet i store dokumentmengder. Om man skal utvikle slike løsninger for norske og samiske skriftspråk og dialekter, må imidlertid teknologien tilpasses disse språkene og de lokale forholdene. Språkteknologi, som verktøy for talegjenkjenning og språkforståelse, er derfor en viktig komponent i kunstig intelligens. Utvikling av KI blir imidlertid ledet an av store internasjonale teknologiselskaper, som ikke nødvendigvis tar hensyn til norsk språk og norske samfunnsforhold.²³

En sentral forutsetning for å kunne trene språkmodellene er tilgang på store mengder språkressurser. Ifølge den nasjonale KI-strategien vil regjeringen legge til rette for at språkressurser samles inn og gjøres tilgjengelig.

Bruk av generativ kunstig intelligens basert på språkmodeller har økt betydelig siden lansering av åpent tilgjengelige ChatGPT i november 2022²⁴. Som vist til tidligere er det mange mulige bruksområder i offentlig sektor. Grunnmodellene til disse KI-verktøyene er imidlertid nesten utelukkende basert på og utviklet av et fåtall globale, private selskaper uten tilknytning til norsk språk eller norske samfunnsforhold og verdier.

Riksrevisjonen merker seg at Kultur- og likestillingsdepartementet peker på at det bør arbeides videre med å sikre språkdata til blant annet bruk i kunstig intelligens, og at flere forhold i denne sammenheng taler for å legge til rette for en felles norsk infrastruktur for kunstig intelligens. Både opphavsrett til åndsverk og tilgang til regnekraft er utfordringer i arbeidet med å sikre språkressurser.

4.3 De etiske prinsippene for ansvarlig bruk av kunstig intelligens ivaretas i ulik grad.

Kontrollmekanismer som sikrer ansvarlig bruk av kunstig intelligens, må være på plass.

Kommunal- og forvaltningskomiteen har pekt på at det er særlig viktig at myndighetene har nødvendig kontroll på og regulerer bruken av kunstig intelligens slik at personvern ivaretas.²⁵ Grunnlaget for dette er en ansvarlig og pålitelig utvikling og bruk av kunstig intelligens. Regjeringens nasjonale strategi for kunstig intelligens legger til grunn at kunstig intelligens i Norge skal bygge på de etiske prinsippene utarbeidet av EU-kommisjonens ekspertgruppe for kunstig intelligens.²⁶ Ansvarlig bruk av kunstig intelligens er dermed forankret i krav til personvern, likebehandling, sikkerhet og teknisk robusthet, transparens og en ansvarlig styring. Dette innebærer blant annet å overholde personvernregler, å overvåke KI-systemet for å avdekke eventuelle feil eller skjevheter, og å kunne forklare hvordan algoritmene

²² Brev fra Kultur- og likestillingsdepartementet av 19. februar 2024.

²³ Brev fra Kultur- og likestillingsdepartementet av 19. februar 2024.

²⁴ Se for eksempel Teknologirådet (2024). *Gjennombruddet for generativ kunstig intelligens – en tidslinje*. [Nettartikkel](#)

²⁵ Innst. 191 S (2020–2021).

²⁶ Independent High Level Expert Group on Artificial Intelligence set up by the European Commission (2019): *Ethics guidelines for trustworthy AI*

fungerer og tar beslutninger. Ved å følge opp de etiske prinsippene og vurdere tiltak kan offentlig sektor utvikle og bruke kunstig intelligens på en måte som gagnar samfunnet, respekterer individets rettigheter og oppfyller etiske standarder.²⁷

Mye av effektiviseringspotensialet med KI er basert på automatisering av arbeidsprosesser. Der kunstig intelligens brukes i forbindelse med automatisert saksbehandling, kan sentrale rettssikkerhetsprinsipper bli utfordret, og det er avgjørende med både tilstrekkelig kunnskap om forvaltningsrett i utvikling av systemene, og med teknologisk kompetanse i kontroll av saksbehandlingen og vedtak.²⁸ KI-systemer i kontakt med borgere, for eksempel basert på språkmodeller som gir tilpasset veiledning om saksbehandling, krever kompetanse i befolkningen, og avklart ansvar ved feil på grunn av misforståelser i kommunikasjonen med KI-systemet. KI-systemer som lærer av data kan videreføre og forsterke historiske skjevheter, eller endre ytelse eller praksis over tid. Videre kan det med avanserte maskinlæringsmodeller være utfordrende å få en god forklaring på hvorfor systemet handlet eller konkluderte slik det gjorde. Dermed kan det være vanskelig å sikre åpenhet og forklaringsevne og dermed en etterprøvbar saksbehandling, for å sikre klageretten.²⁹

Viktigheten av kontroll av KI-systemenes resultater kommer for eksempel frem av EUs KI-forordning, som krever at bruk av høyrisiko KI-systemer er tilrettelagt for at mennesker effektivt kan overvåke systemet.³⁰

Undersøkelsen viser varierende modenhetsnivå i hvordan virksomhetene ivaretar de etiske prinsippene som utgjør ansvarlig bruk av kunstig intelligens. Dette gjelder både virksomhetenes rammeverk for å sikre ansvarlig utvikling og bruk av kunstig intelligens og i det praktiske arbeidet med å utvikle og bruke KI-modeller.

En viktig del av å regulere utviklingen og bruken av kunstig intelligens er at virksomhetene etablerer rammeverk med retningslinjer og rutiner for å sikre ansvarlig bruk av kunstig intelligens. Dokumentasjon av beslutninger, avveininger og resultater i utviklingsprosessen og ved bruk er en viktig del av ansvarlig bruk av kunstig intelligens. Dette bidrar til å skape tillit og sikre at KI-systemene opererer innenfor de etiske rammene.³¹

Undersøkelsen viser at virksomhetenes rammeverk for sikring av ansvarlig bruk av kunstig intelligens ivaretar de etiske prinsippene i ulik grad. Dette kommer blant annet fram gjennom svarene i spørreundersøkelsen Riksrevisjonen sendte til kontaktpersoner som er fagansvarlig for KI-systemer i statlige virksomheter. Vi spurte blant annet om hvorvidt det på et overordnet nivå ble brukt skriftlige retningslinjer og lignende i utviklingen av KI-systemer. Svarene er oppsummert i figur 3.

²⁷ Kommunal- og moderniseringsdepartementet (2020), Nasjonale strategi for kunstig intelligens

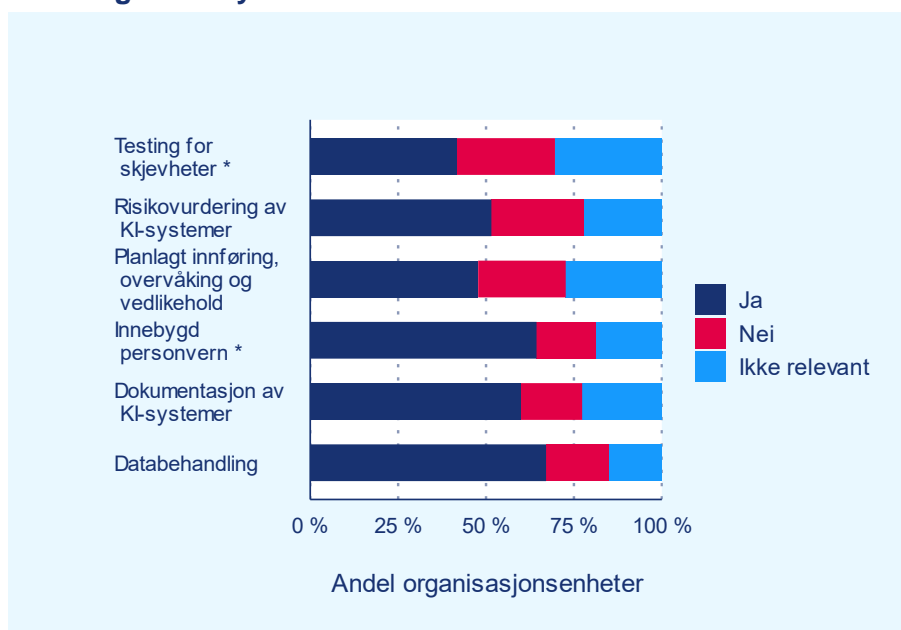
²⁸ Sivilombudsmannen. (2020). *Årsmelding for 2019. Dokument 4 (2019–2020)*.

²⁹ Sivilombudsmannen. (2020). *Årsmelding for 2019. Dokument 4 (2019–2020)*.

³⁰ European Parliament (2024). *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*, [Final draft](#) (2024), artikkel 14

³¹ Kommunal- og moderniseringsdepartementet (2020), Nasjonale strategi for kunstig intelligens

Figur 3 Bruk av skriftlige retningslinjer eller policyer ved utvikling av KI-systemer



Kilde: Riksrevisjonens spørreundersøkelse (n = 104 respondenter, aggregert til 82 organisasjonsenheter).

* Svar for retningslinjer om innebygd personvern og testing for skjevheter i KI-systemets utfall er fremstilt bare for virksomheter med minst ett KI-system som benytter persondata.

Undersøkelsen viser at styringsdokumentene som brukes, i større grad handler om krav til sikkerhet, robusthet og personvern, og i mindre grad handler om krav til likebehandling, forklarbarhet og åpenhet. Figur 3 viser for eksempel at over 60 prosent i spørreundersøkelsen oppgir at de har brukt styringsdokumenter for databehandling og innebygd personvern ved utvikling av kunstig intelligens. Under halvparten har brukt styringsdokumenter som ivaretar krav til blant annet likebehandling som testing for skjevheter i resultater av KI-systemer som benytter persondata.

Videre viser figur 3 at flertallet av virksomhetene benytter standarder for dokumentasjon av KI-systemer. Det varierer i hvilken grad innholdet i standardene for dokumentasjon dekker hensyn til de etiske prinsippene. For flertallet av virksomhetene som har standarder for dokumentasjon, dekker disse datakilder og behandling av data, mens for under halvparten dekker de forklaring av KI-systemets resultater. Vi har gjennomført caseundersøkelser av fire utvalgte KI-prosjekter. Gjennomgangen av disse casene viser tilsvarende svakheter i dokumentasjonen av KI-systemene. For eksempel kommer det ikke klart frem i dokumentasjonen hvilke veivalg som er tatt ved utvikling av KI-modellen, og hvorfor en KI-modell er valgt fremfor andre. I denne sammenhengen etterspør et av KI-prosjektene fra caseundersøkelsen felles maler som kan brukes i KI-utvikling, blant annet basert på risiko ved KI, inkludert det man *ikke* må gjøre eller dokumentere. Prosjektet peker på at dette kunne effektivisere arbeidet med KI i virksomhetene betydelig og åpne for mer innovasjon med KI i små og mellomstore virksomheter.

Det er en rekke risikoer og problemstillinger som særlig gjelder utvikling og bruk av kunstig intelligens og ivaretagelse av de etiske prinsippene. Dette handler for eksempel om hvordan personopplysninger behandles, likebehandling, evne til å forklare hvorfor KI-systemet handlet eller konkluderte slik det gjorde, og algoritmens pålitelighet over tid.

Undersøkelsen av virksomheters retningslinjer for risikovurderinger, som Riksrevisjonen har innhentet fra virksomheter med avanserte KI-systemer, viser at disse er generelt utformet og i ulik grad dekker risikoene forbundet med ansvarlig bruk av kunstig intelligens. Blant annet omfatter retningslinjene hovedsakelig den generelle risikoen ved sikkerhet og beredskap av IT-utstyr, og i ulik grad de særskilte risikoene som er forbundet med automatisering og læring fra data i KI-systemer. Systematisk manipulering av inndata, utilsiktet forskyving av treffsikkerhet over tid og skjevheter som kan gi forskjellsbehandling, er i liten grad dekket i disse retningslinjene. Det varierer dermed om retningslinjene gjør virksomhetene i stand til å avdekke særskilte risikoer forbundet med de etiske prinsippene for ansvarlig bruk av kunstig intelligens.

At de etiske prinsippene ivaretas i ulik grad, går også frem av hvordan det arbeides med kunstig intelligens i virksomhetene. Flertallet av virksomhetene svarte i vår spørreundersøkelse ja på påstander om at de arbeider med sikkerhet og robusthet, mens en lavere andel svarte ja på påstander om åpenhet og likebehandling. Det er for eksempel flere som svarer ja på at de tester for svakheter ved cyberangrep, enn som svarer ja på at KI-systemet begrunner sine beslutninger eller resultater.

For å undersøke nærmere i hvilken grad kravene til de etiske prinsippene blir fulgt opp og ivaretatt i styringen av KI-systemene, gjennomførte Riksrevisjonen en caseundersøkelse av fire KI-systemer. De fire systemene som ble valgt ut, har ulike formål og opererer under ulike rammebetingelser. De fire utvalgte KI-systemene i undersøkelsen brukes som beslutningsstøtte i saksbehandling eller diagnostikk. Tre av KI-systemene er utviklet av virksomheten selv, mens ett er anskaffet fra en leverandør. Undersøkelsen av KI-systemene viser at føringene om ansvarlig bruk av KI ivaretas i varierende grad.

De utvalgte KI-prosjektene har gjennomført flest aktiviteter for å sikre sikkerhet, robusthet og personvern og færre aktiviteter for å sikre likebehandling, forklarbarhet og åpenhet.

To av KI-systemene, som virksomhetene selv har utviklet, behandler og bruker personopplysninger og følger personvernreglene. For eksempel vurderer de hvilke opplysninger som kan brukes i tråd med personvernprinsippene. Arbeidet med personvern i KI-systemer oppleves som tid- og ressurskrevende. Alle de fire KI-systemene gir informasjon internt om funksjon, ytelse og kvalitet og deler informasjon om at KI brukes i virksomheten med eksterne. Imidlertid varierer det hvorvidt KI-systemenes beslutninger og resultater kan forklares. Likebehandling vurderes overordnet i utviklingen av de fire undersøkte KI-systemene. Det er i liten grad gjennomført tester for skjevheter i KI-systemets inndata eller resultater, som forebygging av usaklig forskjellsbehandling. Det er ikke gjennomført noen analyse av hvorvidt KI-systemene fungerer like godt for ulike persongrupper.



De etiske prinsippene for ansvarlig bruk av kunstig intelligens

De etiske prinsippene er i undersøkelsen oppsummert som *personvern, likebehandling, transparens, teknisk sikkerhet og robusthet*, og en prosjektstyring som sikrer at disse prinsippene er ivaretatt i hele livssyklus av KI-systemet (*styring og ansvar*)

KI-prosjektene legger til grunn at likebehandling ivaretas ved at alle er gjenstand for de samme beregningene i KI-systemet, og gjennom at en menneskelig beslutningstaker alltid er involvert. Risikoen for diskriminering er imidlertid også til stede når KI-systemer er brukt som beslutningsstøtte, ikke bare i autonome systemer.

Riksrevisjonens vurdering er at virksomhetene som utvikler og bruker kunstig intelligens, kun delvis har etablert helhetlige rammeverk for ansvarlig bruk av kunstig intelligens. Det varierer hvor langt virksomhetene har kommet i arbeidet med rammeverk som ivaretar ansvarlig KI. De etiske prinsippene blir også i praksis ivaretatt i ulik grad når kunstig intelligens utvikles og brukes. Manglene ved virksomhetenes rammeverk og ved hvordan de utvikler og bruker kunstig intelligens, kan etter Riksrevisjonens vurdering tyde på at det ikke arbeides tilstrekkelig systematisk med prinsipper for likebehandling, forklarbarhet og åpenhet. Videre kan de påviste manglene i virksomhetenes rammeverk føre til at særskilte risikoer forbundet med utvikling og bruk av kunstig intelligens ikke fanges opp eller vurderes. Virksomhetene opplever samtidig at problemstillinger med å ivareta ansvarlig bruk av kunstig intelligens kan være barrierer for bruk og utvikling av kunstig intelligens. Etter Riksrevisjonens vurdering er det mangel på kunnskap og støtte for at statlige virksomheter bedre skal kunne ivareta likebehandling, åpenhet og forklarbarhet i arbeidet med kunstig intelligens. Denne mangelen kan gi risiko for at de etiske prinsippene for kunstig intelligens ikke etterleves.

4.4 Samordningen av arbeidet med kunstig intelligens i offentlig sektor er mangelfull, og den samlede innsatsen er for svak gitt ambisjonen om at Norge skal ha en infrastruktur for kunstig intelligens i verdensklasse

I behandlingen av Meld. St. 30 (2019–2020) var kommunal- og forvaltningskomiteen svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, i form av digitaliseringsvennlig regelverk, gode språkressurser, raske og robuste kommunikasjonsnett og tilstrekkelig regnekraft. Disse overordnede målene er konkretisert i den nasjonale strategien for kunstig intelligens fra 2020. Strategien inneholder blant annet 51 tiltak som regjeringen ville gjennomføre.

Digitaliserings- og forvaltningsdepartementet opplyser at de fleste tiltakene i strategien er gjennomført eller påbegynt, men at det er ikke gjennomført noen systematisk evaluering av gjennomføringen av strategien. Det foreligger heller ingen oversikt over bruken av kunstig intelligens i staten. Etter Riksrevisjonens vurdering har Digitaliserings- og forvaltningsdepartementet dermed ingen god informasjon om utviklingen av bruk av kunstig intelligens i offentlig sektor. Det finnes en ufullstendig oversikt over KI-prosjekter i offentlig sektor som er gjennomført etter initiativ fra flere høyere norske utdannings- og forskningsinstitusjoner (kjent under navnet NORA) og senere i samarbeid med Digitaliseringsdirektoratet. Det er

imidlertid ikke lagt opp til noen systematisk oppdatering av denne kartleggingen.

Digitaliserings- og forvaltningsdepartementet opplyser at det gjennom den gjeldende nasjonale KI-strategien ikke er lagt opp til en satsing på kunstig intelligens. Departementet presiserer imidlertid at det er et politisk mål å ta i bruk kunstig intelligens til effektivisering i offentlig sektor og for verdiskaping i samfunnet. Riksrevisjonen merker seg at digitaliserings- og forvaltningsministeren våren 2024 definerte et mål om betydelig økt bruk av kunstig intelligens i offentlig sektor.

Departementet viser til at selv om det har samordningsansvaret for digitaliseringssatsingen i staten gjelder sektorprinsippet og følgelig at den enkelte sektoren er ansvarlig for utvikling og bruk av kunstig intelligens innenfor eget ansvarsområde. Departementet kan imidlertid legge til rette for bruk av kunstig intelligens gjennom generelle virkemidler – særlig gjennom pedagogiske virkemidler som veiledning. Veiledningsarbeidet skjer først og fremst gjennom Datatilsynet og Digitaliseringsdirektoratet.

Digitaliseringsdirektoratet skal være regjeringens fremste verktøy for effektiv og samordnet digitalisering av offentlig sektor og av samfunnet øvrig. Digitaliseringsdirektoratets eneste oppdrag om å følge opp den nasjonale KI-strategien var å utarbeide en veileder i ansvarlig bruk av kunstig intelligens i offentlig sektor. Veiledningsmateriellet ble publisert i 2023 og er samlet på Digitaliseringsdirektoratets hjemmesider. Veiledningsmaterialet er beskrevet som fortsatt under utvikling med anmodning om innspill. Veilederen er på et overordnet nivå, med lenker til mer detaljert materiell der det er tilgjengelig, som fra Datatilsynet og Likestillings- og diskrimineringsombudet. Det er lite konkret veiledning om hvordan de etiske prinsippene kan ivaretas i praksis. Eksempelvis finnes det ingen omforent metode for å avdekke algoritmeskjevheter.³² Gitt dette og de påviste manglene i virksomhetenes arbeid med ansvarlig kunstig intelligens er det etter Riksrevisjonens vurdering fortsatt et stort behov for konkret veiledning for å kunne sikre at de overordnede føringene om ansvarlig kunstig intelligens etterleveres.

KI-prosjekter kan enkeltvis få mer konkret veiledning om juridiske spørsmål gjennom Datatilsynets regulatoriske sandkasse. Som vist tidligere, har kun fem prosjekter i statlig regi deltatt siden opprettelsen av sandkassen i november 2020. Den eksterne evalueringen viser dessuten at sandkassen har et forbedringspotensial til å utarbeide og spre prosjektovergripende innsikt.

Digitaliserings- og forvaltningsdepartementet viser til at deres pådriverrolle innen kunstig intelligens innebærer å identifisere tverrgående behov og få på plass felles løsninger på tvers av sektorene. Det innebærer også at departementet arbeider for å sikre bedre koordinering ved identifiserte behov. Riksrevisjonen merker seg at Digitaliserings- og forvaltningsdepartementet³³ allerede i 2018 opprettet en interdepartemental



NORA

Norwegian Artificial Intelligence Research Consortium - NORA er et norsk samarbeid mellom åtte universiteter, fem høyskoler og fem forskningsinstitutter innenfor KI, maskinlæring og robotikk

³² Algoritmeskjevheter betyr at algoritmen ikke fungerer like godt for alle, slik at KI-systemet gjør systematisk flere feil for en persongruppe enn for andre, for eksempel at systemet er mindre treffsikkert for noen persongrupper enn for andre. (Datatilsynet (2023). Ahus, sluttrapport: Hjertekom for etisk AI. <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/ferdige-prosjekter-og-rapporter/ahus-sluttrapport-ekg-ai>)

³³ Tidligere Kommunal- og moderniseringsdepartementet.

arbeidsgruppe for kunstig intelligens, som ble videreført med et nytt mandat våren 2024. Alle departementene er representert i den videreførte arbeidsgruppen. Formålet med arbeidsgruppen er å sikre bedre koordinering og informasjonsutveksling på KI-området. Riksrevisjonen merker seg at arbeidsgruppen som regel ikke diskuterer utfordringer med utvikling eller bruk av kunstig intelligens i underliggende etater og virksomheter. Etter Riksrevisjonens vurdering er det vanskelig å se hvordan Digitaliserings- og forvaltningsdepartementet har benyttet denne møteplassen til noe annet enn å utveksle informasjon på tvers av departementene. Deling av informasjon er sentralt, men i seg selv ikke et tilstrekkelig tiltak for å sikre god koordinering. Riksrevisjonen merker seg for øvrig at arbeidsgruppen skal bli brukt videre i arbeidet med den forestående nye nasjonale digitaliseringsstrategien.

I intervju våren 2024 viser Digitalisering- og forvaltningsdepartementet til at de nå ser at kompetanse- og veiledningsbehovet er en felles problemstilling. Departementet mener at sektorlovgivningen er ulik, men er samtidig usikker på hvor langt departementet skal gå inn i den konkrete sektoren, og hvor relevante problemstillingene i én sektor er for andre. Riksrevisjonen vil peke på at alle disse problemstillingene har vært sentrale over flere år.

De enkelte statlige sektorene og etatene må selv vurdere om og eventuelt i hvilket omfang kunstig intelligens er en teknologi som bør benyttes for eksempelvis å forbedre og effektivisere egen oppgaveløsning. Samtidig er det etter Riksrevisjonens vurdering sentralt med sektorovergrepene tiltak for å kunne sikre en god felles infrastruktur for kunstig intelligens. Etter Riksrevisjonens vurdering kan ikke utviklingen overlates til hver enkelt sektor og hvert enkelt departement.

Noen grunnleggende forutsetninger må være på plass for at man skal lykkes. Det gjelder eksempelvis å sikre god datakvalitet og deling av data, teknologisk infrastruktur og kompetanse. Gode nasjonale og sektorielle fellesløsninger kan være et viktig bidrag til det arbeidet. Videre gjelder det å sikre en felles forståelse av og tilnærming til juridiske spørsmål om personvern. Undersøkelsen viser at det på disse områdene er utfordringer som er til hinder for å i større grad kunne ta i bruk kunstig intelligens i offentlig sektor. Riksrevisjonen vil videre peke på uklarheter rundt hvordan de etiske prinsippene kan ivaretas i praksis. Det innebærer for eksempel hva algoritmisk diskriminering kan bety, og hvilke personopplysninger som er nødvendig som grunnlag for å teste og forebygge diskriminering med kunstig intelligens. Etter Riksrevisjonens vurdering er det flere overordnede spørsmål rundt utvikling og bruk av kunstig intelligens som per i dag den enkelte virksomhet og sektor i for stor grad må håndtere på egen hånd.

Gitt de store mulighetene kunstig intelligens gir, og utfordringene med utvikling og bruk av kunstig intelligens i staten som er påvist, mener Riksrevisjonen at Digitaliserings- og forvaltningsdepartementet ikke i tilstrekkelig grad har ivaretatt rollen som samordningsdepartement i tilrettelegging for bruk av kunstig intelligens. Digitaliserings- og forvaltningsdepartementet sikrer informasjonsdeling om kunstig intelligens på overordnet nivå, men det er gjort få tiltak som skal løse de overordnede utfordringene med utvikling og bruk av kunstig intelligens. Selv om alle

sektorer og departementer har et ansvar for å sikre måloppnåelse på dette området, vil Norge ikke få en infrastruktur i verdensklasse for kunstig intelligens uten en tydelig og felles koordinert innsats. Etter Riksrevisjonens vurdering kan ikke dette overlates til den enkelte sektoren.

Riksrevisjonen mener derfor at det ikke er tilfredsstillende at Digitaliserings- og forvaltningsdepartementet gjennom styring og i samarbeid med de andre departementene ikke har lagt bedre til rette for at offentlig sektor kan utnytte mulighetene som kunstig intelligens gir, og ta i bruk kunstig intelligens på en ansvarlig måte. En ren videreføring av dagens samordningsaktiviteter vil etter Riksrevisjonens vurdering ikke kunne sikre at det store potensialet for bruk av kunstig intelligens innenfor ansvarlige rammer blir utnyttet i offentlig sektor. Den samlede innsatsen er for svak gitt ambisjonen om at Norge skal ha en infrastruktur for kunstig intelligens i verdensklasse.

5 Anbefalinger

Riksrevisjonen anbefaler at Digitaliserings- og forvaltningsdepartementet styrker sin samordnings- og pådriverrolle for aktivt å stimulere til og tilrettelegge for ansvarlig bruk av kunstig intelligens i offentlig sektor i samarbeid med andre departementer gjennom å:

- videreutvikle virkemiddelapparatet for å sikre at grunnleggende forutsetninger for ansvarlig bruk av kunstig intelligens er på plass, slik som tilstrekkelig digital infrastruktur, felles digitale løsninger, tilgang til data av god kvalitet og relevant tverrfaglig kompetanse.
- bidra til regulatoriske avklaringer og en samordnet tolkning av regelverket i utviklingen og bruken av kunstig intelligens
- sikre en samordnet forståelse av de etiske prinsippene og veiledning om hvordan prinsippene kan ivaretas i praksis i utviklingen og bruken av kunstig intelligens.

6 Statsrådets svar

Dokument 3:18 (2023–2024) Riksrevisjonens undersøkelse av bruk av kunstig intelligens i staten ble oversendt statsråden i Digitaliserings- og forvaltningsdepartementet.

Svaret fra statsråden følger i sin helhet i vedlegg 2.

7 Riksrevisjonens uttalelse til statsrådets svar

Riksrevisjonen har ingen ytterligere merknader.

Saken sendes Stortinget.

Vedtatt i Riksrevisjonens møte 20. august 2024.

Karl Eirik Schjøtt-Pedersen

Tom-Christer Nilsen

Helga Pedersen

Anne Tingelstad Wøien

Arve Lønnum

Jens A. Gunvaldsen

Vedlegg

Vedlegg 1:

Riksrevisjonens brev til statsråden i Digitaliserings- og forvaltningsdepartementet



Riksrevisjonen

Vår saksbehandler

Carolin Prabhu 22241312

Vår dato

21.06.2024

Deres dato

Vår referanse

2023/00115-229

Deres referanse

Utsatt offentlighet jf. rrevl § 18 (2)

DIGITALISERINGS- OG
FORVALTNINGSDEPARTEMENTET (DFD)
Postboks 8004 Dep
0030 OSLO

Ved statsråd Karianne Oldernes Tung

Riksrevisjonens undersøkelse av bruk av kunstig intelligens i staten - oversendelse av utkast til Dokument 3:X (2023-2024)

Vedlagt er utkast til Dokument 3:X (2023–2024) *Riksrevisjonens undersøkelse av bruk av kunstig intelligens i staten*.

Dokumentet er basert på rapporten oversendt Digitaliserings- og forvaltningsdepartementet ved vårt brev 24. april 2024 og på departementets svar av 27. mai 2024. Dokumentet er også basert på svar på rapporten fra de øvrige departementene som er berørt av denne revisjonen, Arbeids- og inkluderingsdepartementet, Finansdepartementet, Helse- og omsorgsdepartementet, Justis- og beredskapsdepartementet og Kultur- og likestillingsdepartementet.

Statsråden bes redegjøre for hvordan departementet vil følge opp Riksrevisjonens merknader og anbefalinger, og eventuelt om departementet er uenig med Riksrevisjonen.

Statsrådets svar vil i sin helhet bli vedlagt dokumentet. Det bes om at svaret oversendes som PDF lagret fra Word, ikke skannet som bilde, slik at innholdet kan gjøres tilgjengelig for alle i samsvar med krav til universell utforming.

Svarfrist: 9. august 2024

For riksrevisorkollegiet

Karl Eirik Schjøtt-Pedersen
riksrevisor

Brevet er godkjent og ekspedert digitalt.

Vedlegg 2:

Statsrådets svar



DET KONGELIGE DIGITALISERINGS-
OG FORVALTNINGSDEPARTEMENT

Statsråden

Riksrevisjonen
Postboks 6835 St. Olavs plass
0130 OSLO

Unntatt offentlighet,
offl. § 5 andre ledd jf. rr. lov § 18(2)

Deres ref

Vår ref

Dato

24/1791-11

07. august 2024

Riksrevisjonens undersøkelse av bruk av kunstig intelligens i staten

Jeg viser til brev fra Riksrevisjonen 21. juni 2024 om Riksrevisjonens undersøkelse av bruk av kunstig intelligens i staten. Riksrevisjonen ber meg redegjøre for hvordan departementet vil følge opp Riksrevisjonens merknader og anbefalinger, og eventuelt om departementet er uenig med Riksrevisjonen.

Tilbakemelding på Riksrevisjonens konklusjoner og anbefalinger

Det er et stort potensial for offentlig sektor ved å ta i bruk kunstig intelligens (KI). Gjennom å bruke KI kan offentlig sektor redusere behovet for arbeidskraft, effektivisere og øke produktivitet, samt gi bedre tjenester til innbyggere og næringsliv. Jeg er derfor glad for at Riksrevisjonen har undersøkt bruken av KI i staten. Regjeringen har en klar ambisjon om at Norge skal ha en KI-infrastruktur i verdensklasse og at offentlig sektor skal ha høy bruk av KI. Riksrevisjonens rapport bidrar med verdifull innsikt og kunnskap i arbeidet med å realisere denne ambisjonen.

Riksrevisjonens overordnede vurdering er at det ikke er tilfredsstillende at Digitaliserings- og forvaltningsdepartementet gjennom styring og i samarbeid med andre departement ikke har lagt bedre til rette for at offentlig sektor kan utnytte mulighetene med kunstig intelligens og ta det i bruk på en ansvarlig måte. Rapporten fremhever også at viktige forutsetninger for å ta i bruk kunstig intelligens ikke er på plass. Det er et stort behov for å avklare juridiske spørsmål knyttet til bruken av kunstig intelligens, det er mangelfull infrastruktur og tilgang til data, og det mangler kompetanse i offentlig sektor.

Jeg deler Riksrevisjonens vurdering av at offentlig sektor må bli bedre til å utnytte mulighetene som ligger i kunstig intelligens. Riksrevisjonens undersøkelse omfatter perioden 2018 – 2023. I 2024 opprettet regjeringen Digitaliserings- og forvaltningsdepartementet

(DFD). Dette ble gjort fordi regjeringen identifiserte mange av de samme svakhetene og utfordringene som påpekes i Riksrevisjonens rapport, og så behovet for å styrke samordnings- og pådriverkraften i digitaliseringspolitikken. Med opprettelsen av DFD ønsker regjeringen å legge større trykk og kraft inn i digitaliseringsarbeidet, samt sørge for bedre samordning og styring.

Opprettelsen av DFD er viktig for regjeringen fordi det muliggjør en mer helhetlig tilnærming til digitalisering og kunstig intelligens på tvers av sektorer. Sektorprinsippet står sterkt i norsk forvaltning, der hvert departement har ansvar for sin egen sektor. Dette kan skape hindringer for effektiv digitalisering, ettersom utfordringene med digitalisering ofte er felles på tvers av sektorer. Ved å etablere et eget departement med ansvar for digitalisering og forvaltning, ønsker regjeringen å tydeliggjøre koordinerings- og pådriverrollen og sørge for implementering av digitaliseringsinitiativer som gjelder hele offentlig sektor. Dette inkluderer også utvikling og bruk av kunstig intelligens, der potensialet for effektivisering og forbedring av offentlige tjenester er stort. Offentlig sektor vil i årene fremover ha behov for mer arbeidskraft. Økt bruk av KI vil være en del av løsningen på arbeidskraftmangelen. DFD vil derfor ha en sentral rolle i å drive frem en mer sammenhengende og koordinert digitaliseringspolitikk som kan møte de utfordringene og utnytte de mulighetene som digital teknologi og kunstig intelligens gir.

Regjeringen har høye ambisjoner for digitalisering i offentlig sektor og vil legge fram en ny nasjonal digitaliseringsstrategi høsten 2024. Denne strategien skal stake ut kursen for forsterket digitalisering av offentlig sektor og tilrettelegge for økt tempo i digitaliseringen av næringslivet og samfunnet for øvrig. Som en del av dette arbeidet vil vi se på hvordan vi skal utvikle, bruke og regulere kunstig intelligens i Norge, samt hvordan vi kan samordne innsatsen for å oppnå disse målene. Jeg har blant annet satt et mål om at 80 prosent av offentlig sektor skal ta i bruk kunstig intelligens innen 2025, fordi det er verktøyet vi trenger for å sikre bedre, tryggere og enklere tjenester.

Jeg vil trekke frem noen sentrale initiativer og tiltak i den kommende digitaliseringsstrategien som adresserer konklusjonene i Riksrevisjonens rapport.

For å ivareta samordningsansvaret på best mulig måte og styrke pådriverrollen, skal regjeringen videreutvikle eksisterende og utvikle nye virkemidler på digitaliseringsområdet. Flere ulike elementer kan inngå i dette arbeidet, blant annet Digdir, Skate, Stimulab, medfinansieringsordningen, digitaliseringsrundskrivet og annet regelverk. Digitalisering gir også økt behov for å ivareta tverrgående hensyn og investeringer på tvers av sektorer i budsjettprosessen. Jeg vil i denne sammenheng trekke frem medfinansieringsordningen for statlige digitaliseringstiltak som et viktig virkemiddel for å styrke gjennomføringskraften for tverrsektorielle digitaliseringstiltak.

Jeg vil, i samarbeid med forsknings- og høyere utdanningsministeren, kultur- og likestillingsministeren og næringsministeren, arbeide for å etablere en nasjonal infrastruktur for KI. Infrastrukturen vil blant annet omfatte tilgang til regnekraft og språkmodeller tilpasset

norske og samiske språk og norske samfunnsforhold. Superdatamaskiner utgjør en nødvendig infrastruktur for utvikling av norsk KI. Svært stor regnekraft er nødvendig for å trene opp grunnmodeller for KI som deretter kan videreutvikles av næringslivet eller offentlig sektor for å lage og bruke KI-verktøy for mer effektive tjenester og produksjon.

EU vedtok nylig en forordning om kunstig intelligens som legger til rette for trygg, tillitsvekkende og etisk bruk av KI. Departementet vurderer denne forordningen som EØS-relevant og har startet arbeidet med å implementere den i norsk lov. Departementet har også innledet dialog med EØS-partnere for en smidig innlemming av forordningen i EØS-avtalen. Dette regelverket vil gi økt forutsigbarhet for norske virksomheter som benytter KI-teknologi.

For å håndheve forordningen er det krav til nasjonal forvaltningsstruktur. Jeg vil motta en rapport med anbefalinger til nasjonal organisering fra Direktoratet for forvaltning og økonomistyring (DFØ) etter sommeren 2024. Regjeringen vil følge opp DFØs anbefalinger og tar sikte på å ha på plass en nasjonal forvaltningsstruktur innen 2026.

Riksrevisjonens undersøkelse viser at det er vanskelig å håndtere og ivareta personvern hensyn og etiske prinsipper i arbeidet med å utvikle systemer for KI, og at veiledningsbehovet er stort på dette området. Vi vil bygge videre på det viktige arbeidet med KI og personvern som Datatilsynet har gjort i sin regulatoriske sandkasse. Vi vil også videreutvikle veiledningen Digitaliseringsdirektoratet gir i utvikling og bruk av KI og se nærmere på hvordan vi ytterligere kan samordne veiledningen på kunstig intelligens i offentlig sektor.

Det er viktig å bygge kunnskap, blant annet om hvilke konsekvenser teknologiutviklingen har for samfunnet vårt, slik at vi kan styre utviklingen i den retningen vi ønsker. Regjeringen øker derfor forskningsinnsatsen på kunstig intelligens og digitale teknologier med en milliard kroner de neste fem årene. Forskningsmilliarden skal brukes til forskning på selve teknologien, men også bidra til større innsikt om hvilke konsekvenser teknologiutviklingen har for folk og samfunn, samt gi mer kunnskap om muligheter for innovasjon i næringslivet og offentlig sektor. Forskningsrådet forvalter midlene og i løpet av 2025 etableres 4 - 6 KI-senter. Resultatene fra KI-sentrene skal også komme offentlig sektor til gode.

Tilgang på digital kompetanse er en stor utfordring for utvikling og bruk av KI i offentlig sektor. Det viser også Riksrevisjonens undersøkelse. Jeg mener at løsningen ikke bare er å utdanne flere teknologer. Det handler også om innovasjon, ledelse og kultur. For bedre å kunne utnytte mulighetene som ligger i kunstig intelligens vil jeg, i den kommende digitaliseringsstrategien, vurdere hvordan vi kan øke den digitale kompetansen i offentlig sektor.

Deling og bruk av data er avgjørende for å utnytte potensialet i KI og maskinlæring. I juni mottok jeg Viderebruksutvalgets rapport NOU 2024: 14 *Med lov skal data deles*, som foreslår en ny lov om datadeling for å gjøre offentlige data enklere å finne og ta i bruk for å lage nye produkter, tjenester og løsninger. Jeg vil raskt følge opp utvalgets forslag.

Som Riksrevisjonen påpeker, er hvert departement ansvarlig for digitaliseringen i sin sektor. DFD har et viktig mandat om å identifisere sektorovergrepene og koordinere tiltak. Jeg vil bruke dette mandatet til å følge opp Riksrevisjonens anbefalinger i tett samarbeid med andre departementer. Riksrevisjonens undersøkelse viser at det fortsatt gjenstår arbeid for at offentlig sektor skal kunne realisere potensialet i kunstig intelligens. Styrkingen av DFDs samordnings- og pådriverrolle er derfor essensiell. Regjeringen skal gjennom målrettet innsats og tett samarbeid nå ambisjonen om å bygge en infrastruktur for kunstig intelligens i verdensklasse og bruke kunstig intelligens i offentlig sektor til det beste for innbyggerne og samfunnet.

Med hilsen



Karianne Oldernes Tung

Vedlegg 3:

Forvaltningsrevisjonsrapport med vurderinger

Revisjonen er gjennomført som en forvaltningsrevisjon i henhold til

- lov om Riksrevisjonen § 9 tredje ledd
- instruks om Riksrevisjonens virksomhet § 9
- INTOSAI's standard for forvaltningsrevisjon (ISSAI 3000)
- Riksrevisjonens faglige retningslinjer for forvaltningsrevisjon

Innhold

1	Innledning	6
1.1	Bakgrunn	6
1.2	Mål og problemstillinger	8
1.3	Avgrensning	8
1.4	Definisjoner av kunstig intelligens og ansvarlig KI	8
1.4.1	Definisjon av ansvarlig kunstig intelligens.....	9
2	Metodisk tilnærming og gjennomføring	11
2.1	Kartlegging av bruk av KI i statlig forvaltning	11
2.1.1	Brev for å kartlegge KI-aktivitet i staten	11
2.1.2	Spørreundersøkelse om konkrete KI-prosjekter og hvordan virksomhetene arbeider med KI	12
2.2	Caseundersøkelser av utvalgte KI-systemer	14
2.3	Intervju.....	16
2.4	Dokumentgjennomgang	16
3	Revisjonskriterier	18
3.1	Overordnede mål og ansvar.....	18
3.2	Krav til ansvarlig kunstig intelligens	20
3.2.1	Krav til personvern og respekt for menneskets selvbestemmelse og kontroll.....	21
3.2.2	Krav om likebehandling.....	23
3.2.3	Krav om teknisk sikkerhet og robusthet	24
3.2.4	Krav om forklarbarhet og åpenhet	25
3.2.5	Krav om åpenhet i saksbehandling.....	25
3.3	Krav til departementenes styring.....	26
4	Bruk av kunstig intelligens i staten	28
4.1	Stor variasjon i bruk av KI blant statlige virksomheter	28
4.1.1	Hvor mye investerer statlige virksomheter i bruk og utvikling av KI?.....	30
4.2	Over halvparten av KI-systemene bruker kun enkle modeller for å tolke data	31
4.3	Virksomhetenes formål med bruk av kunstig intelligens	33
4.3.1	De fleste virksomhetene oppgir at KI-prosjektene har gitt ønsket forbedring.....	35
4.4	Bruk av personopplysninger og betydning for enkeltpersoner.....	37
5	Utfordringer med å utvikle og ta i bruk kunstig intelligens	40
5.1	Mange opplever at regelverk og personvern er vanskelig i arbeidet med kunstig intelligens .	41
5.1.1	Hjemmelsgrunnlag for å behandle personopplysninger	41
5.1.2	Innebygd personvern og personvernkonsekvenser	43
5.2	Mangelfull tilgang på data av god kvalitet	45
5.2.1	Deling av data	46
5.2.2	Arbeidet med å tilgjengeliggjøre helsedata.....	48
5.3	Behov for infrastruktur og andre tekniske ressurser	49

5.4	Behov for kompetanse på KI.....	51
5.5	Betydningen av språkressurser på norsk.....	54
6	Statlige virksomheter tar hensyn til de etiske prinsippene i ulik grad i utvikling og bruk av kunstig intelligens.....	56
6.1	De utvalgte KI-systemene er mindre utviklingsprosjekter med lav risiko som brukes som beslutningsstøtte i saksbehandlingen	56
6.2	Suksessfaktorer og utfordringer for utvikling av kunstig intelligens	58
6.3	Retningslinjer, policyer og arbeidet med ansvarlig KI ivaretar de etiske prinsippene i ulik grad	59
6.3.1	Styringsdokumenter brukt ved utviklingen av KI dekker i ulik grad de etiske prinsippene for ansvarlig KI.....	59
6.3.2	Virksomhetenes arbeid med KI-utvikling ivaretar ansvarlig KI i ulik grad	63
6.3.3	KI-prosjektene bruker mest generelle retningslinjer og policyer som i ulik grad regulerer ansvarlig KI.....	64
6.4	Etiske prinsipper om personvern, sikkerhet og robusthet ivaretas i større grad enn prinsippene om transparens og likebehandling	65
6.4.1	KI-systemene er tydelig forankret i virksomhetene og kan vise til positive resultater, men dokumentasjon av KI-systemet har svakheter	66
6.4.2	Personvernprinsippene er ivaretatt ved utvikling av KI-systemene	68
6.4.3	Likebehandling vurderes på et overordnet nivå, og testing av skjevheter varierer blant KI-systemene	70
6.4.4	KI-systemene er i varierende grad tilrettelagt for åpenhet og forklarbarhet	71
6.4.5	Robusthet og sikkerhet er i stor grad ivaretatt i KI-systemene	72
6.4.6	Forskjeller i hvordan ansvarlig KI ivaretas ved anskaffelser og selvutvikling av KI-systemer .	73
7	Tilrettelegger Digitaliserings- og forvaltningsdepartementets samordning for ansvarlig bruk av KI i statlige virksomheter?	74
7.1	Digitaliserings- og forvaltningsdepartementets ansvar og sektorprinsippet	74
7.2	Regjeringens strategi for kunstig intelligens og oppfølgingen av denne.....	74
7.2.1	Fremskaffelse av styringsinformasjon.....	75
7.2.2	En gjennomgang av regelverket	76
7.2.3	Datatilsynets regulatoriske sandkasse for å forstå kravene til personvern bedre	78
7.3	Digitaliserings- og forvaltningsdepartementets pådriverrolle og tiltak utover KI-strategien	82
7.3.1	Digitaliseringsdirektoratets arbeid med kunstig intelligens	83
8	Vurderinger	86
8.1	Statlige virksomheter og heleide selskap utnytter mulighetene ved kunstig intelligens ulikt, og kunstig intelligens er fortsatt lite i bruk	86
8.2	Viktige forutsetninger for å ta KI i bruk i stor skala i statlige virksomheter er ikke på plass	87
8.2.1	Regulatoriske uklarheter blir ikke avklart	87
8.2.2	Mangelfull infrastruktur og tilgang på data av god kvalitet.....	89
8.2.3	Stort behov for kompetanse	90
8.3	De etiske prinsippene for ansvarlig bruk av KI ivaretas i ulik grad. Kontrollmekanismer som sikrer ansvarlig bruk av kunstig intelligens må være på plass.....	90

8.3.1	Virksomhetene har kun delvis etablert systemer for sikring av ansvarlig bruk av KI	90
8.3.2	Ved utvikling og bruk av KI ivaretas hensynet til sikkerhet, robusthet og personvern i større grad enn transparens og likebehandling	91
8.4	Manglende samordning i arbeidet med kunstig intelligens i offentlig sektor	92
9	Referanseliste	95
10	Vedlegg	102
10.1	Vedlegg 1: Brev med spørsmål	102
10.2	Vedlegg 2: Spørreskjema.....	104
10.3	Vedlegg 3: Rammeverk for vurdering av ansvarlig KI.....	120
10.3.1	Underspørsmål vurdert i caseundersøkelser	120
10.3.2	Operasjonalisering av rammeverket	123
10.3.3	Styringsprinsipper vurdert i rammeverket	123
10.4	Vedlegg 4: Bruksområder for KI i statlige virksomheter og heleide selskaper	127

Figuroversikt

Figur 1	Antall svar på spørrebrev fordelt på fagdepartement.....	12
Figur 2	Antall KI-systemer rapportert, fordelt på fagdepartementer, fra 2018 til vår 2023.....	29
Figur 3	Antall KI-systemer per virksomhet	30
Figur 4	Type modeller som brukes for å tolke data for de ulike virksomhetstypene	32
Figur 5	Oversikt over hvem som har bidratt til utvikling av modellene, fordelt på virksomhetstype.....	33
Figur 6	Bruk av KI-systemer til ulike oppgaver	34
Figur 7	Formål med bruk av KI-systemer	35
Figur 8	Måloppnåelse for KI-systemer i bruk.....	37
Figur 9	Andel KI-systemer av ulike typer som behandler personopplysninger, helsesektoren unntatt	38
Figur 10	Kan enkeltpersoner bli direkte berørt av resultatet fra KI-systemet?	39
Figur 11	Oppløst utfordringer med å utvikle og bruke kunstig intelligens. Prosent	40
Figur 12	Verktøy/infrastruktur brukt i KI-systemene	50
Figur 13	Bruk av skriftlige retningslinjer eller policyer ved utvikling av KI-systemer	61
Figur 14	Modenhetsnivå av selvutviklede KI-systemene i caseundersøkelser om ansvarlig KI - styring og ivaretagelse av etiske prinsipper	66

Tabelloversikt

Tabell 1	Oversikt over case	14
Tabell 2	Oversikt over sentrale kjennetegn ved de utvalgte KI-systemene	57
Tabell 3	Hovedresultater og gevinster ved utvalgte KI-systemer	67
Tabell 4	Søknader og deltakelse i den regulatoriske sandkassen, 2020–2024.....	80

Faktaboksoversikt

Faktaboks 1	Interdepartemental gruppe for kunstig intelligens.....	52
Faktaboks 2	Regulatorisk sandkasse.....	79

Ordliste og forkortelser

Datakvalitet er et relativt begrep som betegner hvor godt data oppfyller kriterier til nøyaktighet, fullstendighet, gyldighet, konsistens, aktualitet, relevans for og egnethet til et formål.

Dyplæring er en type maskinlæring med nevralt nettverk som består av flere lag, hvor hvert lag abstraherer informasjon videre.

Felles datakatalog skal gi oversikt over alle beskrivelser av datasett, begreper, API-er og informasjonsmodeller norske virksomheter besitter, som de kan dele med andre.

Heleide selskaper betegner i undersøkelsen alle statlige selskaper som er heleid av staten, unntatt helseforetakene.

Helseforetak er i undersøkelsen avgrenset fra andre heleide selskaper og virksomheter. Virksomhetene og selskapene som er regnet som helseforetak, er under Helse- og omsorgsdepartementet (HF, RHF- regionale helseforetak, teknologiforetak), med unntak av helsetjenesteforvaltning (for eksempel Helsedirektoratet), som regnes som virksomheter, og Vinmonopolet, som regnes som heleid selskap.

Helseforetak (HF) er en virksomhet som er eid av et regionalt helseforetak. Helseforetaket yter spesialisthelsetjenester til befolkningen i sitt opptaksområde. I Norge er det 20 helseforetak som yter somatiske helsetjenester. Et helseforetak kan eie ett eller flere sykehus.

Informasjonssikkerhet er hvordan informasjonen blir beskyttet mot uønsket innsyn (konfidensialitet), hvordan informasjonen gjøres tilgjengelig når det er ønsket (tilgjengelighet), og hvordan informasjonen er beskyttet mot manipulering (integritet). **Teknisk sikkerhet** handler om hvordan virksomheten ivaretar sikkerheten til infrastruktur, og hvilke systemer den har implementert for å håndtere krav til informasjonssikkerhet.

Kunstig intelligens (KI) er informasjonsteknologi som utfører handlinger basert på tolkning og behandling av strukturerte eller ustrukturerte data i den hensikt å oppnå et gitt mål. Undersøkelsen følger dermed definisjonen i den nasjonale strategien for kunstig intelligens. Tolking av data kan være fullstendig regelbasert (programvarerobot), basert på en kunnskapsbase (ekspertsystem) eller basert på data gjennom statistisk modellering eller maskinlæring.

Et **KI-prosjekt** omfatter organisering av et KI-system, fra planlegging til produksjon eller avvikling. Begrepet er uavhengig av en formell struktur med prosjektnummer eller lignende, men kan defineres ut fra målet KI-systemet skal bidra til.

Et **KI-system** omfatter hele anvendelsen fra innhenting av data til sluttresultat, ikke bare KI-komponenten som tolker data med for eksempel en maskinlæringsmodell.

1 Innledning

1.1 Bakgrunn

Kunstig intelligens (KI) er et viktig verktøy for utvikling av en bærekraftig, effektiv og brukerrettet offentlig forvaltning.¹ Potensialet for effektivisering av offentlig sektor med bruk av kunstig intelligens er stort. I Norge anslår for eksempel det forskningsbaserte analyse- og rådgivingselskapet Menon at potensiell årlig verdiskapning i norsk økonomi med kunstig intelligens gjennom automatisering er totalt sett mellom 500 og 600 milliarder kroner. Det er særlig stort effektiviseringspotensial innen saksbehandling i offentlig sektor. Ved full utnyttelse av dagens teknolog for kunstig intelligens, kan offentlig sektor effektivisere arbeidsoppgaver som årlig vil tilsvare 155 000 årsverk, ifølge rapporten.² Internasjonalt er det også flere studier som peker i samme retning. Ifølge en rapport fra Alan Turing Institute i Storbritannia kan over 80 prosent av 143 millioner komplekse repetitive handlinger i britisk statsforvaltning automatiseres ved hjelp av KI.³ McKinsey & Company anslår betydelig potensial for produktivitetøkning med generativt kunstig intelligens i offentlig sektor, inklusive helsetjenester og utdanning, med blant annet 40 prosent av arbeidstidene påvirket gjennom KI-systemer som oppsummerer og analyserer tekst, og 60 prosent automatiseringspotensial i kunderettete tjenester innen fem til ti år gjennom teknologi som for eksempel chatbots.⁴

Med utfordringer som vesentlig færre yrkesaktive per pensjonist og oljeinntekter som vil bety mindre enn tidligere, samtidig som utgiftene til alderspensjoner og helserelaterte ytelser øker, vil det være avgjørende at Norge har en *effektiv stat*. Kunstig intelligens har potensial til å gjøre staten betydelig mer produktiv⁵ og bidra til å endre statlig forvaltning på helt grunnleggende måter. Samtidig viser IT i praksis undersøkelsen⁶ at den tradisjonelt høye tilliten i befolkningen til offentlige myndigheter ikke nødvendigvis er overførbart til digitale løsninger fra offentlige myndigheter, og at denne tilliten viser en negativ utvikling fra 2022 til 2023. Generelt er 43 prosent av nordmenn bekymret for hvordan fremtiden påvirkes av kunstig intelligens, som en undersøkelse gjennomført for Nasjonal kommunikasjonsmyndighet (Nkom) i 2023 viser.⁷

Begrepet «kunstig intelligens» kan dekke en rekke ulike systemer i statlig forvaltning. I denne revisjonen brukes den samme definisjonen som er brukt i regjeringens nasjonale strategi for kunstig intelligens (KI-strategien): *Kunstig intelligente systemer utfører handlinger basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål.*⁸ Denne definisjonen av kunstig intelligens kan dermed inneholde enkle algoritmer som ikke faller under KI-definisjonen i for eksempel EUs KI-forordning⁹. Undersøkelsen omfatter slike enkle algoritmer så vidt de er utviklet av eller tilpasset til statlige virksomheter. På lignende måte omfatter undersøkelsen bruk av generativ KI, som for eksempel ChatGPT, kun dersom KI-systemet er videreutviklet eller tilpasset av virksomhetene, eller bygd inn i et eget KI-system.

Potensialet og behovet for å ta i bruk kunstig intelligens i offentlig sektor er tatt opp i en rekke meldinger, blant annet Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet* og Meld. St. 30 (2019–2020) *En innovativ offentlig sektor — Kultur, ledelse og kompetanse*, og i flere strategier¹⁰. I behandlingen av Meld. St. 30 (2019–2020) viste

¹ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonale strategi for kunstig intelligens*. <https://www.regjeringen.no/no/dokumenter/nasjonale-strategi-for-kunstig-intelligens/id2685594/>

² Menon Economics (2023). *KI: Betydning for arbeidsstyrken. En analyse av potensialet for kunstig intelligens-drevet effektivisering i norsk næringsliv*

³ Alan Turing Institute. (2024). *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*.

⁴ McKinsey & Company (2023). *Unlocking the potential of generative AI: Three key questions for government agencies*, [nettartikkel](#)

⁵ Se mer om gevinstene for produktiviteten i Storbritannia i rapporten *Use of Artificial Intelligence in Government* fra den britiske rikskommissjonen, National Audit Office. <https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government/>

⁶ Rambøll Management Consulting AS (2023), *IT i praksis 2023*

⁷ Nasjonal kommunikasjonsmyndighet, nkom.no/aktuelt

⁸ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonale strategi for kunstig intelligens*. <https://www.regjeringen.no/no/dokumenter/nasjonale-strategi-for-kunstig-intelligens/id2685594/>

⁹ European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act), [P9_TA\(2024\)0138](#), fortalen punkt (12)

¹⁰ Kommunal- og moderniseringsdepartementet. (2019). *En digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025*.

kommunal- og forvaltningskomiteen til at kunstig intelligens er et eksempel på en teknologi i rask utvikling, som først kan hente ut sitt fulle potensial gjennom et effektivt 5G-nett. Komiteen er svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, i form av digitaliseringsvennlig regelverk, gode språkressurser, raske og robuste kommunikasjonsnett og tilstrekkelig regnekraft.

I Kommunal- og distriktsdepartementets budsjettproposisjon for 2021 og 2022 (jf. Prop. 1 S 2020–2021 og 2022–2023) vises det til at regjeringen la frem en nasjonal strategi for kunstig intelligens i 2020. I budsjettproposisjonene vises det til at strategien angir retningen for Norges satsing på kunstig intelligens. Strategien inkluderer et sett etiske prinsipper for utvikling og bruk av kunstig intelligens. I KI-strategien vises det videre til at bruk av KI har potensial for betydelige gevinster, blant annet bedre og mer persontilpassede tjenester, større effektivitet og bedre planlegging.¹¹ I strategien står det også tydelig at offentlige virksomheter aktivt bør utforske potensialet i kunstig intelligens, og at grunnlaget skal være «ansvarlig og pålitelig KI». I budsjettproposisjonen til Kommunal- og distriktsdepartementet for 2024 (jf. Prop. 1 S (2023–2024)) vises det til at regjeringen vil se på hvordan man kan legge til rette for at ny teknologi blir brukt best mulig innenfor trygge og ansvarlige rammer.

Generelt presiseres det i KI-strategien at teknologien kan få stor betydning for samfunnsutviklingen og kan gi oss helt nye verktøy for å håndtere samfunnsutfordringer, forbedre offentlige tjenester og styrke verdiskapingen i næringslivet. I Norge er det eksempelvis beregnet at dataøkonomien står for en årlig verdiskaping som tilsvarer 150 milliarder kroner i 2020, og det forventes at dette kan dobles innen 2030.¹²

Våren 2024 uttalte Digitaliserings- og forvaltningsministeren at det er et mål om 80 prosent bruk av kunstig intelligens i offentlig sektor innen 2025. Dette målet skal bidra til å få en bedre og fornyet offentlig sektor, og alle offentlige virksomheter blir oppfordret til å ta i bruk denne teknologien i enda større grad.¹³ Siden kunstig intelligens er under rask utvikling, er studier av produktivetsgevinster oftest basert på et fåtall tidlige brukere, mens langvarig effekt på sektornivå avhenger av utstrekningen av KI-bruken på sikt og av vellykket integrering i organisatoriske prosesser.¹⁴

Digitalisering og arbeid med kunstig intelligens i offentlig sektor følger sektorprinsippet, der hvert departement er ansvarlig for sin sektor. Digitaliserings- og forvaltningsdepartementet har ansvaret for å samordne den offentlige IT-politikken. Samordningsrollen innebærer ansvar for å identifisere sektorovergrepene og for å initiere, koordinere og følge opp tverrgående tiltak i forvaltningen.¹⁵ Departementet skal også være en pådriver¹⁶ for digitalisering i staten, herunder bruk av KI. Departementet ble opprettet 1. januar 2024 og tok over ansvaret for digitaliseringsarbeidet i offentlig sektor fra Kommunal- og distriktsdepartementet. I revisjonen benytter vi i all hovedsak det nye departementsnavnet, også om forhold som tidligere lå under Kommunal- og distriktsdepartementet.

Virksomhetene under Digitaliserings- og forvaltningsdepartementet er også viktige aktører i arbeidet med KI i staten. En av de viktigste aktørene er Digitaliseringsdirektoratet, som skal være regjeringens fremste verktøy for effektiv og samordnet digitalisering av offentlig sektor og av samfunnet øvrig. En annen sentral aktør er Datatilsynet, som utfører blant annet tilsyn og veiledning i skjæringspunktet mellom juss og teknologi. Det skal føre kontroll med at personvernregelverket etterlevs. Datatilsynet er et uavhengig forvaltningsorgan administrativt underordnet Kongen og Digitaliserings- og forvaltningsdepartementet.

¹¹ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*.

¹² Innst. 568 S (2020–2021).

¹³ Muntlig spørretime på Stortinget, 8. mai 2024: [Sak nr. 1 \[10:01:35\] - stortinget.no](#). Hentet 13. juni 2024.

¹⁴ OECD (2024). *The impact of artificial intelligence on productivity, distribution and growth. Key mechanisms, initial evidence and policy challenges*. [OECD artificial intelligence papers No. 15](#)

¹⁵ Prop. 1 S (2023–2024) for Kommunal- og distriktsdepartementet.

¹⁶ Statsministerens kontor. (2023, 16. oktober). *Oppretter nytt departement* [pressemelding]. Regjeringen. <https://www.regjeringen.no/no/aktuelt/opprett-nytt-departement/id3000284/>

1.2 Mål og problemstillinger

Målet med undersøkelsen er å vurdere om statlige virksomheter og heleide selskaper¹⁷ utnytter mulighetene som kunstig intelligens gir, og om de utvikler og bruker KI på en ansvarlig måte i samsvar med Stortingets vedtak og forutsetninger.

Problemstillinger:

1. Bruker statlige virksomheter KI til å effektivisere og forbedre tjenester?
2. Tar statlige virksomheter hensyn til etiske prinsipper i utvikling og bruk av kunstig intelligens?
3. Tilrettelegger Digital- og forvaltningsdepartementets samordning for ansvarlig bruk av KI i statlige virksomheter?

1.3 Avgrensning

Undersøkelsen omfatter KI som enten er utviklet av eller tilpasset til statlige forvaltningsorganer og heleide selskaper, inkludert KI som er utviklet ved bruk av eksterne konsulenter. Undersøkelsen omfatter ikke KI som er innebygd i kommersiell programvare eller ferdige tjenester som virksomhetene kjøper inn (hylleware), med mindre KI-systemet ble spesielt tilpasset til virksomheten eller behandler sensitive personopplysninger. Bruk av for eksempel ChatGPT eller Copilot uten tilpasning i et eget KI-system er dermed ikke del av undersøkelsen.

Undersøkelsen inkluderer ikke bruk av KI som faller inn under sikkerhetsloven, fordi både lovgrunnlaget og risikobildet er annerledes enn for den øvrige forvaltningen.

Undersøkelsen inkluderer tekniske sikkerhetstiltak, som tilgangskontroller, implementert i applikasjoner hvor KI anvendes, og håndtering av endringer i disse applikasjonene. Undersøkelsen omfatter ikke styring av informasjonssikkerhet i virksomheter og organisatoriske og tekniske sikkerhetstiltak som er implementert for å sikre virksomhetens IT-infrastruktur generelt.

Der undersøkelsen berører KI under utvikling, forutsettes det at resultatet av utviklingsprosessen skal kunne brukes i produksjon (det vil si i virksomhetens aktive drift). Utvikling av KI i sammenheng med forskningsprosjekter er derfor ikke en del av undersøkelsen.

1.4 Definisjoner av kunstig intelligens og ansvarlig KI

Definisjonen av kunstig intelligens er angitt ovenfor.

Etter definisjonen regnes et system som kunstig intelligens hvis det inneholder tre aspekter:

- en form av *sansning*, eller registrering og behandling av inndata
- *tolkning* av disse dataene i forbindelse med en gitt oppgave
- en resulterende *handling* som en beslutning eller anbefaling tilpasset oppgaven

Kjernen i et KI-system, som skiller det fra andre automatiseringsprosesser, kan sies å være tolkningen av data, og dette kan implementeres på ulike måter. Tolkningen av data kan være fullstendig regelbasert (eksempelvis en programvarerobot), basert på en kunnskapsbase (ekspertsystem) eller basert på data gjennom statistisk modellering eller maskinlæring:

¹⁷ I det følgende er heleide selskaper inkludert i begrepet statlige virksomheter.

- Et fullstendig regelbasert system består enkelt sagt av et dataprogram basert på klare regler av typen IF $x = a$ THEN Y, IF $x = b$ THEN Z. Et eksempel på et slikt fullstendig regelbasert system er det som ofte blir omtalt som en **programvarerobot eller et RPA-system** (*robotic process automation*). Et slikt system prøver som regel å etterligne menneskelige handlinger, men tolkningen av data er svært begrenset. Eksempler på bruksområder er bokføring, regnskap og fakturahåndtering.
- Tolkning basert på en kunnskapsbase kalles ofte **ekspertsystemer**. Reglene og/eller kunnskapsbasen inneholder heuristisk kunnskap basert på menneskelig ekspertise og erfaring. Resultatene er ofte basert på sannsynligheter om hva som er riktig resultat, selv om de ikke nødvendigvis fremstår slik for sluttbrukerne. Et eksempel på et bruksområde er noen typer chatboter.
- **Statistisk modellering** består av algoritmer som representerer en teori eller antakelser om virkeligheten, og man bruker empiriske data og analyseteknikker, for eksempel logistisk regresjon, for å predikere et utfall. Disse modellene er oftest basert på sannsynlighetsregning. Et eksempel på et bruksområde er enkle analyser av kundedata for å prøve å forutsi hva kundene vil kjøpe neste gang.
- **Maskinlæring (ML)** generaliserer statistisk modellering slik at ingen teori eller antakelse om virkeligheten danner grunnlaget for tolkningen av data. Nye, ukjente data tolkes i all hovedsak ut fra historiske data. Læringsaspektet består altså i at en maskin «lærer» seg reglene for hva som sannsynligvis er riktig resultat, basert på kjennskap til historiske data. En mer avansert form for maskinlæring kalles «dyp læring». De lærte reglene kan være faste, men vanligvis oppdateres de etter en periode eller eventuelt løpende basert på nye data. De har mange bruksområder, for eksempel å identifisere objekter i bilder, finne feil som avvik i store mengder data eller anbefale produkter eller tjenester en kunde mest sannsynlig kan være interessert i.
Generativ KI er basert på en undergruppe av maskinlæring der modellen genererer innhold som tekst eller bilder.

De tre første typene kunstig intelligens blir noen ganger omtalt som «enkel kunstig intelligens» ettersom modellene og systemene er enkle og velkjente og har vært i bruk i mange år. Maskinlæring, herunder dyp læring, kan betraktes som mer «avansert kunstig intelligens» ettersom modellene som brukes, er komplekse og ofte i mindre grad forklarbare, i den forstand at det ikke alltid er mulig å forklare hvorfor man i et bestemt tilfelle fikk et bestemt utfall.¹⁸

Kunstig intelligens er en teknologi i rask utvikling, og siden slutten av 2022 har verktøy basert på generativ kunstig intelligens¹⁹ blitt lett tilgjengelig og brukes av mange. Verktøyene, som for eksempel ChatGPT²⁰ og M365 Copilot²¹ kan blant annet benyttes til å generere tekster, bistå med å lage presentasjoner og å skrive kode. Denne undersøkelsen omfatter bruk av hyllevare som for eksempel ChatGPT *kun* dersom KI-systemet er videreutviklet eller tilpasset av virksomhetene, eller bygd inn i et eget KI-system (jf. avgrensning i avsnitt 1.3).

1.4.1 Definisjon av ansvarlig kunstig intelligens

Den nasjonale KI-strategien følger EUs definisjoner av «tillitvekkende» (*trustworthy*) KI. Det betyr at KI må være lovlige, etiske og sikre. Ifølge EUs ekspertgruppe for kunstig intelligens bygger tillitvekkende KI på fire grunnleggende prinsipper, som danner fundamentet for ansvarlig KI²²:

¹⁸ Manglende forklarbarhet er også kjent som «black box»-problemet.

¹⁹ Generativ KI betegner KI-systemer utviklet for å produsere innhold som tekst, bilder eller kode.

²⁰ ChatGPT er en chatbot av OpenAI basert på en stor språkmodell som kan benyttes som virtuell assistent, se <https://openai.com/index/chatgpt/>

²¹ M365 Copilot er et KI-system fra Microsoft basert på en stor språkmodell og bygd inn i Microsoft 365, som kan brukes som virtuell assistent med tilgang til interne data.

²² Independent High Level Expert Group on Artificial Intelligence set up by the European Commission. (2019). Gruppen har utviklet [Ethics guidelines for trustworthy AI](#) (april 2019), [Policy and investment recommendations for trustworthy AI](#) (juni 2019), [Assessment list for trustworthy AI](#) (juli 2020), og [Sectoral Considerations on Policy and Investment Recommendations for Trustworthy AI](#) (juli 2020).

- respekt for menneskelig autonomi
- forebygging av skade
- likebehandling
- forklarbarhet

Videre foreslår ekspertgruppen syv krav for implementering av disse grunnleggende prinsippene og dermed realisering av tillitvekkende KI, og kravene er tatt inn i den nasjonale strategien for KI.

Ansvarlige KI-systemer må

- respektere menneskets selvbestemmelse og kontroll
- være sikre og teknisk robuste
- ta hensyn til personvernet
- være transparente
- legge til rette for inkludering, mangfold og likebehandling
- være nyttige for samfunn og miljø
- ha mekanismer som sikrer ansvarlighet både før og etter implementering.

Kravene er på et overordnet nivå, og det kan være vanskelig å oppfylle dem i praksis og «oversette» dem til tekniske løsninger. I noen tilfeller kan også prinsippene være i strid med hverandre. Forskning på KI og erfaring med bruk av KI er i rask utvikling. Regulering og standardisering for å sikre at etiske prinsipper ivaretas, er tidkrevende prosesser som naturlig nok følger etter den teknologiske utviklingen. Utviklere av KI-systemer kan derfor være nødt til å samtidig utvikle egne løsninger for utfordrende, tverrfaglige problemstillinger, og det er stor risiko for at virksomheter ikke har god nok kontroll over denne prosessen.

2 Metodisk tilnærming og gjennomføring

For å belyse problemstillingene har vi gjennomført en kartleggingsundersøkelse, en spørreundersøkelse, caseundersøkelser, dokumentanalyser og intervjuer.

Dataene ble samlet inn i perioden februar 2023 til mars 2024. Undersøkellesperioden strekker seg tilbake til 2018. Nedenfor presenterer vi metodene vi har brukt for å undersøke hver enkelt problemstilling.

2.1 Kartlegging av bruk av KI i statlig forvaltning

For å belyse omfanget og effekten av bruk av KI i statlig forvaltning gjennomførte vi kartleggingen i to steg:

- I første steg sendte vi et brev med fem overordnede spørsmål om bruk av KI til statlige virksomheter, heleide selskaper og helseforetakene. Vi bad de som aktivt bruker eller utvikler KI, om å oppgi en eller flere kontaktpersoner som senere kunne svare mer detaljert om KI i en spørreundersøkelse.
- I andre steg fikk virksomheter, heleide selskaper og helseforetak, med aktiv bruk eller utvikling av KI en spørreundersøkelse som går nærmere inn på de innrapporterte KI-systemene.

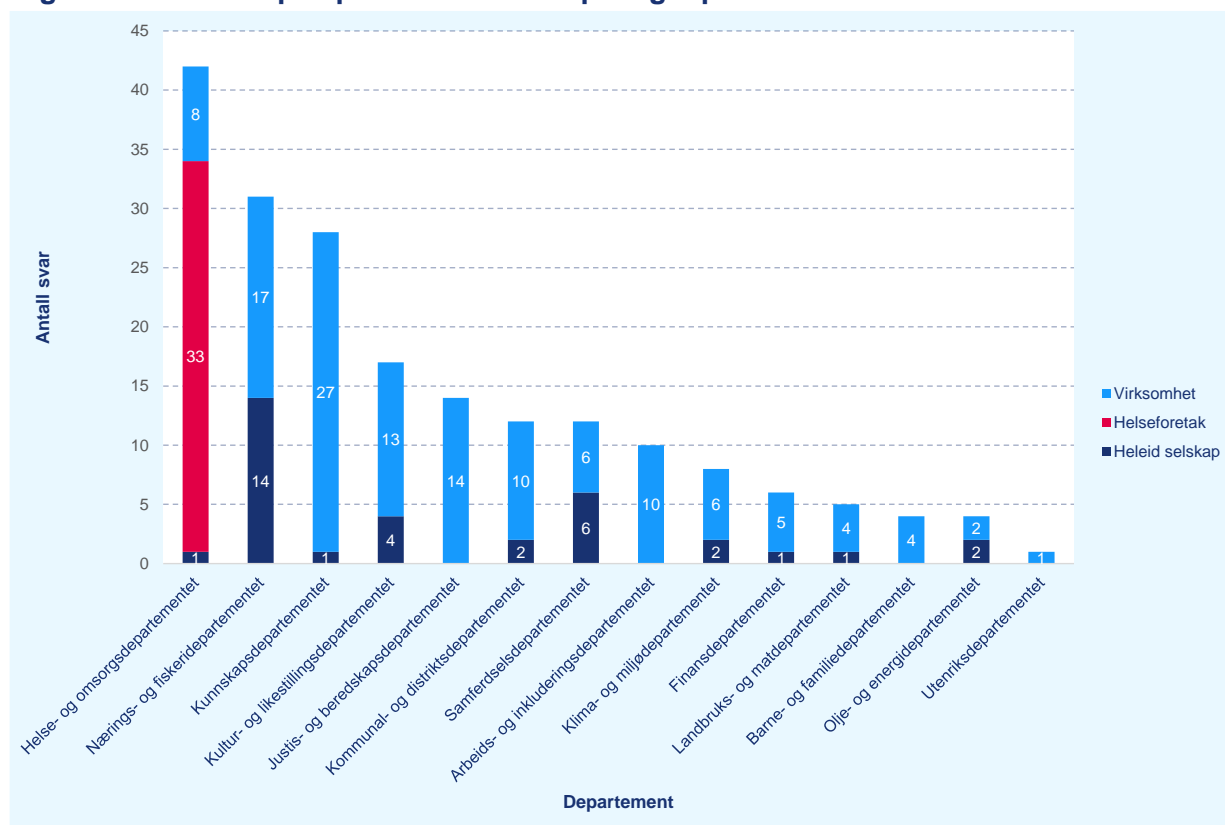
Kartleggingen gir informasjon for å belyse problemstilling 1 om statlige virksomheters bruk av KI til å effektivisere og forbedre tjenester, og deler av problemstilling 2 om virksomhetene som utviklet eller bruker KI, har etablert systemer for å sikre ansvarlig KI. I tillegg var kartleggingen utgangspunktet for utvelgelsen av fire KI-systemer til en caseundersøkelse om ansvarlig KI for problemstilling 2.

2.1.1 Brev for å kartlegge KI-aktivitet i staten

Tidligere kartlegginger av utvikling eller bruk av KI i offentlig sektor har basert seg på et utvalg virksomheter som har deltatt frivillig, og gir derfor ikke et fullstendig bilde. For å få en mest mulig dekkende oversikt over bruken av KI i statlige sektor sendte vi et brev med fem spørsmål til 214 statlige virksomheter, helseforetak og heleide selskaper, fordelt på 14 fagdepartementer. 194 virksomheter (91 prosent) svarte på brevet. Av disse er 127 statlige virksomheter, 34 heleide statlige selskaper og 33 helseforetak.²³

²³ Helsetjenesteforvaltningen (Direktoratet for e-helse, Helsedirektoratet og Statens Helsetilsyn) regnes under virksomheter

Figur 1 Antall svar på spørrebrev fordelt på fagdepartement



Kilde: Riksrevisjonens spørrebrev om kunstig intelligens i staten (2023)²⁴

De følgende sektorene og virksomhetstypene er utelatt fra kartleggingen fordi de faller under sikkerhetsloven eller ikke er relevante for undersøkelsen:

- forsvarssektoren
- alle fagdepartementene (f.eks. Helse- og omsorgsdepartementet, Samferdselsdepartementet etc.)
- statsforvalterne
- Stortinget og Sametinget
- statlige fond (som statens pensjonsfond utland og infrastrukturfondet)

Kartleggingen omfatter ikke bruk av KI som faller inn under sikkerhetsloven, KI-systemer som blir kjøpt inn og brukt uten endringer²⁵, eller forskningsprosjekter²⁶.

Ut fra svarene på brevene kartla vi omfanget av utvikling og bruk av KI-prosjekter (*tatt i bruk, pågående utviklingsprosjekter, tidsavgrensede og avbrutte KI-prosjekter*). Vi spurte også om de vurderte å benytte KI utover det som allerede var planlagt eller implementert. Vi kartla også erfaringer med utfordringer med eller hindringer for å ta i bruk KI. Svarene ga også grunnlag for å velge ut hvilke virksomheter som skulle motta en spørreundersøkelse.

2.1.2 Spørreundersøkelse om konkrete KI-prosjekter og hvordan virksomhetene arbeider med KI

I brevet ba vi om å bli satt i kontakt med personer med faglig ansvar for KI-systemene som har både teknisk forståelse og kunnskap om aktuelle KI-prosjekter. Vi åpnet for flere kontaktpersoner per

²⁴ Figuren er oppdatert etter publisering i vedlegg til Dokument 3:8 (2023–2024). Departementsnavn for KD og NFD var feil i forrige figur, men er nå rettet.

²⁵ KI-systemer som virksomheter kjøper inn (hyllevare), og som verken er utviklet for eller tilpasset til virksomheten, og som er brukt på samme type data som systemet ble utviklet for, og derfor ikke krever validering

²⁶ Undersøkelsen omfatter KI under utvikling i forskningsprosjekter der resultatet av utviklingsprosessen skal kunne brukes i produksjon (f.eks. i utøvende helsetjenester). Delprosjekter forankret i forskning skal inkluderes hvis hovedprosjekt er rettet mot utøvende helsetjeneste, helseforvaltning, saksbehandling, andre kjerneoppgaver, virksomhetens administrasjon eller drift.

virksomhet, slik at alle KI-systemer skulle bli dekket i den påfølgende spørreundersøkelsen, også i store virksomheter med separate fagavdelinger, fakulteter eller lignende.²⁷

Etter vi hadde kvalitetssikret listen over kontaktpersoner og de innrapporterte KI-prosjektene fra brevene, sendte vi ut en spørreundersøkelse til 143 respondenter. Av de 143 respondentene svarte 112 (78 prosent) på spørreundersøkelsen. På bakgrunn av kvalitetssikringen av spørreskjemaene og gjennomgangen av kommentarfelt²⁸ videreførte vi svar fra 104 av respondentene, fordelt på 85 organisasjonsenheter i 66 virksomheter, i analysen av spørreundersøkelsen.

Spørreskjemaet inneholder spørsmål både om måten organisasjonsenheten generelt jobber med KI på, og om de enkelte KI-systemene.

Informasjonen fra spørreundersøkelsen belyser ikke bare om statlige virksomheter bruker KI til å effektivisere og forbedre tjenester (problemstilling 1), men også om virksomhetene har etablert rutiner for å sikre at KI brukes ansvarlig:

- a. Ut fra formålet, hvilke typer algoritmer som brukes, og informasjon om inndata, brukere og så videre har vi kartlagt hvorvidt de ulike aspektene ved ansvarlig KI er relevante for hvert enkelt KI-system.
- b. Vi har kartlagt om virksomhetene har etablert rammeverk eller iverksatt tiltak for å sikre ansvarlig bruk av KI.

Utformingen av spørreundersøkelsen er kvalitetssikret gjennom både intern og ekstern²⁹ pilotering med skriftlige tilbakemeldinger og samtaler. Vi hentet inn innspill om KI i spesialisthelsetjenesten (helseforetakene) ved å intervjuer Helsedirektoratets nasjonale koordineringsprosjekt for bedre bruk av KI i helsesektoren.

Informasjon om organisasjonsenhetsens måte å jobbe med KI på

Når vi har analysert svarene på spørsmålet om hvordan organisasjonsenheten generelt jobber med KI, for eksempel hvilke retningslinjer eller policyer de bruker, har vi slått sammen svarene fra respondenter fra samme virksomhet for å unngå overlapp³⁰. Et unntak er respondenter fra virksomheter som har angitt faglige kontaktpersoner fra ulike divisjoner, fakulteter eller lignende, der det går tydelig frem fra svarene på brevet at det gjelder ulike enheter i samme organisasjon. På denne måten blir svarene om generelt arbeid med KI oppsummert til å representere 85 organisasjonsenheter.

Informasjon om de enkelte KI-systemene

På grunn av visse begrensninger i systemet vi bruker for å sende ut spørreundersøkelser, og for å minimere tidsbruken for respondentene samlet vi inn detaljert informasjon om maksimalt tre KI-systemer per respondent. For å dekke de mest relevante KI-systemene ba vi respondenter som anga flere enn tre KI-systemer totalt, om å velge ut de mest aktuelle systemene for detaljspørsmål etter følgende kriterier:

- det teknologisk mest avanserte prosjektet (f.eks. systemer basert på dyplæring framfør en enkel, regelbasert robot);
- prosjektet som anslagsvis kostet mest å utvikle, enten personalkostnader i forbindelse med egen utvikling eller kostnader til innkjøp av konsulenter eller programvare;

²⁷ I tilfeller hvor flere kontaktpersoner ble angitt uten at disse var tilordnet ulike KI-systemer eller avdelinger, informerte vi kontaktpersonene om dette i forkant av spørreundersøkelsen og ba dem koordinere svarene slik at de dekket alle de ulike systemene.

²⁸ Pga. kommentarer som viser enten at respondenten ikke føler seg kompetente, at spørreundersøkelsen ikke passer til eget bruk av KI, eller at det likevel ikke brukes KI som faller under undersøkelsens definisjon og omfang (scope).

²⁹ Til ekstern validering bidro representanter fra et forskningsinstitutt, to virksomheter og et heleid selskap.

³⁰ Vi vektet svarene rettet på nivå av organisasjonsenheten med 1/(Antall respondenter i virksomhet).

- prosjektet med størst potensielle konsekvenser for enkeltpersoner. Dersom ingen av KI-systemene har konsekvenser for enkeltpersoner, ble de bedt om å velge KI-systemet med størst betydning for virksomheten.

Dette utvalget sikret at vi kunne analysere om respondenten har erfaring med avansert maskinlæring eller med KI-systemer med konsekvenser for enkeltpersoner, også for respondenter som ikke angir detaljer om alle deres KI-systemer. Utvalget er på 189 systemer, mot om lag 300 totalt.

Videre anga alle respondentene et beskrivende navn av KI-systemet, slik at vi kunne identifisere eventuelle overlapp. Vi fant fem KI-systemer som ble dekket av flere enn én respondent. Der svarene for det samme KI-systemet ikke var identiske, tok vi gjennomsnittet av dem. Videre undersøkte vi forskjellen i svarene for å vurdere validiteten på enkelte av svarene.

Analyse av skriftlige retningslinjer eller policyer for risikovurderinger av KI

Respondenter som i spørreundersøkelsen oppga at avansert maskinlæring ble brukt i deres virksomhet,³¹ og dessuten hadde skriftlige retningslinjer eller policyer for risikovurdering for KI-utvikling og/eller bruk, ble bedt om å sende disse retningslinjene/policyene til Riksrevisjonen, samt eventuelle tilhørende maler, sjekklister eller lignende. Respondenter fra 17³² virksomheter har sendt slike retningslinjer, og sammen med de 4 virksomhetene som deltok i caseundersøkelsen (se neste avsnitt), er det innhentet styringsdokumenter for risikovurdering av KI fra til sammen 21 virksomheter. Innsendt dokumentasjon er en variasjon av retningslinjer, strategier, rutinebeskrivelse, maler og prosessbeskrivelser. Ved å analysere disse dokumentene kartla vi hvordan risikoer knyttet til ansvarlig KI ble håndtert.

Videre analyserte vi rutiner for håndtering av personopplysninger ved utvikling og bruk av KI i to av caseundersøkelsene, der virksomhetene brukte persondata til selv å utvikle KI-systemer.

2.2 Caseundersøkelser av utvalgte KI-systemer

For å belyse om utvalgte KI systemer i statlige virksomheter er utviklet og brukt på en ansvarlig måte, som del av problemstilling 2, har vi undersøkt fire KI-systemer i dybden, jf. tabell 1.

Tabell 1 Oversikt over case

Departement	Virksomhet	KI-system brukt for	Leverandør-system	Kjører i sky	Personopplysninger
Finansdepartementet	skatteetaten	utplukk til kontroll av skattefradrag	–	x	x
Arbeids- og inkluderingsdepartementet	Arbeidstilsynet	utvalg av virksomheter til tilsyn	–	x	–
Arbeids- og inkluderingsdepartementet	Statens pensjonskasse	etterkontroll av pensjonsutbetalinger	–	–	x
Helse- og omsorgsdepartementet	St. Olavs hospital HF	klassifisering av celler (diagnose)	x	–	x

³¹ Unntatt rene forskningsprosjekter, eller leverandørsystemer uten tilpasning eller validering.

³² I spørreundersøkelsen anga 56 respondenter (fra 43 organisasjonsheter) at de bruker skriftlige retningslinjer for risikovurdering av KI-systemer, og 25 av disse (23 organisasjonsheter) besvarte undersøkelsen med bakgrunn i minst ett avansert KI-system. I tillegg til virksomheter som var del av caseundersøkelser, ble respondenter fra 19 ulike virksomheter bedt om å tilsende tilsvarende retningslinjer for risikovurdering, 17 svarte.

KI-systemene ble valgt ut på følgende måte:

Basert på svar fra spørreundersøkelsen har vi laget en modell som etter ulik score måler hvor relevant problemstillinger om ansvarlig KI er for hvert KI-system. Scorene er basert på risikoen knyttet til de etiske prinsippene og brukes til sammenligningene av KI-systemene. Høyere score gjenspeiler høyere risiko, som kan skyldes for eksempel bruk av persondata, kompleksitet i algoritmen, automatiserte beslutninger, bruk av eksterne konsulenter (som indikasjon på lav intern kompetanse) eller at KI-systemets resultater har konsekvenser for personer.

I utvalget av KI-systemer til caseundersøkelsen la vi mest vekt på systemenes relevans for samfunnet og at systemene allerede er i bruk. I tillegg vektla vi variasjon i utvalget, slik at KI-systemene dekker ulike sektorer, med minst ett system fra helsesektoren (da vår kartlegging viser at det er størst KI-aktivitet innen helsesektoren), et leverandørsystem og systemer med henholdsvis høy og lav risiko basert på scorene fra den ovennevnte modellen.

Etter oppfølgingsmøter om KI-systemene med åtte kandidater til caseundersøkelsen, falt fire systemer bort. Noen var uaktuelle fordi den underliggende modellen er så enkel at det er på grensen til å kunne kalles KI, og andre fordi KI-systemet bruker et leverandørsystem der det ikke er behov for noen tilpasning til virksomheten.

De fire KI-systemene i caseundersøkelsen ble undersøkt gjennom dokumentanalyser, samtaler med KI-prosjektet og skriftlige oppfølgingsspørsmål. Dokumentene vi hentet inn, var blant annet virksomhetsintern dokumentasjon i ulike formater, rapporter, uttrekk av kodenært dokumentasjon, møtepresentasjoner, prosjektbeskrivelse, rammeverk og styringsdokumenter. Dokumentasjonen beskriver planlegging, utvikling, implementering, overvåking og vedlikehold av KI-systemer samt eventuelle policyer, rammeverk eller maler for prosesser knyttet til ansvarlig KI (for eksempel en mal for vurdering om det burde gjennomføres en vurdering av personvernkonsekvenser). Dokumentasjonen for leverandørsystemet i sykehuset omfatter anskaffelses- og valideringsprosessene istedenfor utvikling av KI-systemet. Dokumentasjonen som gjelder IT-sikkerhet, beskriver tilgangsstyring, logging, backup og endringshåndtering i IT-systemer brukt for KI-systemet. Der det manglet dokumentasjon, avklarte vi forholdene gjennom samtaler med og skriftlige oppfølgingsspørsmål til prosjektteamet.

For å vurdere om prinsippene for ansvarlig KI er ivaretatt i KI-systemene i caseundersøkelsen, laget vi et eget rammeverk basert på to eksisterende rammeverk:

1. [Auditing machine learning algorithms](#) er et rammeverk for revisjon av KI-systemer basert på maskinlæring, utviklet av riksrevisjonene i Finland, Tyskland, Nederland, Storbritannia og Norge i 2020 (oppdatert 2023), med tilhørende [Excel-verktøy](#).
2. [Audit framework for algorithms](#) er et rammeverk utviklet av den nederlandske riksrevisjonen (Netherlands Court of Audit) i forbindelse med revisjonen [Understanding algorithms](#) i 2021, og brukt i revisjonen [An audit of 9 algorithms used by the Dutch government](#) i 2022³³. Dette rammeverket legger ikke bare maskinlæring til grunn, men også enklere KI-systemer.

Vi har tilpasset rammeverket vårt til spørsmål som er mest relevante for KI-systemene i dybdeundersøkelsen. Vi brukte spørsmålene i de eksisterende rammeverkene som kunne bli besvart gjennom dokumentanalyser og intervju. Vi har ikke tatt med spørsmål i rammeverkene som forutsetter gjennomgang av for eksempel programkoden. Dermed kan vi trekke konklusjoner om det er etablert rutiner og en dokumentert prosess for å ivareta prinsippene om ansvarlig KI, men ikke om et KI-system i praksis bryter med noen av prinsippene.

³³ Vi har jobbet med den nyeste versjonen av rammeverket, som ikke enda er publisert på engelsk.

Rammeverket er nærmere beskrevet i vedlegg 3. Revisjonsspørsmålene i rammeverket ble kategorisert etter prinsippene om *styring og ansvar*, *personvern*, *likebehandling*, *transparens*, og *robusthet og sikkerhet*³⁴. Vi vurderer KI-systemets modenhetsnivå, eller KI-prosjektets aktivitetsnivå, innenfor de ulike etiske prinsippene ut fra hvor mange krav i de underliggende revisjonsspørsmålene som er oppfylt.

Oppsummerende rapporter av hver caseundersøkelse har blitt verifisert av de respektive prosjektteamene i virksomhetene.

2.3 Intervju

Vi har gjennomført intervjuer for å undersøke og belyse alle problemstillingene. Intervjuene har særlig blitt brukt til å utdype og få synspunkter på problemstilling 1 og 3. Vi har intervjuet de følgende departementene, direktoratene, virksomhetene og ombudene:

- Akershus universitetssykehus HF (alle problemstillinger)
- Arbeids- og inkluderingsdepartementet (problemstilling 1 og 3)
- Arbeids- og velferdsdirektoratet (problemstilling 1 og 3)
- Arbeidstilsynet (caseundersøkelse)
- Datatilsynet (alle problemstillinger)
- Digitaliserings- og forvaltningsdepartementet (alle problemstillinger)
- Digitaliseringsdirektoratet (alle problemstillinger)
- Finansdepartementet (problemstilling 1 og 3)
- Helsedirektoratet (problemstilling 1 og 2)
- Justis- og beredskapsdepartementet (problemstilling 1 og 3)
- Likestillings- og diskrimineringsombudet (problemstilling 2)
- skatteetaten (alle problemstillinger og caseundersøkelse)
- Statens pensjonskasse (alle problemstillinger og caseundersøkelse)
- St. Olavs hospital HF (caseundersøkelse)
- tolletaten (problemstilling 1 og 3)

Intervjuene med virksomhetene har i all hovedsak vært avholdt med avdelinger og enheter med ansvar for digitalisering, kunstig intelligens og juridiske spørsmål. Referatene fra alle intervjuene som er gjennomført i undersøkelsen, er verifisert. Det er sendt brev med spørsmål til:

- Helse- og omsorgsdepartementet (problemstilling 1 og 3)
- Kultur- og likestillingsdepartementet (problemstilling 1 og 3)

2.4 Dokumentgjennomgang

Dokumentene har blitt analysert i forbindelse med alle problemstillingene. Dokumentene har vært brukt både som utgangspunkt for intervju spørsmål og til å belyse informasjon innhentet gjennom brev med spørsmål, spørreundersøkelsen og caseundersøkelsen.

³⁴ Jf. vedlegg 3.

Noen eksempler på dokumenttyper, jf. referanselisten:

- tildelingsbrev og årsrapporter fra statlige virksomheter under Digitaliserings- og forvaltningsdepartementet, Finansdepartementet og Arbeids- og inkluderingsdepartementet
- forskningsrapporter
- rapporter fra andre riksrevisjoner
- rapporter fra Datatilsynets KI-sandkasse
- rapport fra den hurtigarbeidende arbeidsgruppen for å vurdere behov for nasjonal regulering utover KI-forordningen (november 2023).

I tillegg har vi mottatt en Excel-oversikt fra Digitaliserings- og forvaltningsdepartementet som viser status i gjennomføringen av de 51 tiltakene i den nasjonale strategien for kunstig intelligens.

3 Revisjonskriterier

3.1 Overordnede mål og ansvar

Digitaliserings- og forvaltningsdepartementet har ansvaret for å koordinere regjeringens IT-politikk, jf. Prop 1 S (2023–2024). I Prop. 1 S (2023–2024) for Kommunal- og distriktsdepartementet går det frem at formålet med IT-politikken er å utvikle rammebetingelser som støtter digitaliseringen i de ulike sektorene, på tvers av sektorene og i samfunnet som helhet. Samordningsansvaret, som er lagt til Digitaliserings- og forvaltningsdepartementet, innebærer blant annet å identifisere sektorovergrepene utfordringer og å initiere, koordinere og følge opp tverrgående tiltak. Departementet har et særskilt ansvar for å arbeide for en styrket og mer helhetlig tilnærming til informasjonssikkerheten i forvaltningen. Videre presiserer proposisjonen at digitaliseringen av samfunnet skal skje på en bærekraftig og inkluderende måte og bidra til forenkling og effektivisering i både offentlig og privat sektor, men også hindre en utvikling som sentraliserer oppgaver og arbeidsprosesser, svekker personvern eller forbrukervern, gjør samfunnet mer sårbart for digitale angrep, svekker konkurransen i de digitale markedene eller forsterker digitalt utenforskap.

Behovet for å tilrettelegge for og regulere bruken av kunstig intelligens i offentlig sektor er tatt opp i flere meldinger, budsjettproposisjoner og innstillinger. I Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet* vises det til at man må anta at nye teknologier for automatisering og kunstig intelligens kan gi endringer i hvordan tjenesteproduksjonen i offentlig sektor vil foregå i fremtiden, og at dette er en utvikling som også offentlig sektor må evne å dra nytte av. I behandlingen av meldingen (jf. Innst. 84 S (2016–2017)) viser transport- og kommunikasjonskomiteen generelt til at de teknologiske endringene først og fremst åpner for store fremskritt, men at det kreves en aktiv politikk for å høste og fordele gevinstene og begrense ulempene. Et flertall i komiteen viser til at «[a]utomatisering og kunstig intelligens har et enormt potensial for å øke produktiviteten i samfunnet, særlig i offentlig sektor».³⁵

I Meld. St. 30 (2019–2020) *En innovativ offentlig sektor — Kultur, ledelse og kompetanse* og i flere strategier³⁶ vises det til at regjeringen vil arbeide for én digital offentlig sektor på tvers av forvaltningsnivåene og utnytte mulighetene kunstig intelligens gir for å løse oppgavene på nye måter. Regjeringen mener skrittvis innovasjon ikke alltid er tilstrekkelig for å utnytte mulighetene som blant annet ligger i ny teknologi som kunstig intelligens og datadeling. I behandlingen av meldingen (jf. Innst. 191 S (2020–2021)) viser flertallet³⁷ i kommunal- og forvaltningskomiteen til at regjeringen vil legge til rette for økt bruk av kunstig intelligens i offentlig sektor. Det samme flertallet viser til at tillit til systemet og de som forvalter, drifter og utvikler det, vil være avgjørende for at også fremtidens velferdsmodell aksepteres av innbyggerne.

I behandlingen av Meld. St. 30 (2019–2020) viste kommunal- og forvaltningskomiteen til at kunstig intelligens er et eksempel på en teknologi i rask utvikling, som først kan hente ut sitt fulle potensial gjennom et effektivt 5G-nett. Komiteen er svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, i form av digitaliseringsvennlig regelverk, gode språkressurser, raske og robuste kommunikasjonsnett og tilstrekkelig regnekraft. Kommunal- og forvaltningskomiteen viste videre til at det er særlig viktig at myndighetene har nødvendig kontroll på og regulerer bruken av kunstig intelligens og «big data», slik at personvern ivaretas, og at utviklingen ikke fører til enda skjevare maktforhold og økende økonomiske forskjeller.³⁸

³⁵ Arbeiderpartiet, Kristelig Folkeparti, Senterpartiet og Venstre.

³⁶ Kommunal- og moderniseringsdepartementet. (2019). *En digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025*, Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*.

³⁷ Arbeiderpartiet, Høyre, Senterpartiet, Sosialistisk Venstreparti og Kristelig Folkeparti.

³⁸ Innst. 191 S (2020–2021).

I Kommunal- og distriktsdepartementets budsjettproposisjoner for perioden 2021–2023 (jf. Prop. 1 S 2020–2021 og 2022–2023) vises det til den nasjonale strategien for kunstig intelligens fra 2020, og at strategien angir retningen for Norges satsing på kunstig intelligens. Ett viktig tiltak er tilgang på data av god kvalitet, ifølge Prop. 1 S (2021–2022). God datakvalitet henger sammen med målene om å utnytte de store mengdene data som finnes, til det beste for samfunnet (jf. f.eks. Meld. St. 22 (2020–2021) *Data som ressurs – Datadrevet økonomi og innovasjon*). I Prop. 1 S (2021–2022) vises det også til at strategien inkluderer et sett etiske prinsipper for utvikling og bruk av kunstig intelligens. I behandlingen av Prop. 1 S (2020–2021) for Kommunal og distriktsdepartementet (jf. Innst. 16 S (2020–2021) viser kommunal og forvaltningskomiteen til at både de store personvernutfordringene den økte digitaliseringen gir, og til at det er behov for å gjøre det enklere for brukerne å forstå hvordan personopplysninger brukes. Derfor er det stort behov for sterkt myndighetsfokus på dette, ifølge komiteen.

Meld. St. 22 (2020–2021) *Data som ressurs — Datadrevet økonomi og innovasjon*, jf. Innst. 568 S (2020–2021), viser til at arbeidet for å forhindre diskriminering, manipulasjon og misbruk av informasjon blir særlig viktig, og at åpenhet, likebehandling og rettssikkerhet er viktige demokratiske verdier som også må gjelde i dataøkonomien. Meldingen definerer *dataøkonomi* som verdiskapingen som skjer når data er en viktig innsatsfaktor i produksjon av varer og tjenester, eller når data er en driver for innovative løsninger.

I budsjettproposisjonen til Kommunal- og distriktsdepartementet for 2024 (jf. Prop. 1 S (2023–2024)) vises det til at regjeringen vil se på hvordan man kan legge til rette for at ny teknologi som kunstig intelligens blir brukt best mulig innenfor trygge og ansvarlige rammer.

De påleggene og anbefalingene som gjelder for digitalisering i offentlig sektor på overordnet nivå, blir årlig sammenstilt i Digitaliseringsrundskrivet. Dette er et rundskriv som Digitaliserings- og forvaltningsdepartementet sender til de øvrige departementene og de statlige forvaltningsorganene hvert år. I Digitaliseringsrundskrivet H5-21 av 11. januar 2022 går det frem at digitalisering handler om å bruke teknologi til å tilby bedre tjenester og å tilby digitale tjenester som er enkle å bruke, effektive og pålitelige, og at virksomheter som deler ansvarsområder bør samarbeide for å gi brukerne et hensiktsmessig og helhetlig tilbud, som er uavhengig av måten forvaltningen er organisert på.³⁹

Digitaliserings- og forvaltningsdepartementet forvalter IT-politikken i all hovedsak gjennom Digitaliseringsdirektoratet og Datatilsynet. Digitaliseringsdirektoratet skal være regjeringens fremste verktøy for å gjøre digitaliseringen av samfunnet raskere og mer samordnet. Direktoratet har en rolle som både premissgiver, iverksetter, fellestjenesteleverandør og tilsynsmyndighet. Noen av direktoratets oppgaver som kan relevante for bruken av kunstig intelligens hos statlige virksomheter, er det overordnede, nasjonale ansvaret for informasjonsforvaltning i offentlig sektor og deling av data og oppgaven med å samordne og være pådriver i offentlig sektors arbeid med forebyggende informasjonssikkerhet. Digitaliseringsdirektoratet skal særlig bidra til at alle statlige virksomheter har et system for internkontroll av informasjonssikkerhet.⁴⁰ Direktoratet har, med utgangspunkt i utredningsinstruksen, laget en prosjektveiviser. Dette er den anbefalte prosjektmodellen for alle digitaliseringstiltak i offentlig sektor⁴¹. Veiviseren tar utgangspunkt i prosjektstyringsfasene i en utviklingsmodell og gir råd om å etablere prosjektet, beskrive og forankre målene og hvordan prosjektet skal styres. Digitaliseringsrundskrivet presiserer at Prosjektveiviseren bør brukes dersom virksomheten ikke har etablert en prosjektmodell for IT-relaterte tiltak for å styre digitaliseringsprosjekter i tråd med god forvaltningsskikk.

³⁹ Kommunal- og distriktsdepartementet. (2022). *Digitaliseringsrundskrivet* [Rundskriv], H-5/21. Hentet 30. november 2023 fra <https://www.regjeringen.no/no/dokumenter/digitaliseringsrundskrivet/id2895185/>

⁴⁰ Kommunal- og moderniseringsdepartementet. (2021). *Økonomi- og virksomhetsinstruks for Digitaliseringsdirektoratet*.

⁴¹ Prosjektveiviseren.no (DigDir) > Hvorfor ble Prosjektveiviseren laget? Hentet 23. november 2023 fra [Prosjektveiviseren \(digdir.no\)](https://prosjektveiviseren.no/digdir)

Datatilsynet fører tilsyn på personvernområdet og er også ombud i personvernsaker. Mer spesifikt på området kunstig intelligens finnes det et testmiljø som skal sørge for etisk og forsvarlig bruk av kunstig intelligens, *Datatilsynets regulatoriske sandkasse for kunstig intelligens*.⁴² Formålet med den regulatoriske sandkassen er å bidra til at virksomheter som tar i bruk kunstig intelligens, følger regelverket og ivaretar personvernet på en god måte. Virksomheter kan selv søke om å få delta i den regulatoriske sandkassen.

Regelverket relevant for digitalisering og ansvarlig kunstig intelligens blir delvis forvaltet av andre fagdepartementer. Justis- og beredskapsdepartementet har ansvaret for forvaltningsloven og personvernloven, og Kultur- og likestillingsdepartementet for likestillings- og diskrimineringsloven og arkivloven.

3.2 Krav til ansvarlig kunstig intelligens

Per i dag finnes det ingen lovhjemmel som direkte dekker kunstig intelligens eller algoritmebruk i forvaltningen, og forvaltningsloven stiller ingen krav til dokumentasjon av slik bruk. Det er imidlertid mulig å utlede krav fra grunnleggende etiske prinsipper i menneskerettsloven og Grunnloven, fra alminnelige forvaltningsrettslige lover og prinsipper og fra generelle krav om en effektiv og god forvaltning.

Digitaliserings- og forvaltningsdepartementet, Datatilsynet og Digitaliseringsdirektoratet har gitt krav, anbefalinger og retningslinjer som gjelder der statlige virksomheter møter problemstillinger som gjelder spesielt for digitaliseringsprosjekter, digital saksbehandling og for eksempel der kunstig intelligens blir brukt som beslutningsstøtte. De viser også til relevante normer, standarder og prinsipper for bruk av kunstig intelligens i offentlige virksomheter.

EU-kommisjonen har nedsatt en ekspertgruppe som har utarbeidet etiske retningslinjer for bruk av kunstig intelligens. Retningslinjene, som er basert på EUs pakt om grunnleggende rettigheter og internasjonal menneskerettslovgivning, er ment å legge til rette for den graden av autonomi som kjennetegner kunstig intelligens, slik at handlinger og beslutninger kan iverksettes uten at mennesker er direkte involvert⁴³. De etiske prinsippene for kunstig intelligens tar utgangspunkt i at alle beslutninger og avveiiinger som er gjort, skal være begrunnet og dokumentert, og at det bør finnes mekanismer som rapporterer urimelige negative virkninger.

EUs pakt om grunnleggende rettigheter fastsetter fire grunnleggende etiske prinsipper: respekt for menneskelig autonomi, forebygging av skade, likebehandling og forklarbarhet⁴⁴. Selv om pakten i seg selv ikke har status som rettskilde i norsk rett, har det betydning for både den europeiske menneskerettskonvensjonen og avtalen om det europeiske økonomiske samarbeidsområdet (EØS) som igjen har betydning for norsk rett, og påvirker tolkning av norske rettskilder, inklusive Grunnloven^{45,46}. De etiske prinsippene fører til følgende krav, som er forankret i norsk lov:

1. personvern, inklusive respekt for menneskets selvbestemmelse og kontroll
2. likebehandling og ikke-diskriminering
3. teknisk sikkerhet og robusthet
4. transparens (forklarbarhet og åpenhet)

⁴² Datatilsynet. (u.å.). *Sandkassesiden*. Hentet 12. februar 2024 fra <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/>

⁴³ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*, kap 5.2.

⁴⁴ Charter of Fundamental Rights of the European Union, EUR-Lex, Document 12012P. Hentet 19. april 2024 fra <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>

⁴⁵ Dokument 16 (2011–2021) Rapport til Stortingets presidentskap fra Menneskerettighetsutvalget om menneskerettigheter i Grunnloven, avgitt 19. desember 2011

⁴⁶ Elgesem, F. (2021). Tolkning av den europeiske menneskerettighetskonvensjonen, *Jussens Venner*, 55(6), s. 441–487. Universitetsforlaget.

Disse aspektene, og de fundamentale rettigheter de bygger på, danner grunnlaget for å vurdere om kunstig intelligens brukes ansvarlig i forvaltningen, og vil sammen med relevant særlovgivning danne utgangspunktet for vurderingene for de enkelte virksomhetene og departementene.⁴⁷

Hvor viktige de enkelte aspektene er, varierer med bruksområde. I bruksområder med kategorisering av personer for eksempel for kontrollformål, er rettssikkerhet og likebehandling viktige faktorer. Der kunstig intelligens brukes i forbindelse med automatiserte vedtak som ikke bruker personopplysninger, er det en relevant risiko for feilaktig saksbehandling, for eksempel på grunn av manglende hensyn til skjønnsvurderinger⁴⁸. I tillegg blir det viktig å stille tekniske krav til sikkerhet og robusthet samt å sikre åpenhet og forklaringsevne for å sikre at saksbehandlingen er etterprøvable, og for å sikre klageretten.

3.2.1 Krav til personvern og respekt for menneskets selvbestemmelse og kontroll

Respekt for menneskelig autonomi har en nær sammenheng med menneskets rett til verdighet og frihet i EUs pakt om grunnleggende rettigheter artikkel 1 og 6. Av menneskets iboende verdi følger det at mennesker ikke skal bli redusert til objekter som blir sortert og vurdert av teknologi på en automatisert måte. Mennesker skal ha oppsyn med KI-systemer som interagerer med eller har direkte virkning på mennesker, med et grensesnitt utformet slik at mennesker kan ta meningsfulle valg og KI-systemer er til støtte for menneskelig autonomi.

Kunstig intelligens som er rettet mot personer eller bygger på personopplysninger, skal ivareta personvernet. Grunnloven slår fast at «[e]nhver har rett til respekt for sitt privatliv og familieliv, sitt hjem og sin kommunikasjon», og at «[s]tatens myndigheter skal sikre et vern om den personlige integritet».⁴⁹ Mer generelt presiserer Grunnloven at «statens myndigheter skal respektere og sikre menneskerettighetene slik de er nedfelt i denne grunnlov og i for Norge bindende traktater om menneskerettigheter».⁵⁰ Barn har et enda sterkere integritetsvern.⁵¹ Den europeiske menneskerettskonvensjonen av 1950, som ble tatt inn i norsk lov ved menneskerettsloven av 1999,⁵² pålegger staten å sørge for at begrensninger i retten til personvern er forankret i lov og er nødvendig i et demokratisk samfunn for å oppnå andre mål og rettigheter.⁵³ Den europeiske menneskerettsdomstolen har over tid lagt til grunn at dette vernet inkluderer vern av personopplysninger, også for offentlig sektor.⁵⁴

Personvernforordningen⁵⁵ inneholder bestemmelser om behandling av personopplysninger på de fleste samfunnsområder. Personopplysningsloven⁵⁶ slår fast at personvernforordningen gjelder som norsk lov og ved konflikt går foran annen norsk lovgivning. Den inneholder også bestemmelser som bare gjelder i Norge, men særregler finnes også i annen norsk lovgivning. Personvernforordningen gjelder for behandlingen av personopplysninger, altså opplysninger som *kan* knyttes til en fysisk person. Det er altså tilstrekkelig at det er mulig å knytte en opplysning til en person som kan identifiseres, og opplysningen trenger heller ikke å bli oppfattet som sensitiv.

Virksomhetene har en plikt til å kunne påvise at behandlingen av personopplysninger skjer i samsvar med personvernforordningen.⁵⁷ Virksomheter som har mer enn 250 ansatte, eller som behandler personopplysninger på en måte som sannsynligvis vil medføre en risiko for de registrertes rettigheter og friheter, skal dokumentere behandlingen av personopplysninger i såkalte behandlingsprotokoller.

⁴⁷ Jf. personvernforordningen, diskrimineringsloven, forvaltningsloven, åpenhetsloven, sikkerhetsloven, mfl.

⁴⁸ Sivilombudsmannen. (2020). *Årsmelding for 2019. Dokument 4 (2019–2020)*.

⁴⁹ Grunnloven § 102.

⁵⁰ *Ibid* § 92.

⁵¹ *Ibid* § 104 tredje ledd.

⁵² Menneskerettsloven av 1999 § 2 nr. 1.

⁵³ Den europeiske menneskerettskonvensjonen artikkel 8.

⁵⁴ Se *Niemietz v. Tyskland*, [J], no. 13710/88, (1992) 16 EHRR 97 og *von Hannover v. Tyskland*, [J] no. 59320/00 ECHR 2004-VI ECHR 294.

⁵⁵ Forordning (EU) (2016/679) 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).

⁵⁶ Lov 15. juni 2018 nr. 38 om behandling av personopplysninger (personopplysningsloven). Resten av avsnittet er basert på samme kilde.

⁵⁷ Personvernforordningen

Generelle retningslinjer for virksomhetene

Personvernforordningen inneholder seks generelle personvernprinsipper for behandling av personopplysninger:⁵⁸

- *Prinsippet om lovlighet, likebehandling og åpenhet* innebærer at behandlingen må være i samsvar med personvernregelverk og menneskerettigheter. Virksomheten må i tillegg avveie personvern mot andre grunnleggende rettigheter og friheter på en forholdsmessig, lyttende og ikke-forutinntatt måte. Åpenhet er nødvendig for at den registrerte skal kunne bruke rettighetene sine, og at kontrollorganer kan vurdere om regelverket etterleves.
- *Formålsbegrensningsprinsippet* krever at virksomheten definerer et formål med behandlingen før den kan samle inn personopplysninger. Virksomheten kan i utgangspunktet ikke bruke personopplysninger til andre formål. Forordningen har også en egen bestemmelse for dette.
- *Dataminimeringsprinsippet* innebærer at virksomheten ikke skal behandle flere personopplysninger enn det som er nødvendig for å oppnå formålet.
- *Riktighetsprinsippet* stiller krav til virksomheten om at personopplysninger skal være så riktige og fullstendige som formålet tilsier.
- *Lagringsbegrensningsprinsippet* innebærer at virksomheten ikke skal lagre personopplysningene lenger enn det som er nødvendig for å oppnå formålet.
- *Prinsippet om integritet og konfidensialitet* stiller krav til virksomheten om informasjonssikkerhet, integritet, tilgjengelighet og konfidensialitet i behandling av personopplysninger. I tillegg til det overordnede prinsippet, gir forordningen konkrete regler om sikkerhet, som vi beskriver senere.

Begrepet behandling av personopplysninger omfatter alle former for bruk, men i hovedsak de som er helt eller delvis automatisert, og reglene er strengere for helautomatisert bruk. Automatisk behandling av personopplysninger, slik som ved bruk av kunstig intelligens, krever at virksomheten utformer systemet slik at den etterlever bestemmelsene i forordningen. Retten og plikten til skjønnsvurderinger er et grunnleggende prinsipp i norsk forvaltning. Der avgjørelser basert på automatisert behandling av personopplysninger er tillatt i en særlov, kan avgjørelser ikke bygge på skjønnsmessige vilkår i en lov eller forskrift, med mindre avgjørelsen er utvilsom, og den registrerte personen har rett til manuell overprøving.⁵⁹ Det er den øverste ledelsen i virksomheten som er ansvarlig for at virksomheten følger både personvernprinsippene og de enkelte bestemmelsene.

Personer som virksomheten har opplysninger om, har rettigheter overfor virksomheten. Det innebærer for eksempel at virksomheten skal gi den registrerte opplysninger om hvilke personopplysninger de har, og hvordan de behandler dem. Virksomheten skal ha systemer som gjør det lettere for de som er registrert, å ivareta sine rettigheter.

Krav til virksomhetene om vurdering av risiko og til innebygd personvern

Personvernforordningen stiller krav om at virksomheten vurderer risikoen når den skal behandle personopplysningene, og at den vurderer tiltak som kan redusere risikoen til et *akseptabelt* nivå.⁶⁰ Dersom det er «sannsynlig» at behandlingen vil medføre «høy risiko for å fysiske personers rettigheter og friheter», skal virksomheten gjennomføre en vurdering av personvernkonsekvenser (DPIA) *før* den samler inn og tar i bruk personopplysninger.⁶¹ Virksomheten må sette inn flere eller større tiltak dersom krenkelsene kan være store eller sannsynlige. Den må vurdere alle risikoreducerende tiltak som kan være virksomme eller hensiktsmessige.

Personvernforordningen gir virksomhetene plikt til å bygge inn personvern i tekniske og organisatoriske løsninger, omtalt som innebygd personvern og personvern som standardinnstilling,

⁵⁸ Personvernforordningen artikkel 5 nr. 1. Referansen gjelder frem til neste overskrift.

⁵⁹ F.eks. lov om Statens pensjonskasse, §45 b, lov om arbeids- og velferdsforvaltning (NAV), § 4 a, lov om skatteforvaltning, § 5-11.

⁶⁰ Personvernforordningen artikkel 35 nr. 1.

⁶¹ Ibid.

men rekkevidden av plikten avhenger av en avveing av ulike forhold.⁶² Plikten springer ut av resultatet av risikovurderingen, kostnadene, de tekniske mulighetene, omfanget av personopplysninger, formålet og typen personopplysning.

Personvernforordningen stiller blant annet krav om at «[t]ilsynsmyndigheten skal utarbeide og offentliggjøre en liste over hvilke typer behandlingsaktiviteter som omfattes av kravet om vurdering av personvernkonsekvenser» (artikkel 35, punkt 4). Datatilsynet har utarbeidet en veileder for vurdering av personvernkonsekvenser. Veilederen oppgir en rekke aktiviteter som **alltid** krever en slik vurdering, eller som kan bidra til høy risiko for de registrertes rettigheter og friheter. Denne listen er godkjent av Det europeiske personvernrådet. Veilederen skal hjelpe virksomheter med å vurdere om virksomheten skal gjennomføre en analyse av personvernkonsekvenser, og *hvordan* de da skal gå frem.⁶³ Veilederen er utformet med utgangspunkt i anbefalingene fra EUs rådgivende organ i personvern, Artikkel 29-gruppen⁶⁴, men Datatilsynet har tilpasset anbefalingene til norske forhold. Datatilsynet har også utarbeidet en veileder for innebygd personvern og personvern som standard.⁶⁵ Den skal bidra til at systemene oppfyller personvernprinsippene. Veilederen er basert på Personvernrådets (EDBP) retningslinjer⁶⁶. Vi tar utgangspunkt i at Datatilsynets veiledere er godkjente standarder som skal operasjonalisere kravene i personvernforordningen, deriblant for når en vurdering av personvernkonsekvenser bør gjennomføres, forståelse av *innebygd personvern* og relevante personvernprinsipper.

3.2.2 Krav om likebehandling

Grunnleggende rettigheter som likhet for loven, ikke-diskriminering og respekt for mangfold og hensyn til sårbare grupper⁶⁷ betyr at KI-systemer må være *rettferdige* og sikre at like saker behandles likt og fører til lignende utfall⁶⁸. Videre inkluderer likebehandling muligheten å klage på beslutninger truffet av eller med hjelp av KI, og derfor at ansvaret for følgene av KI-systemet er avklart.

Virksomhetenes bruk av kunstig intelligens skal ivareta retten til å ikke bli diskriminert og den ulovfestede forvaltningsrettslige rettsregelen som innebærer at vedtak ikke skal innebære usaklig forskjellsbehandling (likhetsprinsippet)⁶⁹.

Virksomhetene kan ikke diskriminere på bakgrunn av kjønn, graviditet, permisjon ved fødsel eller adopsjon, omsorgsoppgaver, etnisitet, religion, livssyn, funksjonsnedsettelse, seksuell orientering, kjønnsidentitet, kjønnsuttrykk, alder eller kombinasjoner av disse.⁷⁰ Det er heller ikke tillatt å diskriminere på grunn av tilknytning til en annen person. Virksomhetene kan ikke behandle en person dårligere enn andre blir, har blitt eller ville blitt behandlet i en tilsvarende situasjon.⁷¹ Virksomhetene har heller ikke lov til å forskjellsbehandle gjennom et tilsynelatende nøytralt utformet system.⁷² Virksomhetene kan forskjellsbehandle individer dersom det har et saklig formål, er nødvendig for å oppnå formålet og ikke er uforholdsmessig inngripende overfor den eller de som forskjellsbehandles.⁷³

⁶² Personvernforordningen artikkel 25.

⁶³ Datatilsynet. (2023, 27. juli). Vurdering av personvernkonsekvenser (DPIA). Hentet fra <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/vurdering-av-personvernkonsekvenser/>

⁶⁴ Artikkel 29-gruppen ble erstattet av Personvernrådet (EDPB, European Data Protection Board) ved ikrafttredelsen av EUs personvernforordning (GDPR) 25. mai 2018. Retningslinjene til Artikkel 29-gruppen finner du her: <https://ec.europa.eu/newsroom/article29/items/611236>.

⁶⁵ Datatilsynet. (2022). *Innebygd personvern og personvern som standard*. Hentet 23. november 2023 fra <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/innebygd-personvern-og-personvern-som-standard/>.

⁶⁶ Retningslinjene til EDPB finner du her: <https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and-en>.

⁶⁷ EUs Charter of Fundamental Rights artikkel 20-26, Grunnloven § 98, Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven) kapittel 2

⁶⁸ Vi er oppmerksomme på at det finnes ulike definisjoner av rettferdighet i sammenheng med KI, og at det kan være umulig å oppfylle ulike definisjoner samtidig (jf. Miconi, T. (2017). *The impossibility of "fairness": a generalized impossibility result for decisions*, <https://arxiv.org/abs/1707.01195>).

⁶⁹ Smith, E. & Eckhoff, T. (2018). *Forvaltningsrett* (11. utgave). Universitetsforlaget.

⁷⁰ Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven) § 6. Forskjellsbehandling og diskriminering som vi omtaler videre, er knyttet til denne paragrafen.

⁷¹ Ibid § 7.

⁷² Ibid § 8.

⁷³ Ibid § 9.

Virksomhetene skal arbeide aktivt, målrettet og planmessig for å fremme likestilling og hindre diskriminering.⁷⁴ De har en plikt til å redegjøre for hvordan de tar hensyn til likestilling og ikke-diskriminering. Dette gjelder også i arbeidet med kunstig intelligens. Det innebærer at de skal kunne redegjøre for hvordan de innarbeider prinsipper, prosedyrer og standarder for likestilling og ikke-diskriminering i KI-systemene sine. Virksomhetene skal redegjøre for resultater og planer i årsrapport eller lignende.

Dersom virksomheten har behov for å behandle opplysninger om etnisitet, religion, politisk oppfatning, helseopplysninger, seksuell legning eller andre opplysningstyper som man generelt anser som spesielt sensitive, må de ha et særlig lovgrunnlag, siden dette i utgangspunktet er forbudt.⁷⁵ Bestemmelser i personvernforordningen eller i nasjonalt regelverk kan åpne for at virksomheten kan behandle slike opplysninger.

3.2.3 Krav om teknisk sikkerhet og robusthet

Den grunnleggende retten til beskyttelse av menneskelig verdighet og respekt for både fysisk og mental integritet⁷⁶ tilsier at KI-systemer ikke skal være til skade for mennesker. *Forebygging av skade* omfatter teknisk sikkerhet og robusthet, sikring mot manipulasjon eller misbruk av systemet, og at KI er utformet og implementert på en måte som tar særlig hensyn til sårbare grupper.

Teknisk sikkerhet handler om hvordan virksomheten ivaretar sikkerheten til infrastruktur, og hvilke systemer den har implementert for å håndtere krav til informasjonssikkerhet. *Handlingsplan for informasjonssikkerhet i statsforvaltningen* definerer informasjonssikkerhet som hvordan informasjonen blir beskyttet mot uønsket innsyn (konfidensialitet), at informasjonen er tilgjengelig når det er ønsket (tilgjengelighet) og at informasjonen er beskyttet mot manipulering (integritet).

Digitaliseringsdirektoratet bruker den samme definisjonen i sin begrepsliste om informasjonssikkerhet.^{77,78} Videre går informasjonssikkerhet ut på at organisasjonen og systemene er motstandsdyktige og evner å gjenopprette normaltilstand ved hendelser (robusthet)⁷⁹.

KI-systemer bygges på store mengder data, og bruker ofte maskinlæringsmodeller som inneholder informasjon fra data brukt for trening av modellen. Det kan være mulig å rekonstruere informasjon fra treningsdataene ut fra KI-systemet, og systemet bør derfor sikres med minimum samme sikkerhetsnivå som treningsdataene. Der bruken av KI bygger på personopplysninger, stiller i tillegg personvernforordningen krav om at det er gjort tekniske og organisatoriske sikkerhetstiltak som er egnet med hensyn til risikoen, blant annet kryptering av data, evne til gjenopprettelse og regelmessig vurdering av sikkerhetstiltak, jf. personvernforordningen artikkel 5 nr. 1 bokstav f og artikkel 32 nr. 1.

Digitaliseringsdirektoratet viser til at NS-ISO/IEC 27002 gir beste praksis for å oppfylle krav knyttet til administrasjon av informasjonssikkerhet (sikringstiltak).⁸⁰ Som anerkjente standarder vil denne revisjonen ta utgangspunkt i relevante anbefalinger i NS-ISO/IEC 27002:2017 (ISO 27002)⁸¹ og tilsvarende grunnprinsippene for IT-sikkerhet fra Nasjonal sikkerhetsmyndighet (NSM)⁸². Basert på disse standardene skal revisjonen vurdere følgende sikkerhetstiltak knyttet til KI applikasjoner: tilgangskontroll, håndtering av endringer, samt logging og gjenoppretting.

⁷⁴ Ibid § 24. Avsnittet er basert på samme kilde.

⁷⁵ Personvernforordningen artikkel 9 nr. 1. Avsnittet er basert på samme kilde.

⁷⁶ EUs Charter of Fundamental Rights artikkel 1 og 3, Lov om styrking av menneskerettighetenes stilling i norsk rett (menneskerettsloven)

⁷⁷ Kommunal- og moderniseringsdepartementet. (2015). *Handlingsplan for informasjonssikkerhet i statsforvaltningen*.

https://www.regjeringen.no/contentassets/b7d0918e555b418abda2993a71969cdc/handlingsplan_informasjonssikkerhet_staten.pdf

⁷⁸ Digitaliseringsdirektoratet, Digdir.no. (u.å.) *Begrepsliste*. Hentet 18. november 2023 fra

<https://www.digdir.no/informasjonssikkerhet/begrepsliste/3230#informasjonssikkerhet>

⁷⁹ Datatilsynet. (2018, 30. oktober). *Iverksette styringssystem for informasjonssikkerhet*. <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/informasjonssikkerhet-internkontroll/etablere-internkontroll/iverksette-styringssystem-for-informasjonssikkerhet/>

⁸⁰ Digitaliseringsdirektoratet.no, ISO/IEC 27002 Fra: <https://www.digdir.no/standarder/isoiec-27002/1696> (Hentet 17. november 2023).

⁸¹ Jf. referansekatalogen til Difi.

⁸² NSM: *Grunnprinsipper for IKT-sikkerhet 2.0*, særlig relevante tiltak under 2.1, 2.5, 2.6, 2.7, 2.9, 2.10, 3.2.

3.2.4 Krav om forklarbarhet og åpenhet

Det er en nødvendig forutsetning for rettssikkerheten at mekanismene i et KI-system som bestemmer utfallet eller resultatet av maskinell behandling, er forklarbare. Forklarbarhet er også nødvendig for å avgjøre om KI-systemets utfall er rettferdig og muliggjøre klage⁸³. Prinsippet om transparens innebærer at det skal være åpenhet om bruk av KI samt formålet og begrensningene. Transparens er nødvendig for å muliggjøre meningsfulle valg og ivareta menneskelig autonomi.

Grunnloven § 100 femte ledd fastsetter at enhver har rett til innsyn i statens dokumenter. Offentleglova gir allmennheten rett til innsyn i offentlig informasjon, og presiserer at dette gjelder saksdokument, journaler og lignende registre dersom ikke annet følger av annen særlovgivning.⁸⁴ Dette omfatter også sammenstilling av databaser.⁸⁵ Loven presiserer at vi med dokument kan forstå «ei logisk avgrensa informasjonsmengde som er lagra på eit medium for seinare lesing, lytting, fremsyning, overføring eller liknande». I forarbeidene til loven understrekes det at «[b]lant anna vil tekstar, lyd, bilete, teikningar og modellar vere omfatta av omgrepet».⁸⁶

Kunstig intelligens inkluderer per definisjon algoritmer som beskriver hvordan en modell skal lære, og hvordan en gitt oppgave skal løses. Kildekoden som bygger og instruerer modellen, er en logisk avgrenset informasjonsmengde, og kunstig intelligens er en metodikk som beskriver og automatiserer videreutvikling av modellen.

Åpenhetsprinsippet i personvernforordningen artikkel 5 nr. 1 bokstav a fremhever prinsippet om transparens som et av de mest sentrale prinsippene i forordningen. Det vil imidlertid ikke være hensiktsmessig å gi innsyn i algoritmen slik den er skrevet som programkode, men åpenhetsprinsippet betyr at offentligheten må forstå hva som skjer med de opplysningene som blir behandlet av kunstig intelligens, herunder at bruken skal være forståelig på en lett tilgjengelig måte. Kravet til åpenhet i forordningen handler både om informasjon om selve behandlingsaktiviteten (innholds krav) og om at informasjonen skal være forståelig (formkrav).⁸⁷

Ansvarlig bruk av kunstig intelligens innebærer at virksomheten må opplyse om at kunstig intelligens er brukt som beslutningsstøtte, i saksbehandling eller på andre måter i virksomhetens informasjonsbehandling, der det er rimelig/hensiktsmessig.⁸⁸ Det er avgjørende at offentlige virksomheter er åpne om slik bruk, for å bevare tillit til myndighetenes arbeid.

At bruken av kunstig intelligens skal være sporbar, forklarbar og gjennomiktig, følger også av alminnelige forvaltningsrettslige prinsipper. Sett i sammenheng med *lov om behandlingssåten i forvaltningssaker* (forvaltningsloven) § 17, som krever at «forvaltningsorganet skal påse at saken er så godt opplyst som mulig før vedtak treffes», lov om rett til innsyn i dokument i offentlig verksemd (offentleglova) og lov om arkiv (arkivloven) er hensikten å sikre at forvaltningen er transparent om de avgjørelsene som den fatter. Forvaltningsloven skal gi forvaltningens avgjørelser legitimitet og bidra til at borgerne og private virksomheter har tillit til forvaltningen. Arkivloven skal bidra til at disse avgjørelsene, og dokumentasjonen som ligger til grunn for dem, ikke mister sin legitimitet og bevisverdi i et langsiktig perspektiv.

3.2.5 Krav om åpenhet i saksbehandling

De ulovfestede reglene for god forvaltningsskikk følger i stor grad av langvarig rettspraksis. De ulovfestede reglene skal i hovedsak hindre at det blir tatt utenforliggende hensyn i saksbehandlingen, at noen usaklig forskjellsbehandles, og at avgjørelsene er urimelige eller vilkårlige. Det er viktig at

⁸³ Independent High Level Expert Group on Artificial Intelligence set up by the European Commission (2019), [Ethics guidelines for trustworthy AI](#)

⁸⁴ Offentleglova § 3.

⁸⁵ Offentleglova § 9.

⁸⁶ Ot. prp nr. 102 (2004–2005), s. 120.

⁸⁷ Personvernforordningen artikkel 5 (2).

⁸⁸ Europaparlamentet (2021) Artificial intelligence and public services

saksbehandlingen og vurderingene som blir gjort, er basert på saklighet (saklighetsprinsippet), og at vurderingene kan etterprøves. Med etterprøvbarehet menes det at det i ettertid kan være mulig å finne tilbake til de avveiningene og vurderingene som er gjort, og forstå hvorfor beslutningen er tatt.⁸⁹

Forvaltningsloven presiserer at enkeltvedtak skal inneholde en begrunnelse,⁹⁰ og at denne skal vise til de reglene vedtaket er fattet på bakgrunn av, og de faktiske forholdene som vedtaket bygger på.⁹¹ Sivilombudet har videre uttalt at det er i strid med god forvaltningsskikk å ikke gi veiledning om begrunnelsen dersom forvaltningen avviser anmodninger eller krav om begrunnelser.⁹² Dette er hensyn som må ivaretas hvis kunstig intelligens brukes i saksbehandling.

Nasjonal strategi for kunstig intelligens tar særlig opp utfordringen med åpenhet med henblikk på dyp læringsløsninger.⁹³ Dette blir særlig drøftet fordi slike løsninger er svært lite transparente. I strategien blir det foreslått å velge alternativer til dyp læringsløsninger for å ivareta åpenheten der forklarbarhet⁹⁴ er viktig.⁹⁵

Personvernkommissjonen uttrykker at de mener at «det er nødvendig at resultater av automatiserte prosesser som har direkte betydning for innbyggernes plikter, rettigheter, friheter og muligheter kan forklares».⁹⁶

Alle har rett til å reservere seg mot at avgjørelser som gjelder en selv, skjer i form av en helautomatisert behandling, altså uten involvering fra et menneske.⁹⁷ Virksomhetene kan likevel ta helt automatiserte avgjørelser dersom de har en lov- eller forskriftshjemmel til å gjøre det.⁹⁸ Virksomhetene har en plikt overfor de registrerte til å gi innsyn i eller gi informasjon om «den underliggende logikken» i helautomatiserte avgjørelser.⁹⁹

3.3 Krav til departementenes styring

Digitaliserings- og forvaltningsdepartementet har et overordnet ansvar for IT-politikken og digitalisering av offentlig forvaltning. Samordningsansvaret, som innebærer blant annet å identifisere sektorovergrepene og å initiere, koordinere og følge opp tverrgående tiltak, er lagt til Digitaliserings- og forvaltningsdepartementet. Datatilsynet og Digitaliseringsdirektoratet er virksomheter underlagt departementet. Det enkelte sektordepartementet har ansvaret for IT og digitalisering innenfor sin sektor.

Departementene har ifølge *reglement for økonomistyring i staten* (økonomireglementet) og *bestemmelser om økonomistyring i staten* (økonomibestemmelsene) det overordnede ansvaret for at underliggende virksomheter gjennomfører aktiviteter i samsvar med målene i Stortingets vedtak og forutsetninger. De grunnleggende styringsprinsippene er at det skal fastsettes mål- og resultatkrav, og at disse skal oppnås på en effektiv måte. Departementets styring må tilpasses virksomhetens egenart samt risiko og vesentlighet.

⁸⁹ Se bl.a. Boe, E.M. (2022) *Læren om myndighetsmisbruk i forvaltningen: krav til saklighet og forsvarlighet*. Fra: <https://www.idunn.no/doi/10.18261/jv.57.4.1> (Henretet 21. november 2023) eller NOU 2019:5, kap 34: *Bør deler av den ulovfestede forvaltningsretten lovfestes?*

⁹⁰ Forvaltningsloven § 24

⁹¹ Forvaltningsloven § 25

⁹² Sivilombudsmannen. (2021). *Årsmelding for 2020. Dokument 4 (2020–2021)*. https://www.sivilombudet.no/wp-content/uploads/2021/03/ARSMELDING_2020_Sivilombudsmannen_Nett.pdf

⁹³ Dyp læring er en kategori av maskinlæring med nevrale nettverk som består av flere lag, hvor hvert lag abstraherer informasjon videre.

⁹⁴ Forklarbar KI er et eget forskningsområde som utvikler metodikker for forklaringer om hvordan en modell kom frem til et visst resultat for gitt inndata. For kompliserte modeller kan man benytte separate forklaringsmodeller i tillegg til hovedmodellen, som igjen inneholder antakelser og forenklinger som kan være vanskelig å begrunne.

⁹⁵ Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*. <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/>

⁹⁶ NOU 2022: 11. (2022). *Ditt personvern – vårt felles ansvar*. Kommunal- og distriktsdepartementet, s. 17.

⁹⁷ Personvernforordningen artikkel 22.

⁹⁸ Personvernforordningen artikkel 22 nr. 2 bokstav b slår fast at det må være «fastsatt egnede tiltak for å verne den registrertes rettigheter, friheter og berettigede interesser».

⁹⁹ Personvernforordningen artikkel 15 nr. 1 bokstav h.

I økonomireglementet og økonomibestemmelsene går det videre frem at departementet har et overordnet ansvar for at underliggende virksomheter har et fungerende og forsvarlig system for internkontroll, bruker tildelte ressurser effektivt og forvalter oppgaver på en forsvarlig måte. Departementet skal sørge for at det blir gjennomført evalueringer for å skaffe kunnskap om måloppnåelsen og resultatene på området. Frekvensen og omfanget av evalueringer skal bestemmes ut fra virksomhetens egenart, risiko og vesentlighet. Behovet må vurderes ut fra kvaliteten på og omfanget av den øvrige rapporteringen.

4 Bruk av kunstig intelligens i staten

For å belyse problemstilling 1 om statlige virksomheter bruker kunstig intelligens til å effektivisere og forbedre tjenester, har vi kartlagt omfang og bruk av KI i staten. Totalt 194 statlige virksomheter, helseforetak og heleide selskaper i statlig sektor (heretter samlet omtalt som «virksomheter») svarte på et brev med spørsmål. 85 av disse svarte bekreftende på at de har brukt, bruker eller utvikler KI, i tråd med definisjonen som vi bruker i denne undersøkelsen. I tillegg svarte 15 virksomheter at de vurderer å benytte KI i fremtiden. Som omtalt i metodedelens fulgte vi opp svarene på brevet med en spørreundersøkelse rettet mot faglige kontaktpersoner i virksomhetene, som handlet konkret om KI-systemene som er under utvikling eller i bruk.

Enkelte deler av revisjonen som omhandler kartlegging av utvikling og bruk av KI, ble rapportert som et vedlegg til Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen*.¹⁰⁰

4.1 Stor variasjon i bruk av KI blant statlige virksomheter

I brevet om kunstig intelligens stilte vi virksomhetene disse tre spørsmålene:

- Har dere tatt i bruk, eller har dere pågående prosjekter for å ta i bruk, KI?
- Har dere brukt KI til oppgaver i et tidsavgrenset prosjekt (engangsbruk) som er avsluttet, i tidsrommet 2018 og frem til nå?¹⁰¹
- Har dere avbrutt eller stoppet en prosess hvor dere vurderte, utviklet eller brukte KI, i tidsrommet 2018 og frem til nå?

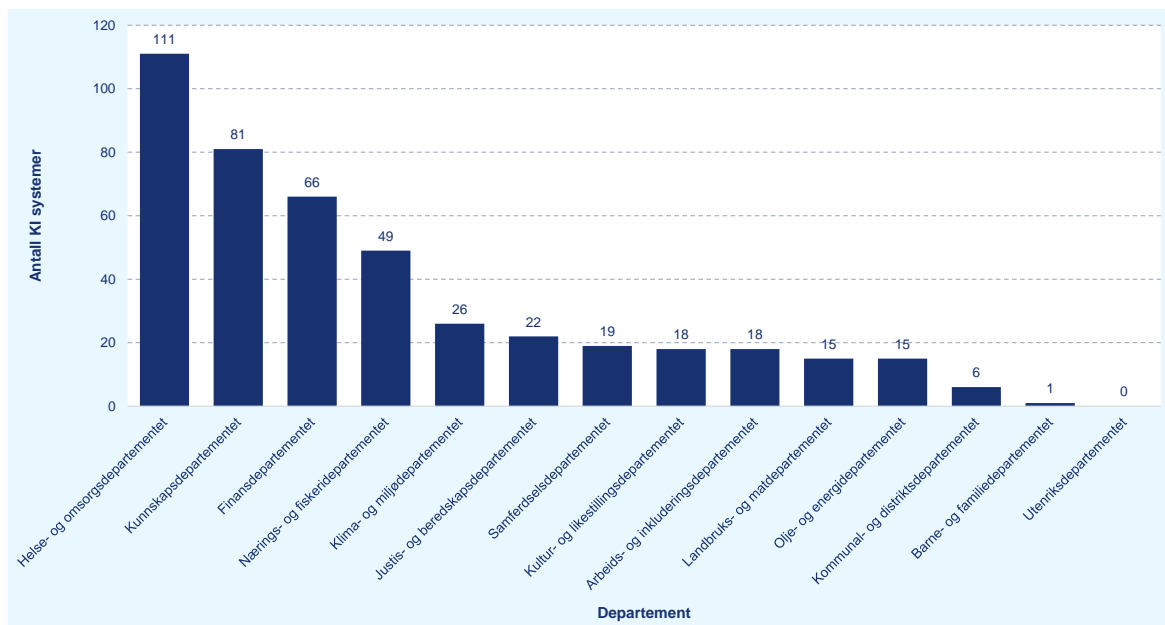
44 prosent av virksomhetene rapporterer at de har noe erfaring med KI, det vil si at de svarte «ja» på minst ett av disse tre spørsmålene. 56 prosent svarer at de ikke har noe erfaring med kunstig intelligens i perioden fra 2018 til våren 2023. Helseforetakene har høyest andel (73 prosent) som har erfaring med utvikling eller bruk av kunstig intelligens, etterfulgt av heleide selskaper (44 prosent) og statlige virksomheter (36 prosent). Flertallet av KI-systemene er under utvikling eller i bruk. De resterende er tidsavgrensede prosjekter som er avsluttet, eller prosjekter som er stoppet.

Det er stor spredning i antallet rapporterte KI-systemer mellom departementenes underliggende virksomheter og selskaper. Det er størst KI-aktivitet, i form av rapporterte prosjekter, i virksomheter under Helse- og omsorgsdepartementet, jf. Figur 2. Deretter følger Kunnskapsdepartementet, Finansdepartementet og Nærings- og fiskeridepartementet (inkludert de heleide selskapene). Nærmere 70 prosent av de innrapporterte KI-prosjektene ligger under et av disse fire departementene.

¹⁰⁰ Riksrevisjonen. (2023). Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen*, vedlegget «Kartlegging av utvikling og bruk av kunstig intelligens i statsforvaltningen og statseide selskaper».

¹⁰¹ Brevene med spørsmål ble sendt i våren 2023, så tidsrommet dekket av undersøkelsen er fra 2018 til vår 2023.

Figur 2 Antall KI-systemer rapportert, fordelt på fagdepartementer, fra 2018 til vår 2023¹⁰²



Kilde: Svar på brev til statlige virksomheter og heleide statlige selskaper (n = 130)

Tas det hensyn til at antall virksomheter i undersøkelsen varierer per sektor finner man at i gjennomsnitt er antall KI-prosjekter høyest blant virksomhetene under Finansdepartementet etterfulgt av Klima- og miljødepartementet og Helse- og omsorgsdepartementet.

Finansdepartementet og Arbeids- og inkluderingsdepartementet viser i intervju til at de ikke har gitt noen føringer til underliggende virksomheter om å ta i bruk kunstig intelligens. Justis- og beredskapsdepartementet viser i intervju til at de ikke har definert spesifikke mål for bruk av kunstig intelligens, men har gitt enkelte underliggende virksomheter føringer om vurdering av ny teknologi inklusive kunstig intelligens. Ingen av disse departementene har imidlertid noen formell rolle i å følge opp særskilte deler av den nasjonale strategien for kunstig intelligens. Alle de tre ovennevnte departementene legger vekt på at virksomhetene skal sikre effektiv ressursbruk, men at det er opp til virksomhetene å velge og bruke de virkemidlene og den teknologien som gir de beste løsningene innenfor sine ressursrammer.¹⁰³ Generelt legger for eksempel Finansdepartementet stor vekt på at de underliggende virksomhetene skal være utviklingsorienterte, men ønsker ikke å detaljstyre disse med utgangspunkt i spesifikke teknologier.¹⁰⁴ I tildelingsbrev til samtlige underliggende virksomheter for 2023 viser Finansdepartementet til Digitaliseringsrundskrivet for føringer om digitalisering og strategisk IT-utvikling.

Helse- og omsorgsdepartementet viser på sin side til kunstig intelligens i brev til det tverrfaglige koordineringsprosjektet «Bedre bruk av kunstig intelligens», som ble startet som en del av arbeidet med Meld. St. 7 (2019–2020) *Nasjonal helse- og sykehusplan 2020–2023*. Arbeidet er videreført som en del av etatens ordinære ansvar. Helse- og omsorgsdepartementet har også omtalt kunstig intelligens i oppdragsbrevet for 2024 til de regionale helseforetakene på følgende måte¹⁰⁵:

¹⁰² Riksrevisjonen. (2023). *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen* (Dokument 3:8 (2023–2024)), vedlegg «Kartlegging av utvikling og bruk av kunstig intelligens i statsforvaltningen og statseide selskaper».

¹⁰³ Verifisert referat fra intervju med Finansdepartementet 29. november 2023, Arbeids- og inkluderingsdepartementet 18. desember 2023 og Justis- og beredskapsdepartementet 5. februar 2024.

¹⁰⁴ Verifisert referat fra intervju med Finansdepartementet 29. november 2023.

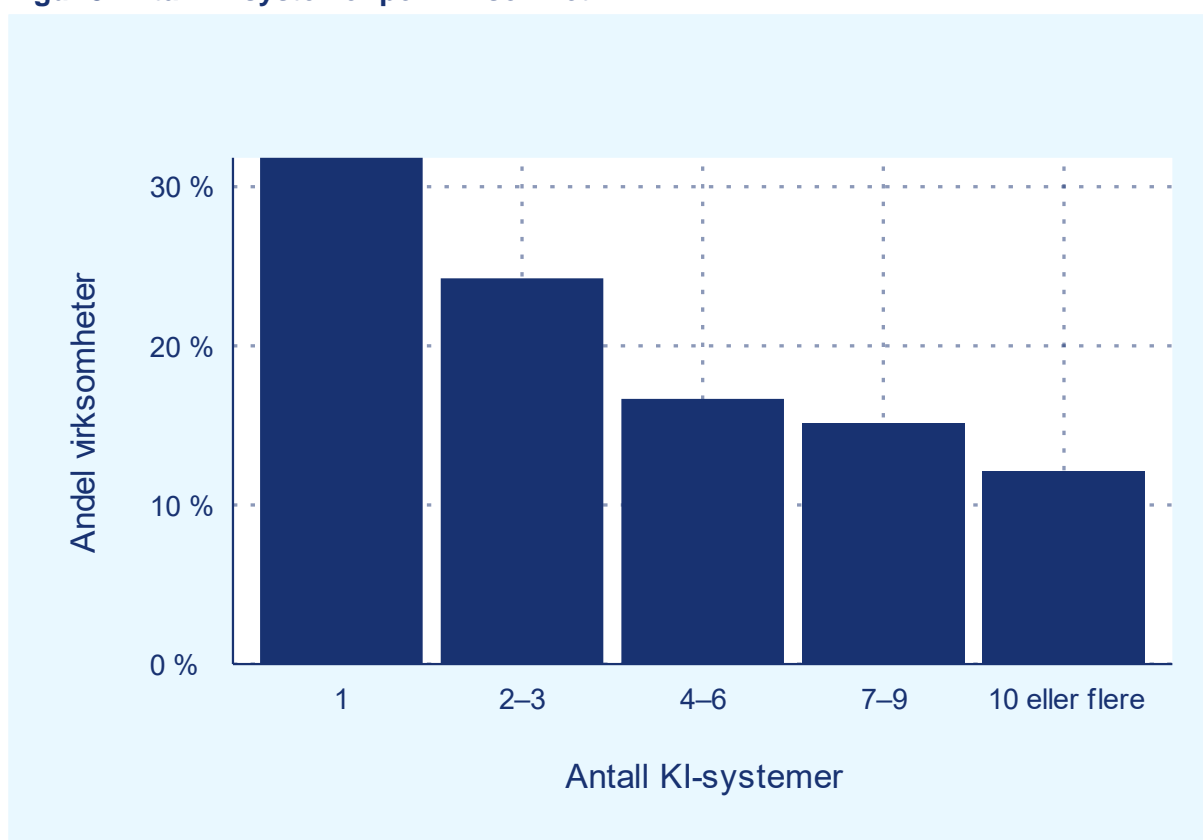
¹⁰⁵ Helse- og omsorgsdepartementet. (2024). *Oppdragsdokument 2024. Helse Sør-Øst RHF*. <https://www.regjeringen.no/contentassets/bd8a11644b744dec8a8dc452794000e4/oppdragsdokument-2024-helse-sor-ost-rhf-2601.pdf>

Utvikling, implementering og spredning av teknologi, kunstig intelligens, nye arbeidsformer og mer presis diagnostikk og behandling gjennom innovasjon og næringslivssamarbeid, er sentralt for å skape en bærekraftig helsetjeneste for alle.

Teknologirådet (2021) anslo at data om helseopplysninger utgjør en stor andel av verdens lagrede data, og at omfanget øker raskt. Dette er et viktig grunnlag for utvikling av tjenester basert på kunstig intelligens i helse- og omsorgstjenesten, ifølge rådet.¹⁰⁶

De faglige kontaktpersonene som fikk tilsendt spørreundersøkelsen, fikk spørsmål om hvor mange KI-systemer de kunne uttale seg om i sin virksomhet. Totalt kunne de uttale seg om omtrent 300 systemer¹⁰⁷. Figur 3 viser hvor mange KI-systemer som er oppgitt per virksomhet. Her ser vi at om lag 30 prosent av virksomhetene oppgir at de har ett KI-system. Over halvparten (56 prosent) har ett til tre KI-systemer.

Figur 3 Antall KI-systemer per virksomhet



Kilde: Riksrevisjonens spørreundersøkelse (n = 104 respondenter, aggregert til 66 virksomheter).

4.1.1 Hvor mye investerer statlige virksomheter i bruk og utvikling av KI?

Digitaliserings- og forvaltningsdepartementet viser i korrespondanse¹⁰⁸ med Riksrevisjonen til at det har ønsket å få en oversikt over størrelsen på investeringer i KI i Norge, fortrinnsvis som en del av EUs oversikt, eller i det minste i form av sammenlignbare tall. Departementet var i 2021 i kontakt med SSB om dette, men lyktes aldri å få noen konkrete svar fra SSB.

¹⁰⁶ Teknologirådet. (2021). *Kunstig intelligens i klinikken – seks trender for fremtidens helsetjeneste*. [Kunstig intelligens i klinikken – Seks trender for fremtidens helsetjeneste - Teknologirådet \(teknologiradet.no\)](https://www.teknologiradet.no/kunstig-intelligens-i-klinikken)

¹⁰⁷ Et av svaralternativene er «Flere enn 10», dette ble satt til «11». Antakelig vil tallet reelt være høyere, men dette er et minimumsestimert.

¹⁰⁸ Epost fra daværende Kommunal- og Distriktsdepartementet, 9. mars 2023

I svar på brevene med spørsmål er det flere av virksomhetene som oppgir at det er dyrt å utvikle og drifte KI-systemer. Flere har ikke nok midler til å gjøre nødvendige investeringer. Det er heller ikke alltid enkelt å dokumentere at de nødvendige investeringene vil gi ønsket effekt eller en gevinst. Mange virksomheter mangler intern kompetanse og må derfor kjøpe kompetanse til både utvikling og drift av teknologien. Det trekkes også frem at enkelte hyllevareløsninger har betydelige kostnader. Noen virksomheter påpeker at det er store vedlikeholdskostnader, og det er kostnadskrevenende å endre og overvåke KI-systemene. Mange virksomheter har dermed ikke et økonomisk handlingsrom til å utarbeide og benytte KI-systemer.

For å få mer kunnskap om kostnader knyttet til å utvikle og ta i bruk KI-systemer ba vi kontaktpersonene i spørreundersøkelsen om å oppgi hvor mye hvert av KI-systemene hadde kostet i kroner, ansatte og tid frem til systemene ble tatt i bruk. Slike beregninger er ofte krevende å gjøre, og de kan være upresise. Det generelle bildet er imidlertid at de fleste KI-prosjektene er små. Mer enn en tredjedel av prosjektene anslås å ha brukt under 1 million kroner. I tillegg kan høye kostnader til å kjøpe nødvendig infrastruktur som servere og sky-løsninger til KI-systemene være utelatt fra disse beregningene. Kontaktpersonene oppgir at ca. 70 prosent av KI-systemene har fem eller færre ansatte som har jobbet på prosjektet. Kun 12 prosent oppgir å ha brukt mer enn to år på et KI-prosjekt.

4.2 Over halvparten av KI-systemene bruker kun enkle modeller for å tolke data

I spørreundersøkelsen ble respondentene bedt om å svare på hvilke typer teknologi som er brukt i KI-systemene, og som innebærer *tolkning og behandling av strukturerte eller ustrukturerte data*, jf. definisjonen av kunstig intelligens i kapittel 1.1 og 1.4.

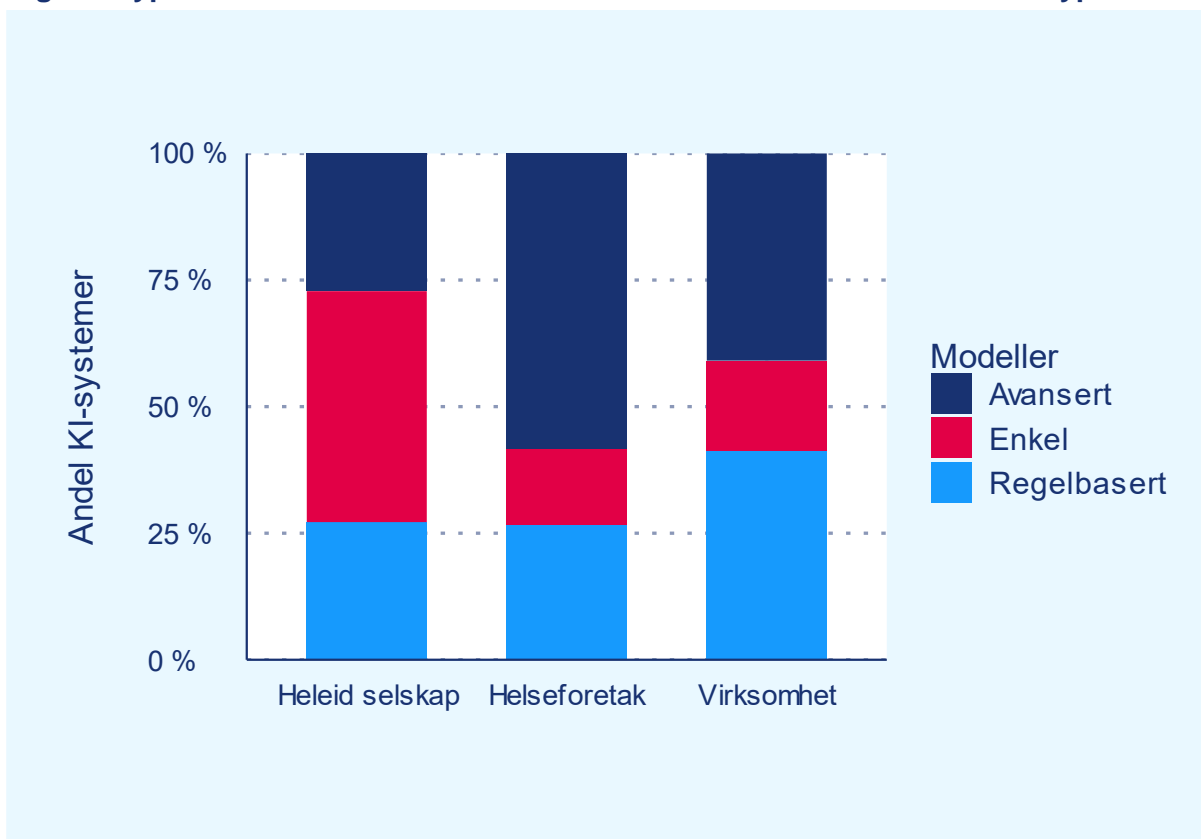
Av de ca. 300 KI-systemene er det 56 prosent som bruker regelbaserte systemer eller enkle modeller for å tolke og behandle data, 44 prosent av KI-systemene bruker mer avansert maskinlæring. På *virksomhetsnivå* er det 39 prosent av virksomhetene (26 av 66) som kun har regelbaserte eller enkle KI-systemer. 9 prosent (6 virksomheter) har bare regelbaserte systemer, som etter de fleste definisjoner ikke regnes som KI.¹⁰⁹ Med en strengere definisjon på et KI-system ville antallet oppgitte KI-systemer ha vært lavere.

Virksomheter med mange KI-systemer har flere enkle modeller, slik at regelbaserte modeller er enda hyppigere og utgjør 36 prosent av alle KI-systemer. Av de sju virksomhetene som har oppgitt at de har ti eller flere KI-systemer, er det kun to som har avanserte modeller som utgjør en betydelig del av deres KI-portefølje, en virksomhet bruker begge deler, og flertallet (fire virksomheter) bruker primært regelbaserte systemer.

Skillet mellom ulike typer systemer illustreres i figuren nedenfor, som viser fordelingen av typer modeller som brukes til å tolke data i de ulike KI-systemene, for de tre ulike virksomhetstypene.

¹⁰⁹ Ikke alle respondentene var enige i at enkle, regelbaserte systemer bør være en del av undersøkelsen. De mener det de benytter, ikke er avansert nok, eller ikke er en stor nok del av prosjektet. En respondent oppgir for eksempel at deres prosjekt ikke ville vært definert som bruk av KI basert på Store norske leksikons definisjon av KI. De er likevel bedt om å svare.

Figur 4 Type modeller som brukes for å tolke data for de ulike virksomhetstypene



Kilde: Riksrevisjonens spørreundersøkelse (n = 104 respondenter, herav 15 i heleide selskaper, 29 i helseforetak og 60 i virksomheter).

Figur 4 viser at helseforetakene i større grad enn de andre benytter seg av avanserte modeller. De oppgir at 58 prosent av KI-systemene deres benytter seg av avanserte modeller for å tolke og behandle data. Tilsvarende tall for heleide selskaper og virksomheter er henholdsvis 27 prosent og 41 prosent.

Gjennom svarene på brevene ser vi at helseforetakene har stor bredde i KI-prosjektene og helsesektoren fremstår som mer moden enn andre sektorer. Eksempelvis er Vestre Viken HF det første helseforetaket i Norge til å ta i bruk bildediagnostikk av bruddpasienter utført av et KI-system til klinisk beslutningsstøtte. KI-systemet analyserer om pasientene har beinbrudd eller ikke, på ett minutt.¹¹⁰ Et annet eksempel er fra Akershus universitetssykehus HF, som har utviklet en modell for å bedre utnytte informasjonen fra EKG-målinger som foretas på sykehuset.

Tallene hittil i undersøkelsen baserer seg på svar om alle KI-systemer i virksomhetene. Ut fra gitte kriterier ba vi hver kontaktperson om å ikke svare om flere enn tre KI-systemer.¹¹¹ Dette utvalget er på 189 systemer (mot ca. 300 totalt), og består i større grad av mer avanserte KI-systemer.¹¹² Til forskjell fra totalpopulasjonen som besto av i underkant 44 prosent KI-systemer som bruker avansert maskinlæring til å tolke dataene sine, består dette utvalget nå av 51 prosent. Omtrent halvparten¹¹³ av KI-systemene i dette utvalget er i bruk, mens den andre halvparten er under utvikling. Helseforetakene

¹¹⁰ Eriksen, K. F. (2024, 1. mars). Vestre Viken har innført benbrudd-KI på alle sykehus - over 11.000 pasienter sjekket. *Dagens medisin*. <https://www.dagensmedisin.no/kunstig-intelligens-vestre-viken-hf/vestre-viken-har-innfort-benbrudd-ki-pa-alle-sykehus-over-11000-pasienter-sjekket/62>

¹¹¹ Se kapittel 2 om metode for en beskrivelse av hvordan avgrensningen ble gjort.

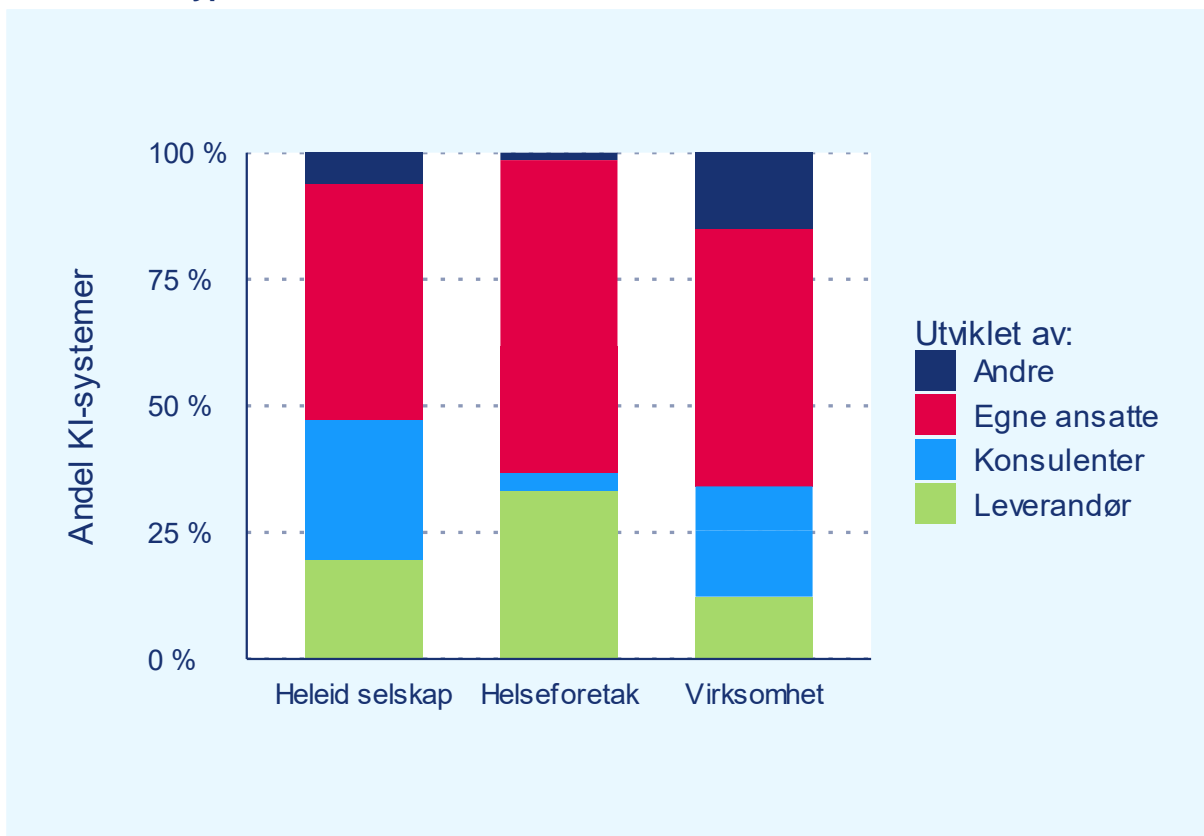
¹¹² Jf. fotnoten ovenfor vil det derfor være mindre uenighet om at det faktisk bør være inkludert i undersøkelsen.

¹¹³ I utvalget var 101 av 189 KI-systemer (det vil si ca. 53 prosent) allerede i bruk, 88 (47 prosent) under utvikling. Ytterligere 9 KI-systemer var angitt som lagt på is, avsluttet eller som bare ble vurdert anskaffet, og ble ikke tatt med i undersøkelsen, som i utgangspunktet skulle omfatte KI-systemer i bruk eller under utvikling.

skiller seg ut, der 63 prosent av prosjektene er under utvikling, i heleide selskaper er det motsatt med i underkant av 70 prosent som er i bruk.

Det er i varierende grad egne ansatte som utvikler KI-systemene, både i de heleide selskapene, helseforetakene og andre virksomheter, jf. figur 5. Ifølge respondentene i spørreundersøkelsen er ca. 50 prosent av KI-systemene selvutviklet blant de heleide selskapene og virksomhetene. For helseforetakene er 64 prosent av alle KI-systemene selvutviklet. Helsesektoren bruker i tillegg i større grad ferdigutviklede leverandørsystemer enn andre statlige virksomheter.

Figur 5 Oversikt over hvem som har bidratt til utvikling av modellene, fordelt på virksomhetstype¹¹⁴



Kilde: Riksrevisjonens spørreundersøkelse (n = 189 KI-systemer).

4.3 Virksomhetenes formål med bruk av kunstig intelligens

I brevet med spørsmål virksomhetene fikk tilsendt, stilte vi et åpent spørsmål om formålet med KI-prosjektene. Virksomhetene oppgir at formålet med KI-systemene i stor grad handler om å effektivisere og forbedre sine oppgaver. Særlig mange oppga å forbedre kjerneoppgaver eller lage nye tjenestetilbud. Å frigjøre tid og/eller effektivisere i ulike varianter og bedre informasjonsgrunnlag var også viktig.

Respondentene i spørreundersøkelsen ble spurt om hvilke oppgaver KI-systemene brukes eller utvikles til. Figur 6 viser andelen KI-systemer fordelt på de ulike svaralternativene, med separate søyler for om KI-systemet allerede er i bruk eller er under utvikling. Det ble i spørsmålet presisert hva som var ment med de ulike kategoriene¹¹⁵, men svaralternativene overlapper noe. For eksempel ble

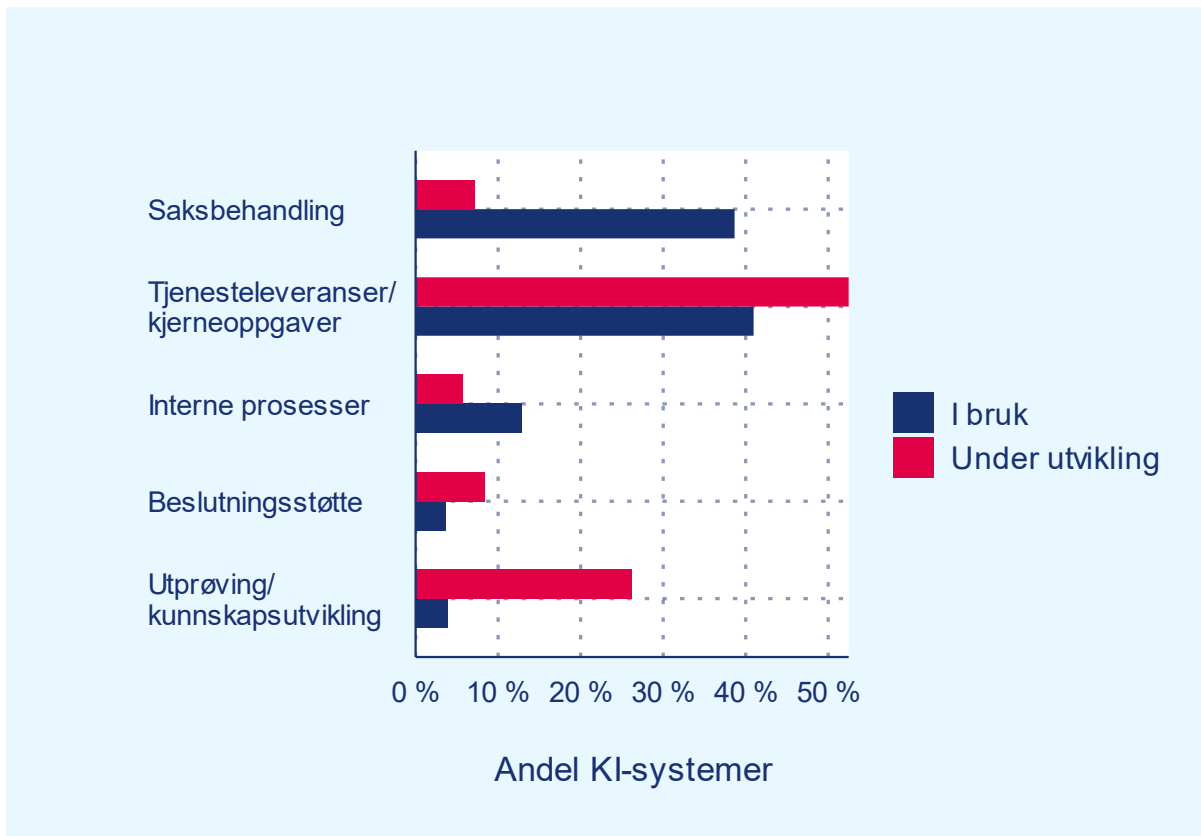
¹¹⁴ Kategorien «andre» står her for «andre samarbeidspartnere». Dette kan for eksempel være universiteter, eller lignende.

¹¹⁵ Jf. spørreundersøkelse, spm. 19, vedlegg 2.

«interne prosesser» definert til å handle om drift, administrasjon, HR og så videre.

«Tjenesteleveranser eller andre kjerneoppgaver» ble definert til det som vanligvis er rettet mot eksterne brukere, og er et alternativ til kategorien «saksbehandling» for virksomheter som ikke gjennomfører saksbehandling.

Figur 6 Bruk av KI-systemer til ulike oppgaver



Kilde: Riksrevisjonens spørreundersøkelse (n = 189 KI-systemer).

Flesteparten av KI-systemene både i bruk og under utvikling er laget for tjenesteleveranser eller andre kjerneoppgaver. Blant de KI-systemene som allerede er tatt i bruk, er nærmere 40 prosent angitt å være utviklet for saksbehandling. Få systemer oppgis å være rettet mot å gi beslutningsstøtte for strategi og styring av virksomheten eller selskapet.

Nærmere informasjon fra spørreundersøkelsen om KI-systemene viser at

- over halvparten av KI-systemene er delvis automatisert og i underkant av 30 prosent er fullautomatisert¹¹⁶
- nesten halvparten av KI-systemene brukes til klassifisering (tildeling til en gruppe)
- i underkant av 40 prosent er brukt av eksterne brukere¹¹⁷
- ca. en tredjedel av KI-systemene analyserer tekst eller bilder¹¹⁸, mindre enn 5 prosent tale

Figur 7 viser hva virksomhetene oppga som formål¹¹⁹ med hvert KI-prosjekt. Merk at ett KI-system kan ha flere formål. Figuren skiller ikke mellom hovedformål og sideformål, og det er mulig at respondentenes forståelse av de ulike formålscategoriene overlapper. Det er imidlertid tydelig at hovedformålet er å bidra til en form for effektivisering: Størst andel av KI-systemene har som formål å

¹¹⁶ Spørreundersøkelse, spørsmål 25

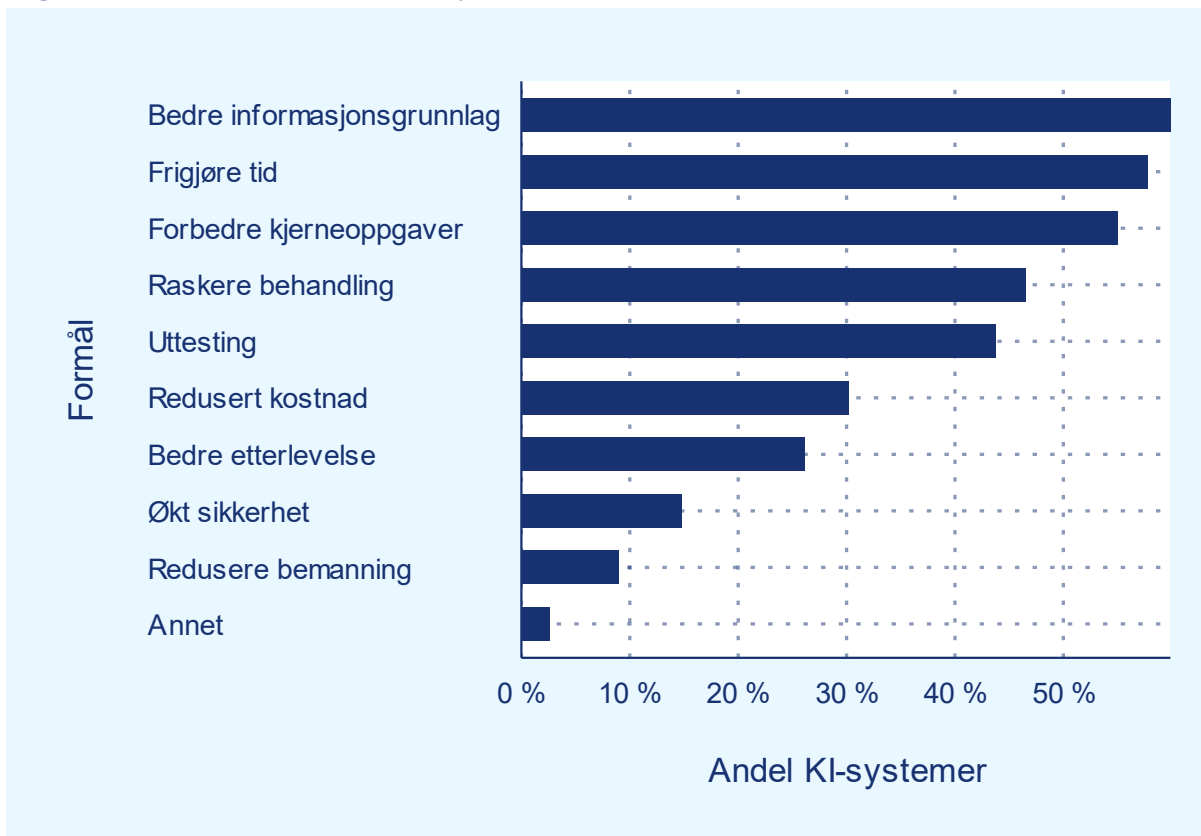
¹¹⁷ Spørreundersøkelsen, spørsmål 24. Flere typer brukere per KI-system mulig.

¹¹⁸ Spørreundersøkelsen, spørsmål 11: Hhv. 34 prosent og 31 prosent analyserer tekst og/eller bilder.

¹¹⁹ Spørreundersøkelsen, spørsmål 20 og 22. Figur 7 skiller ikke mellom hovedformål og andre formål.

gi bedre informasjonsgrunnlag for beslutninger, frigjøre tid til å løse andre eller mer krevende oppgaver, eller å forbedre kjerneoppgavene eller lage nye tjenester/tilbud, med 55–60 prosent hver. Nesten halvparten av KI-systemer har også raskere saksbehandling eller bearbeidelse av henvendelser som formål. Mindre enn 10 prosent av KI-systemene har som formål å redusere bemanningen i virksomheten, for kun 1 prosent oppga respondenter det som hovedformål.

Figur 7 Formål med bruk av KI-systemer



Kilde: Riksrevisjonens spørreundersøkelse (n = 189 KI-systemer).

4.3.1 De fleste virksomhetene oppgir at KI-prosjektene har gitt ønsket forbedring

Cirka 75 prosent av respondentene i spørreundersøkelsen oppgir at målet om forbedring er nådd for prosjekter som allerede er tatt i bruk, jf. Figur 8. I tillegg sier ca. 15 prosent at målet er delvis nådd. De resterende sier at de ikke vet, eller at det for tidlig å si. Ingen svarer at målet ikke er nådd. Her ser vi ingen store forskjeller mellom virksomhetstypene.

Hvis vi ser nærmere på hvem som har bidratt til utvikling av KI-systemene, så fremkommer det at de som har benyttet en leverandør, i noe lavere grad rapporterer om måloppnåelse som forventet. For disse systemene er det ca. 55 prosent som svarer «Ja, som forventet», mens nesten 25 prosent svarer «Delvis».

Blant virksomhetene som er med i vår dybdeundersøkelse, som har erfaring med å utvikle og bruke kunstig intelligens, var det flere eksempler på positive erfaringer og resultater:¹²⁰

¹²⁰ Jf. tabell 3 i kapittel 6.2 for gevinster oppnådd i KI-systemene i caseundersøkelsen

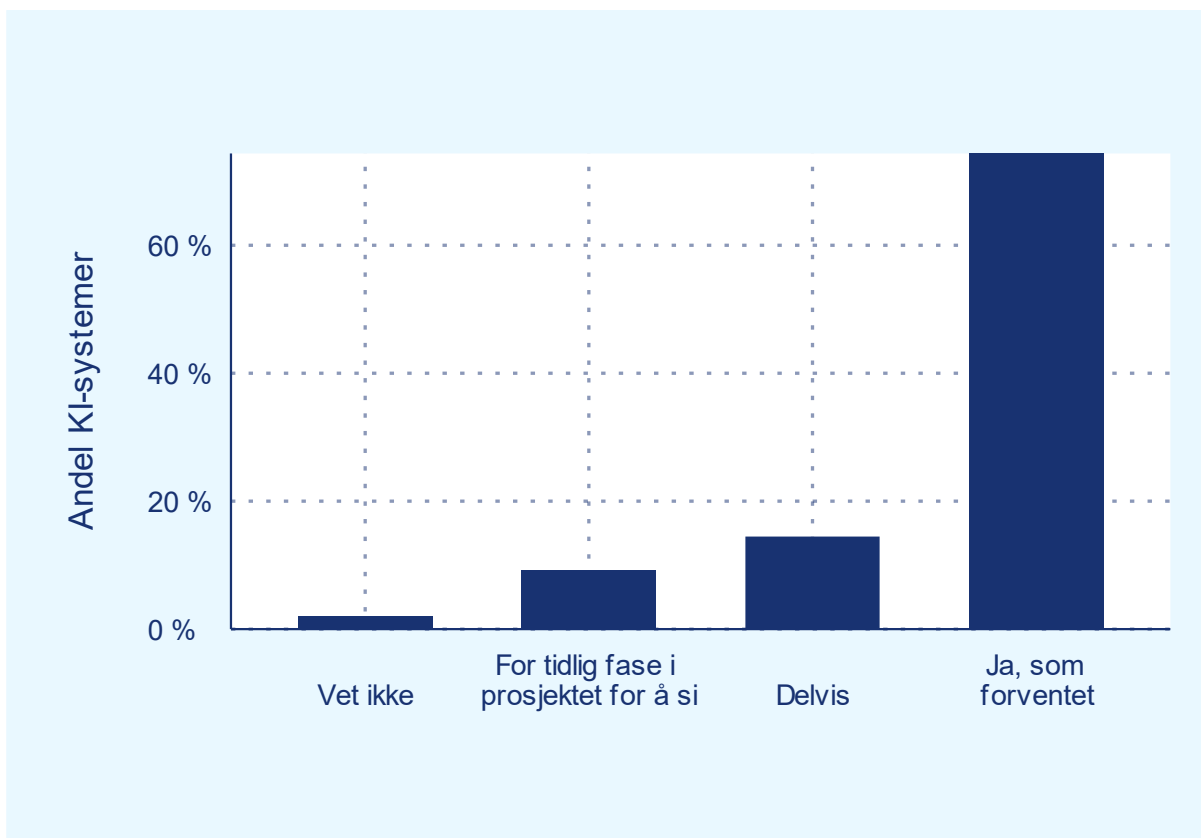
- Statens pensjonskasses bruk av maskinlæringsmodeller treffer vesentlig bedre enn de tidligere modellene de hadde for å predikere produksjonsutbetalinger. Erfaringen er at KI er god til å levere innsikt i prosesser og flaskehalsar i deres systemer, og at KI dermed kan brukes bredt til kun å utvikle bedre tjenester og øke effektiviteten.
- Skatteetaten viser til muligheter til å frigjøre ressurser til andre oppgaver ved at KI-systemene kan overta noen rutineoppgaver. Her er det imidlertid viktig med god endringsledelse overfor de som blir berørt. Som ett eksempel har skatteetatens fradragsmodell for risikobasert kontrollutplukk av skattefradrag gitt svært lønnsomme kontroller med en skattevirkning på 309 millioner kroner i 2021.
- Arbeidstilsynets bruk av KI som beslutningsstøtte ved utvalg av virksomheter for tilsyn kan vise til mer treffsikre og derav effektive tilsyn. For eksempel ved at valg av tilsyn med virksomheter i høyere risikogrupper øker sannsynligheten for å avdekke flere og mer alvorlige lovbrudd.

Flere nasjonale strategier og utredninger peker på kunstig intelligens som en teknologi som kan gi bedre helse- og omsorgstjenester, gi bedre grunnlag for forskning og innovasjon, og bidra til mer effektiv ressursbruk, noe som kan gi helt nye typer helsetjenester og måter å jobbe på.¹²¹ Akershus universitetssykehus HF viser i intervju til at det satser mye på innovasjons- og forskningsarbeid innenfor kunstig intelligens. Det er et stort behov for å bygge opp KI-baserte tjenester internt ved sykehuset, og ambisjonen deres er at alle interne KI-prosjekter skal kunne settes inn i produksjon ved sykehuset. Forslag til KI-prosjekter som ikke vurderes å ha et potensial for å bli satt inn i produksjonen, vil ikke bli prioritert. Ambisjonen er videre at KI-prosjektene skal kunne munne ut i nyttige beslutningsverktøy for klinikerne som jobber ved sykehuset. Modellene skal ikke bare kunne brukes til klinisk utprøving.¹²²

¹²¹ Meld. St. 7 (2019–2020). *Nasjonal helse- og sykehusplan 2020–2023*. Helse- og omsorgsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-7-20192020/id2678667> og NOU 2023: 4. (2023). *Tid for handling*. <https://www.regjeringen.no/contentassets/3371ef958f2148bebd326f0749a1213d/no/pdfs/nou202320230004000dddpdfs.pdf>

¹²² Verifisert referat fra intervju med Akershus universitetssykehus HF 20. november 2023.

Figur 8 Måloppnåelse for KI-systemer i bruk



Kilde: Riksrevisjonens spørreundersøkelse (n = 101 KI-systemer som er i bruk).

4.4 Bruk av personopplysninger og betydning for enkeltpersoner

Om KI i staten utvikles på en ansvarlig måte, omtales nærmere i kapittel 6. Helt sentralt her står kravet om at personopplysninger må håndteres på en god og lovlig måte. Det er derfor interessant å se på i hvilken grad eksisterende KI-systemer bruker personopplysninger, og om enkeltpersoner kan bli direkte berørt av utfallet generert av et KI-system.

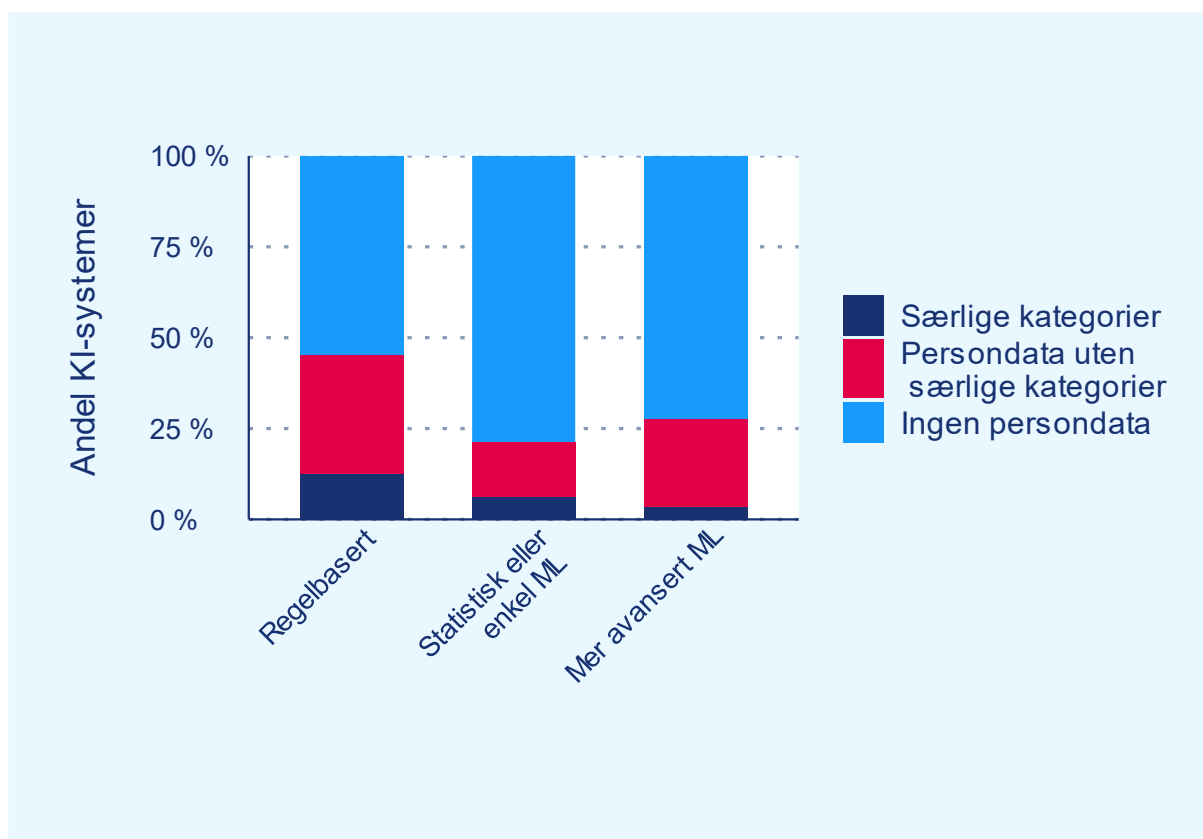
Spørreundersøkelsen viser at totalt 43 prosent av KI-systemene bruker persondata, hvorav omtrent halvparten av disse (en femtedel totalt) bruker særlige kategorier¹²³ av personopplysninger som definert i personvernforordningen. For fire KI-systemer har respondenten oppgitt at hen ikke vet om persondata er brukt: Disse KI-systemene er utviklet av en leverandør eller andre samarbeidspartnere.

Ser man på KI-systemer utenfor helsesektoren, er det kun 25 prosent som bruker personopplysninger uten særskilte kategorier, og 7 prosent som bruker én eller flere særskilte personopplysninger. Personopplysninger brukes også noe oftere i forbindelse med regelbaserte KI-systemer enn med KI-systemer basert på maskinlæring, jf. figur 9¹²⁴.

¹²³ Særlige kategorier av personopplysninger er definert i personvernforordningen artikkel 9 som «rasemessig eller etnisk opprinnelse, politisk oppfatning, religion, filosofisk overbevisning eller fagforeningsmedlemskap, samt behandling av genetiske opplysninger og biometriske opplysninger med det formål å entydig identifisere en fysisk person, helseopplysninger eller opplysninger om en fysisk persons seksuelle forhold eller seksuelle orientering» (jf. lov om behandling av personopplysninger artikkel 9 nr. 1).

¹²⁴ Figuren utelater to KI-systemer fordi respondentene ikke vet hvilken type modell som er brukt, og to KI-systemer fordi respondentene ikke vet om persondata er brukt.

Figur 9 Andel KI-systemer av ulike typer som behandler personopplysninger, helsesektoren unntatt



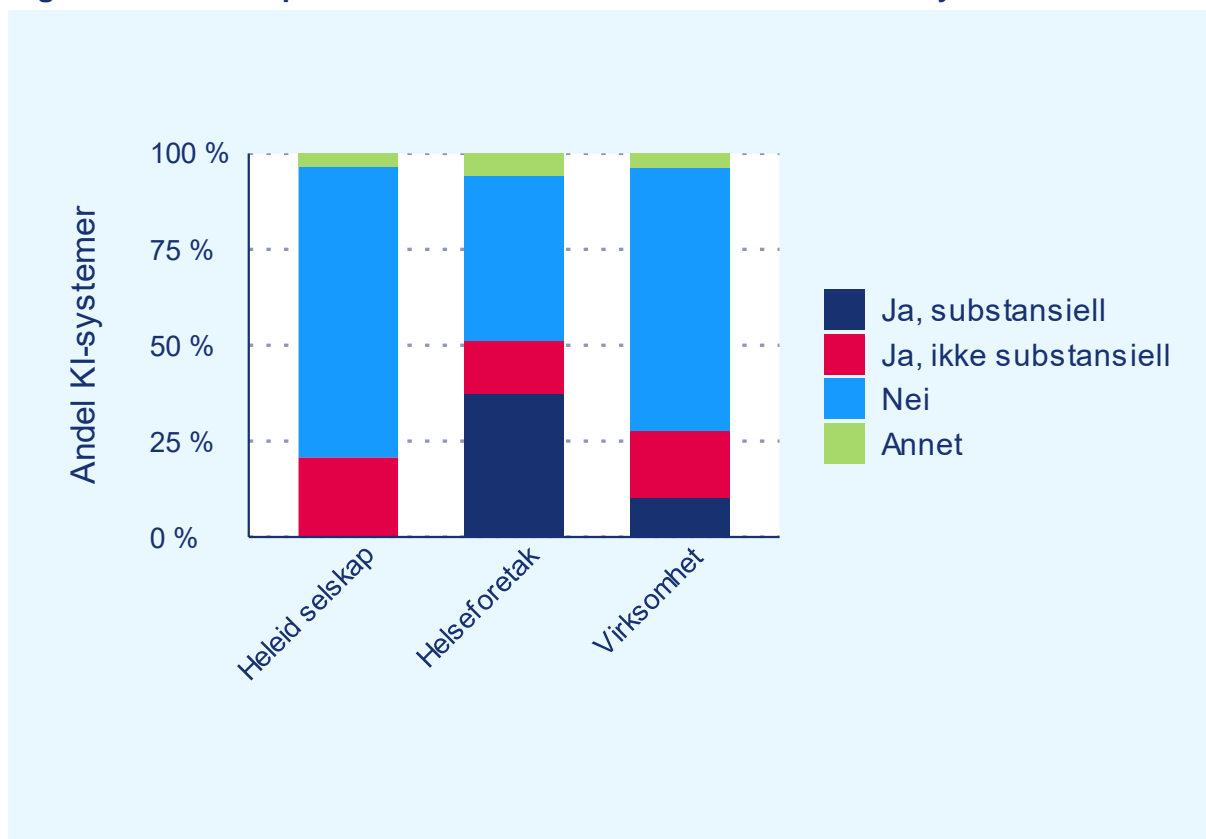
Kilde: Riksrevisjonens spørreundersøkelse (n = 134 KI-systemer).

Spørreundersøkelsen inneholdt også et spørsmål om respondentene mente at enkeltpersoner kunne bli direkte berørt av resultatet fra KI-systemet. Konsekvensene for enkeltpersoner vurderes som betydelige (*substansielle*) dersom KI-systemets utfall kan føre til økonomiske, helsemessige eller rettslige konsekvenser eller til tiltak som berørte personer kan oppfatte som inngripende.

En tredjedel (33 prosent) av alle KI-systemene har ifølge respondentene konsekvenser for enkeltpersoner. Omtrent halvparten av KI-systemene som har konsekvenser for enkeltpersoner, kan ha konsekvenser for enkeltpersoner som er betydelige. Ikke overraskende er det helsesektoren som har størst andel av KI-systemer der enkeltpersoner kan bli direkte berørt av resultatet, jf. figur 10¹²⁵. Men det er også eksempler på slike systemer andre steder. Dette tyder på at KI blir brukt på områder som har direkte betydning for befolkningen, og at ansvarlig bruk av KI er viktig.

¹²⁵ Annet-kategorien i figur 10 ble delvis kommentert med at KI-systemet er under utvikling og ikke enda har konsekvenser, at konsekvensene kan være ulike, at KI-systemet brukes som beslutningsstøtte og menneskelig beslutning er det som har konsekvenser, eller usikkerhet hvor vidt enkeltpersonforetak kan regnes som personer.

Figur 10 Kan enkeltpersoner bli direkte berørt av resultatet fra KI-systemet?



Kilde: Riksrevisjonens spørreundersøkelse (n = 189 KI-systemer).

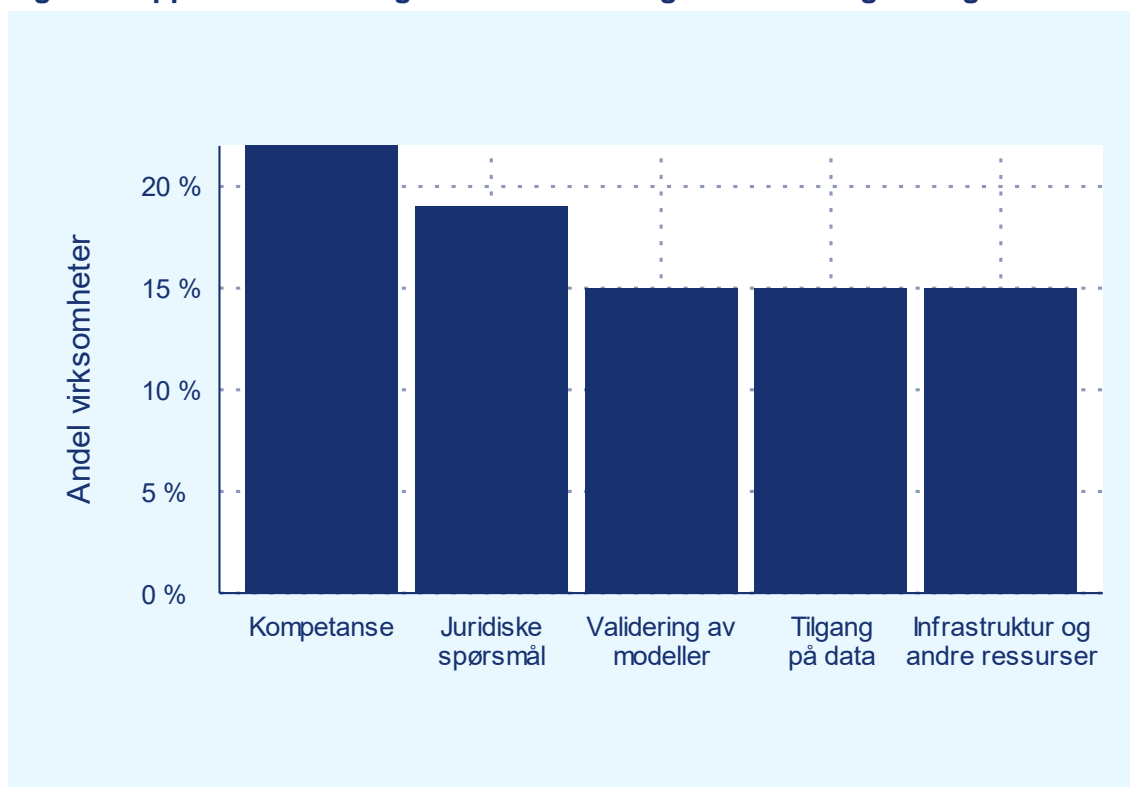
5 Utfordringer med å utvikle og ta i bruk kunstig intelligens

I brev, spørreundersøkelsen og intervjuene ble det stilt spørsmål om eventuelle utfordringer eller hindringer i arbeidet med å utvikle og å ta i bruk kunstig intelligens. I spørreundersøkelsen oppga respondentene i tillegg konsekvensene av utfordringene.

Svarene som pekte seg ut i brevene er oppsummert i Figur 11 nedenfor. Flere av utfordringene må anses som velkjente og er blant annet omtalt i den nasjonale strategien for kunstig intelligens. Flere av hovedutfordringene er også omtalt i Riksrevisjonens Dokument 3:8 (2022–2023) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen*.¹²⁶

Hvordan Digitaliserings- og forvaltningsdepartementet og andre departementer har arbeidet med noen av disse utfordringene, er også berørt i dette kapitlet. Digitaliserings- og forvaltningsdepartementets samordning av og tilrettelegging og pådriverrolle for bruk av KI i staten er imidlertid omtalt i større bredde i kapittel 7 (jf. problemstilling 3).

Figur 11 Opplevde utfordringer med å utvikle og bruke kunstig intelligens. Prosent



Kilde: Svar på brev til statlige virksomheter og heleide statlige selskaper (n = 130).

Figur 11 viser at tilgangen på kompetanse og usikkerhet rundt hjemmels- og behandlingsgrunnlag til å sammenstille og behandle data (særlig knyttet til personopplysninger) blir trukket frem som noen av de sentrale utfordringene. En annen utfordring er tilgang på data av god kvalitet. For flere er det også vanskelig å teste om KI-modeller fungerer, gjennom eksisterende datagrunnlag (validere modeller). I tillegg nevner flere av virksomhetene at de ikke har ressurser til å kjøpe inn nødvendige tekniske løsninger, infrastruktur og lignende. Ressursmangel kan også bety at de ikke kan avse nok ansatte til

¹²⁶ Vi har justert noe på kategoriene sammenlignet med oversikten over utfordringer som ble presentert i Dokument 3:8.

å utvikle og eventuelt vedlikeholde KI-løsninger. Basert på intervjuene, svarene i brevene og spørreundersøkelsen kan de opplevde utfordringene med utvikling og bruk av kunstig intelligens avhenge av hvor langt virksomhetene har kommet i arbeidet med å utvikle KI-prosjekter, hvor avhengig de er av personsensitive data, og hvor avanserte modeller virksomheten arbeider med.

5.1 Mange opplever at regelverk og personvern er vanskelig i arbeidet med kunstig intelligens

I Norge har vi i stor grad et teknologinøytralt regelverk. Det innebærer blant annet at offentlig sektor står fritt til å velge hvilken teknologi som passer best til å løse oppgavene, og at regelverket ikke skal være et hinder for bruk av teknologi. Et teknologinøytralt regelverk kan dermed også bidra til innovasjon. I Norge finnes det heller ingen spesifikk regulering av kunstig intelligens. EU kommer imidlertid med en forordning om regulering av kunstig intelligens (KI-forordningen), som ble vedtatt av EU-parlamentet i mars 2024. Når forordningen er endelig vedtatt i EU, må Norge i henhold til EØS-avtalen ta stilling til om forordningen skal gjelde som norsk rett. Digitaliserings- og forvaltningsdepartementet, Digitaliseringsdirektoratet og Datatilsynet følger EUs arbeid, særlig med hensyn til hvordan KI-forordningen vil påvirke norsk regelverk, og hvordan det kan påvirke norsk forvaltning.

5.1.1 Hjemmelsgrunnlag for å behandle personopplysninger

Utvikling og bruk av kunstig intelligens fordrer ofte bruk av personopplysninger. Bruk av personopplysninger er regulert gjennom EUs personvernforordning (GDPR), og i Norge gjennomfører personopplysningsloven personvernforordningen. Personopplysningsloven stiller flere krav til virksomheter som tar i bruk kunstig intelligens. Regelverket krever at det tas hensyn til personvern i utvikling og bruk av KI-systemer (jf. forordningen artikkel 25). Ifølge personopplysningsloven må all behandling av personopplysninger også ha et rettslig grunnlag for å være lovlig. Virksomhetene må derfor ha et behandlingsgrunnlag **før** opplysninger om personer hentes inn. Hvis det ikke finnes et behandlingsgrunnlag, er bruken av personopplysningene ulovlig (jf. forordningen artikkel 6).

I brevene og i spørreundersøkelsen peker de statlige virksomhetene og de heleide selskapene på flere og sammensatte faktorer ved personvernregelverket som kan komplisere arbeidet med å utvikle og å ta i bruk kunstig intelligens. De opplevde utfordringene synes å være delvis betinget av hvor langt virksomhetene og selskapene har kommet i arbeidet med kunstig intelligens. Jo mer inngripende behandlingen av personopplysningene er, desto strengere kan kravene være til en lovmessig behandling av opplysningene.¹²⁷

Flere av de statlige virksomhetene mener de mangler, eller er usikre på om de har, tydelige nok hjemler til å kunne sammenstille og behandle data som inneholder personopplysninger til bruk i KI-systemer, altså om de i utgangspunktet har et behandlingsgrunnlag i henhold til kravene i personopplysningsloven og personvernforordningen. I spørreundersøkelsen svarte rundt 40 prosent av respondentene at det er vanskelig å ivareta personvern hensyn i arbeidet med å utvikle systemer for kunstig intelligens. Det fremgår av intervjuene, svarene på brev og spørreundersøkelsen at dette kan være krevende avveininger for den enkelte virksomhet. For eksempel viser Finansdepartementet i intervju til at det er krevende å bygge opp den nødvendige kompetansen for å være på trygg juridisk grunn når teknologi som kunstig intelligens skal tas i bruk. I svarene på spørreundersøkelsen oppgir mange av virksomhetene også at det var vanskelig å få avklart om de kunne bruke persondata i det aktuelle KI-systemet. I flere av prosjektene som respondentene i spørreundersøkelsen har opplyst om,

¹²⁷ Følger av personvernforordningen og personopplysningsloven, jf. etterfølgende omtale.

opplevde virksomhetene at arbeidet med slike avveininger resulterte i forsinkelser og til dels høyere kostnader.

Flere departementer har imidlertid i senere år særlig arbeidet med denne problemstillingen ved å innarbeide hjemler i sektorlovgivningen som gir underliggende virksomheter et tydeligere rettslig grunnlag til å kunne behandle personopplysninger. Det gjelder for eksempel Helse- og omsorgsdepartementet og Finansdepartementet. Finansdepartementet og Helse- og omsorgsdepartementet viser til, i henholdsvis intervju og brev, at de har arbeidet med å få tydelige hjemler til sine sektorer. Finansdepartementet viser til at arbeidet har pågått over lang tid, og at det har skjedd i dialog med de underliggende virksomhetene, blant annet for å få synliggjort hvilke områder av regelverket som var uklart, og hvilke behov for hjemler virksomhetene hadde. Begge departementene understreker at teknologinøytralitet ble lagt til grunn i arbeidet med hjemlene, og hjemlene er derfor ikke utarbeidet med sikte på å kun legge til rette for kunstig intelligens. Et teknologinøytralt regelverk gir fleksibilitet og rom for nye tekniske løsninger. Helse- og omsorgsdepartementet viser videre til at det gjeldende regelverket som stiller krav til blant annet personvern og forsvarlighet, også gjelder der et KI-system inngår som en del av helsehjelpen. Helse- og omsorgsdepartementet viser til at i tilfeller der KI-systemet er en del av eller er et nødvendig tilbehør til medisinsk utstyr, vil lov om medisinsk utstyr gjelde. Det innebærer at medisinsk utstyr skal utprøves og brukes på en faglig og etisk forsvarlig måte (jf. lov om medisinsk utstyr § 2).

Det finnes også eksempler på enkeltvirksomheter som har jobbet med å få et tydeligere rettslig grunnlag for å behandle personopplysninger. Et slikt eksempel er Statens pensjonskasse, som tok initiativ overfor Arbeids- og inkluderingsdepartementet til å få nødvendige hjemler på plass da EUs personvernforordning ble vedtatt og inntatt i norsk lov i 2018. Statens pensjonskasse sier i intervju at det har vært en stor fordel å få et slikt hjemmelsgrunnlag tidlig på plass for å kunne satse på digitalisering og kunstig intelligens. Statens pensjonskasse viser samtidig til at det kan finnes en rekke bruksområder for KI som ikke vil ha betydning for personvern, for eksempel intern effektivisering. Statens pensjonskasses erfaring er derfor at man bygger best kompetanse og modenhet i analyse, personvern og andre sentrale hensyn ved å faktisk ta i bruk KI.

I intervjuer med Arbeids- og inkluderingsdepartementet og Arbeids- og velferdsdirektoratet (Nav) vises det til at Nav imidlertid på sin side ikke har et avklart rettslig grunnlag for å behandle personopplysninger til utvikling av KI-systemer (dette omtales også under i omtalen av Navs prosjekt i Datatilsynets sandkasse, se kapittel 7). Arbeids- og inkluderingsdepartementet viser i intervju til at et spesifikt hjemmelsgrunnlag kan medføre at det etableres et dobbeltregelverk – ved at eksisterende regelverket er tilstrekkelig. Departementet er i dialog med Justis- og beredskapsdepartementet om dette og viser til at spørsmålet om særregler og generelle regler blir diskutert i kjernegruppen for et bedre kunnskapssystem for håndtering av kriser. En ny forvaltningslov vil også kunne virke inn på disse spørsmålene, ifølge departementet. Arbeids- og inkluderingsdepartementet viser til at det rettslige grunnlaget er komplisert i en virksomhet som har et stort tilfang av personsensitive data, og legger til at det er Nav selv som må avklare det nødvendige rettslige grunnlaget. Nav fikk på plass en egen enhet for kunstig intelligens i 2017, men Arbeids- og inkluderingsdepartementet og Nav viser i intervju til at per i dag har Nav ennå ingen produksjonssatte løsninger som baserer seg på egenutviklet kunstig intelligens. Nav forvalter omtrent en tredel av statsbudsjettet.¹²⁸

I intervju viser Datatilsynet til at det er klar over problemet med ulik sektorlovgivning, og at et uklart regelverk kan hindre innovasjon. Datatilsynet har imidlertid ingen rolle i å utvikle eller å utfordre et regelverk. Datatilsynet kan tolke regelverket og bistå med å forstå hva sektormyndigheter og andre eventuelt mangler i eget regelverk. Sektormyndighetene må ifølge Datatilsynet selv ta initiativ til å endre regelverk som er til hinder for innovasjon, og samtidig sikre at regelverket ivaretar personvernet.

¹²⁸ Arbeids- og velferdsetaten (NAV), hentet 7. juni 2024 [regjeringen.no](https://www.regjeringen.no)

Datatilsynet understreker at det alltid må finnes et behandlingsgrunnlag for å håndtere personopplysninger, og at dette er et ufravikelig krav.¹²⁹

5.1.2 Innebygd personvern og personvernkonsekvenser

Selv om flere statlige virksomheter og heleide statlige selskaper har tydelige hjemler for å behandle personopplysninger, vil det kunne oppstå ulike spenninger mellom kravene til personvernet og behovet for å bruke personopplysninger for å utvikle og bruke KI-systemer. Personvernregelverket krever for eksempel at det ikke skal behandles flere opplysninger enn nødvendig (prinsippet om dataminimering). Opplysningene som det er ønskelig å bruke i et KI-system, kan i tillegg ha blitt samlet inn til andre formål. Det vil kunne innebære at et KI-system behandler personopplysninger på en måte som er uforenlig med det opprinnelige formålet opplysningene ble innsamlet for. En slik behandling vil da sannsynligvis ikke være lovlig.

Siden bruk av kunstig intelligens ofte kan eller vil innbefatte personopplysninger, må virksomhetene som regel gjøre en grundig vurdering av kravene til personvern i hvert enkelt prosjekt. Sentrale spørsmål som må håndteres, er for eksempel om virksomheten er i ferd med å bruke data på en måte som er uforenlig med grunnlaget for innhenting av opplysningene, og om behandlingen er nødvendig og proporsjonal.

I intervju viser flere til at en stegvis fremgangsmåte er en god mulighet for å imøtekomme de ulike kravene til personvern. Finansdepartementet viser blant annet til at flere av de underliggende etatene prøver seg frem på små avgrensede områder som er trygge, av hensyn til både etiske spørsmål og spørsmål om personvern. Etatene kan da bygge erfaring og utnytte opparbeidet kunnskap på nye områder.¹³⁰ Skatteetaten viser til at interne avklaringsprosesser om juridiske spørsmål kan ta svært lang tid, selv om etaten har gode hjemler for å kunne bruke persondata. Tolletaten legger til i intervju at hvis den ønsker å bruke personopplysninger mer enn i dag, må det gjøres et betydelig arbeid med å vurdere rekkevidden av eksisterende regelverk og behovet for nye hjemler. Statens pensjonskasse viser til at det er viktig å ikke bare se begrensningene – da blir det fort til at man ikke gjør noe som helst. Kompetanse og modenhet blant annet i personvern bygges ved å faktisk ta i bruk KI.

Skatteetaten og Nav etterlyser i intervju en mer overordnet tilnærming til å tolke og bruke regelverket. I dag skjer det meste ad hoc i den enkelte virksomhet. Nav legger til at det er positivt at det er et stort engasjement på grasrota, men skatteetaten og Nav peker på at arbeidet med å utvikle KI ville bli mer effektivt hvis det var en mer overordnet og tverrsektoriell tilnærming til noen sentrale regelverksspørsmål. Akershus universitetssykehus HF viser i intervju på sin side til Helsedirektoratets arbeid med å få på plass hjemler til å drive med forskning innen helsesektoren, og at Helsedirektoratet svarer raskt på overordnede spørsmål om KI og er flinke til å gi nødvendige avklaringer og godkjenninger.

Vurdering av personvernkonsekvenser (DPIA)

Datatilsynet viser til at etter personvernregelverket må personvernet ivaretas i alle utviklingsfaser av et system, dersom utviklingen innebærer behandling av personopplysninger. På samme måte har virksomhetene plikt til å vurdere personvernkonsekvensene for alle løsninger de tar i bruk. Hvis behandlingen gir høy risiko for enkeltpersoners rettigheter og friheter, har virksomheten plikt til å gjennomføre en vurdering av personvernkonsekvenser. En slik vurdering kalles en *Data Protection Impact Assessment* (DPIA), det vil si en vurdering av personvernkonsekvenser. Vurderingen skal sikre personvernet til de registrerte hvis data inngår i et slikt system, for eksempel KI-løsninger. Ifølge Datatilsynets skal en eventuell vurdering av personvernkonsekvenser gjennomføres før databehandlingen settes i gang (jf. artikkel 35). Med utgangspunkt i kravene i artikkel 35 nr. 7 i

¹²⁹ Verifisert referat fra intervju med Datatilsynet 13. oktober 2023.

¹³⁰ Verifisert referat fra intervju med Finansdepartementet 29. november 2023, Statens pensjonskasse 1. desember 2023 og Arbeids- og inkluderingsdepartementet 18. desember 2023.

personvernforordningen har Datatilsynet laget en veileder og en sjekklister¹³¹ for vurdering av personvernkonsekvenser.

Enkelte av de intervjuede virksomhetene viser til at kravene til personvern i hele utviklingsprosessen oppleves som en utfordring. Arbeidet med å utvikle KI-løsninger krever utforskning, eksperimentering og bruk av ulike datakilder – gjerne på tvers i den enkelte virksomheten eller på tvers av ulike virksomheter.¹³² I noen tilfeller er virksomhetene kun ute etter å teste om det er mulig å ta i bruk KI på spesifikke områder. Flere virksomheter og selskaper som arbeider med å utvikle KI-prosjekter, opplever at de juridiske prosessene for å ivareta personvern hensyn i en slik utforskningsfase er for strenge. Videre trekker de intervjuede virksomhetene frem at det blir vanskelig å komme i gang bare med å kartlegge mulighetene for bruk av KI når en rekke juridiske spørsmål må avklares for hvert enkelt utviklingsprosjekt, og når slike prosesser i tillegg tar lang tid.¹³³

I intervju viser blant annet Statens pensjonskasse til at de ønsket å sjekke om det var mulig å bruke KI til å forbedre personvernet i personskadeforsikringsarbeidet. For å teste om KI kunne bidra til en forbedring startet Statens pensjonskasse arbeidet med en vurdering av personvernkonsekvenser¹³⁴. Det tok ifølge Statens pensjonskasse to år å få denne vurderingen på plass, og resultatet er at de aldri gikk videre med dette prosjektet.¹³⁵

Virksomhetene understreker viktigheten av at juristene er med i hele utviklingsprosessen av KI. Som et eksempel sier Nav at selv om det ikke nødvendigvis oppstår noen vanskelige prinsipielle juridiske spørsmål ved bruken av KI i starten av prosessen, er det viktig med gode juridiske avklaringer fra start. Slik sikrer man at de endelige modellene og løsningene er godt juridisk forankret.¹³⁶ Ifølge skatteetaten kan juridiske avklaringsprosesser generelt ta lang tid, og det er ikke nødvendigvis slik at virksomhetene har nok ressurser til å avklare relevante juridiske spørsmål raskt.¹³⁷

Også i helse- og omsorgssektoren fremholder flere av aktørene at den største utfordringen de har møtt på for å kunne ta i bruk KI, er de regulatoriske hindringene som gjelder bruk av helsedata. Det er også her en rekke juridiske utfordringer rundt begrensninger ut fra formålet med innsamling av data og sekundærbruk av disse dataene. Dette fremkommer tydelig i svarene på brevene sendt til helseforetakene.

Virksomheter i helse- og omsorgssektoren trekker frem en annen utfordring i arbeidet med KI-utvikling: Det mangler juridisk hjelp og veiledning til å sette opp gode databehandleravtaler. Alle virksomheter som benytter seg av en underleverandør, har plikt til å ha en databehandleravtale. Den skal sikre at personopplysningene blir behandlet i samsvar med regelverket, og setter en klar ramme for hvordan databehandleren kan behandle opplysningene.¹³⁸

Helse- og omsorgsdepartementet påpeker at det er opprettet en egen tverretattlig veiledningstjeneste i helsesektoren for de som forsker eller utvikler produkter basert på kunstig intelligens, eller som ønsker å ta i bruk utstyr basert på kunstig intelligens. Gjennom tjenesten kan det søkes om en-til-en-veiledning fra jurister fra de ulike etatene i sektoren. Etatene vil på bakgrunn av disse erfaringene kunne videreformidle rettslige utfordringer og eventuelle behov for regelverksendringer til departementet.¹³⁹

¹³¹ Datatilsynets DPIA sjekklister. Lenke: <https://www.datatilsynet.no/contentassets/8b767689abb14926af27820c9c2fb89e/sjekklister-for-dpiafaser.pdf>

¹³² Verifisert referat fra intervju med skatteetaten 17. oktober 2023.

¹³³ Dette er også omtalt i Dokument 3:8 (2023–2024).

¹³⁴ Datatilsynet. (2023, 27. juli). Vurdering av personvernkonsekvenser (DPIA). Hentet fra <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/vurdering-av-personvernkonsekvenser/>

¹³⁵ Verifisert referat fra intervju med Statens pensjonskasse 1. desember 2023.

¹³⁶ Verifisert referat fra intervju med Nav 8. november 2023, Akershus universitetssykehus HF 20. november 2023 og skatteetaten 17. oktober 2023.

¹³⁷ Verifisert referat fra intervju med skatteetaten 17. oktober 2023.

¹³⁸ Datatilsynet. (2023, 27. juli). Databehandleravtale. <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/hvordan-lage-en-databehandleravtale/>

¹³⁹ Brev fra Helse- og omsorgsdepartementets av 29. februar 2024.

5.2 Mangelfull tilgang på data av god kvalitet

Det å utvikle og bruke kunstig intelligens forutsetter ofte bruk av store mengder data. Dataene brukes til å trene opp maskinene, men kan være vel så viktige for å teste og evaluere utfall i KI-modeller som er satt i produksjon. Flere av de statlige virksomhetene og de heleide selskapene trekker frem at tilgang på data av god kvalitet er helt sentralt i arbeidet med kunstig intelligens. Mangel på data av god kvalitet gjør det vanskelig å utvikle og bruke kunstig intelligens. For å kunne utvikle og bruke KI bør virksomheten ha

- data av god kvalitet
- tilgang på riktige data
- tekniske muligheter til å dele data

Prop. 1 S (2021–2022) for Kommunal- og distriktsdepartementet viser til den nasjonale strategien for kunstig intelligens og at tilgangen på data av god kvalitet er ett viktig tiltak i arbeidet med kunstig intelligens.

I svarene på brevene, spørreundersøkelsen og intervjuene viser flere av de statlige virksomhetene og heleide statlige selskapene til at manglende tilgang på data av god kvalitet er et hinder i arbeidet med kunstig intelligens. I spørreundersøkelsen svarer ca. halvparten av respondentene at manglende tilgang på data og utilstrekkelig datakvalitet har vært en utfordring i arbeidet med å utvikle og å ta i bruk kunstig intelligens. Konsekvensen for de aktuelle prosjektene har vært forsinkelser i arbeidet, forringet kvalitet i løsningen og økte kostnader. Gjennom svarene kommer det også frem at flere prosjekter ble stanset som en følge av mangel på data og mangelfull datakvalitet.

Tilgang på data av god kvalitet som grunnlag for kunstig intelligens kan gjelde vel så mye tilgang på egne data av god kvalitet som data fra andre offentlige og private virksomheter. Virksomhetenes og selskapenes modenhet og ambisjoner for bruk av kunstig intelligens kan også påvirke om de har vanskeligheter med å få tilgang på data av god kvalitet.

Behovet for deling og tilgang på data av god kvalitet har over flere år vært løftet frem som viktig i arbeidet med digitalisering. Digitaliseringsdirektoratet utarbeidet blant annet en veileder (2018) om å sikre god kvalitet på virksomhetenes egne data gjennom prinsippet «orden i eget hus».¹⁴⁰ God datakvalitet henger sammen med målene om å utnytte de store mengdene dataene som eksisterer til det beste for samfunnet (jf. Meld. St. 22 *Data som ressurs – Datadrevet økonomi og innovasjon* (2020–2021)). God datakvalitet handler om sikre at dataene som samles, lagres og brukes innenfor en organisasjon eller et system, er nøyaktige, pålitelige, fullstendige og relevante for formålet de skal brukes til. Prinsippet om orden i eget hus er pålagt statlige virksomheter gjennom Digitaliseringsrundskrivnet og er løftet frem i den nasjonale strategien for kunstig intelligens. Orden i eget hus er derfor et nødvendig *utgangspunkt* for at offentlig og privat sektor skal kunne utnytte potensialet i data til å skape nye produkter, gi innovativ tjenesteutvikling og en effektiv forvaltning¹⁴¹. Både Finansdepartementet og Arbeids- og inkluderingsdepartementet legger i intervju vekt på at underliggende virksomheter må sikre god kvalitet på egne data.

Digitaliseringsdirektoratet viser imidlertid til at en tredjedel av offentlige virksomheter opplever at datakvalitet er et hinder for utvikling av digitale tjenester.¹⁴² Enkelte av de intervjuede virksomhetene i undersøkelsen påpeker det samme og at de selv har et potensial for å sikre bedre kvalitet på egne data.¹⁴³

¹⁴⁰ Digitaliseringsdirektoratet. (2018). *Veileder for orden i eget hus*. Hentet 12. februar 2024 fra <https://www.digdir.no/informasjonsforvaltning/veileder-orden-i-eget-hus/2716>

¹⁴¹ Digdirs nettside: <https://www.digdir.no/informasjonsforvaltning/hva-er-orden-i-eget-hus/2792> Hentet: 7. februar 2024.

¹⁴² Digitaliseringsdirektoratets årsrapport for 2022.

¹⁴³ Verifisert referat fra intervju med Nav 8. november 2023 og Akershus universitetssykehus HF 20. november 2023.

5.2.1 Deling av data

Tilgang på data av god kvalitet kan gjelde deling av data med andre virksomheter og sektorer. I Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen, jf. Dokument 3:8 (2023–2024), ble det pekt på at potensialet for deling av data er langt større enn det som utnyttes. I 2016 ble den offentlige tjenesten Felles datakatalog etablert. Den skulle bidra til å samle og tilgjengeliggjøre metadata om offentlige datakilder i Norge.

Digitaliseringsdirektoratet er ansvarlig for drift og utvikling av nettstedet. Felles datakatalog skal ha en oversikt over hvilke data de ulike offentlige virksomhetene har, hvordan de henger sammen, og hva de betyr. Datakatalogen skal gjøre det mulig å søke i opplysningene hos offentlige virksomheter og bruke dem til for eksempel analyseformål. Ifølge Digitaliseringsrundskrivet¹⁴⁴ skal statlige virksomheter registrere informasjon om sine datasett i Felles datakatalog. Denne registreringen skal virksomhetene som et minimum gjøre når de endrer eller etablerer tjenester.¹⁴⁵

Den nasjonale KI-strategien sier at regjeringen vil at «det skal legges til rette for deling av data innenfor og på tvers av bransjer og sektorer». Digitaliserings- og forvaltningsdepartementet viser i intervju til at dataenes kvalitet, så vel som tilgjengelighet, er viktig for om dataene kan brukes til KI.¹⁴⁶ Digitaliserings- og forvaltningsdepartementet viser til at Felles datakatalog er et viktig virkemiddel for å få oversikt over tilgjengelige data, men peker på at det er virksomhetene som legger ut informasjon om datasett fra Felles datakatalog, som selv må sørge for at dataene er gode nok.

I tillegg til Felles datakatalog opprettet Digitaliseringsdirektoratet et nasjonalt ressurscenter for deling av data i september 2020. Senteret skal fremme deling og bruk av data ved å spre kunnskap og utvikle regelverket på området. Senteret veileder virksomheter blant annet om juridiske spørsmål om personvern- eller forvaltningsregelverket og sammenhengen mellom disse regelverkene. Ressurscenteret er en kompetanseressurs for juridiske spørsmål om digitalisering og har også fagfolk med bakgrunn i informasjonsforvaltning som arbeider med andre spørsmål.¹⁴⁷ I Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen* trekker flere virksomheter frem at Nasjonalt ressurscenter for deling og bruk av data er et viktig virkemiddel i arbeidet med å få til juridiske avklaringer. Digitaliseringsdirektoratet uttaler imidlertid at virksomhetene i varierende grad bruker de tilgjengelige verktøyene.¹⁴⁸

På spørsmål til Digitaliserings- og forvaltningsdepartementet om den konkrete oppfølgingen av tiltakene i den nasjonale strategien for kunstig intelligens som er rettet mot datadeling og datakvalitet, peker departementet på en oversikt som viser at mange av de aktuelle tiltakene skissert i KI-strategien er gjennomført.¹⁴⁹ Departementet viser til at Digitaliseringsdirektoratet har utviklet et godt grunnlag for datadeling av åpne data i offentlig sektor gjennom Felles datakatalog. Direktoratet har etablert det ovennevnte ressurscenteret for deling og bruk av data¹⁵⁰, laget en Nasjonal verktøykasse for deling av data¹⁵¹, utformet et rammeverk for informasjonsforvaltning¹⁵² og utarbeidet en veiledning for hvordan man registrerer datasett i Felles datakatalog¹⁵³. I Dokument 3:8 (2023–2024) om Riksrevisjonens undersøkelse om deling og gjenbruk av data er det påvist at mange virksomheter bruker verktøyene som Digitaliseringsdirektoratet har utarbeidet, og synes at de er nyttige.

I årsrapporten for 2022 viser Digitaliseringsdirektoratet til at det ved utgangen av 2022 var registrert 1 658 datasett i Felles datakatalog, som var en økning på 83 fra 2021. Likevel er økningen mindre enn

¹⁴⁴ Kommunal- og distriktsdepartementet. (2022). *Digitaliseringsrundskrivet H-5/21*.

¹⁴⁵ Riksrevisjonen. (2023). *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen. Dokument 3:8 (2023–2024)*. <https://www.riksrevisjonen.no/globalassets/rapporter/no-2023-2024/myndighetenes-tilrettelegging-for-deling-og-gjenbruk-av-data-i-forvaltningen.pdf>

¹⁴⁶ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

¹⁴⁷ Verifisert referat fra intervju med Digitaliseringsdirektoratet 20. september 2023.

¹⁴⁸ <https://www.riksrevisjonen.no/globalassets/rapporter/no-2023-2024/myndighetenes-tilrettelegging-for-deling-og-gjenbruk-av-data-i-forvaltningen.pdf>

¹⁴⁹ Mottatt dokumentasjon fra Digitaliserings- og forvaltningsdepartementet 13. november 2023.

¹⁵⁰ Lenke: <https://www.digdir.no/datadeling/nasjonalt-ressurscenter-deling-og-bruk-av-data/1914> Hentet 24. mars 2024

¹⁵¹ Lenke: <https://www.digdir.no/datadeling/nasjonalt-ressurscenter-deling-og-bruk-av-data/2243> Hentet 11. mars 2024

¹⁵² Lenke: <https://www.digdir.no/informasjonsforvaltning/rammeverk-informasjonsforvaltning/3626> Hentet 11. mars 2024

¹⁵³ Lenke: <https://data.norge.no/about-registration> Hentet 19. april 2024

de foregående årene. Digitaliseringsdirektoratet viser til at det verdt å bemerke at kun 64 statlige virksomheter har datasett i datakatalogen, og at dette anslagsvis svarer til om lag en fjerdedel av de virksomhetene som er omfattet av Digitaliseringsrundskrivets krav om deling og videre bruk av informasjon. Digitaliseringsdirektoratet viser også til at flere offentlige virksomheter opplever at manglende kvalitet på tilgjengelige data er en utfordring for gjenbruk av data fra andre offentlige virksomheter.¹⁵⁴ Det fremgår av Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse om deling og gjenbruk av data* at det er mange virksomheter som ikke benytter Felles datakatalog, og det blir pekt på at det er et potensial for betydelig mer deling og gjenbruk av data.

Mange av virksomhetene forvalter i tillegg betydelige mengder personopplysninger og det er derfor store mengder data som ikke kan deles uten videre eller som kan tilgjengeliggjøres gjennom ulike åpne løsninger.

Basert på svar fra brevene, spørreundersøkelsen og intervjuer er det enkelte store virksomheter og selskaper som trekker frem at deling av data på tvers av virksomheter og selskaper er en hindring i å utvikle og å ta i bruk kunstig intelligens. Eksempelvis påpeker Nav at de har jobbet mye med å få til avtaler om deling av data med andre virksomheter. Dataene fra andre er ikke nødvendigvis på en form, i et format og av en kvalitet som gjør dem anvendelige i Navs arbeid. Eksempelvis Statistisk sentralbyrå (SSB) kan ifølge Nav ha relevante data, men det er vanskelig å få detaljerte nok data fra SSB tilpasset Navs behov. Nav sender på lik linje som andre offentlige virksomheter data til SSB, men det er ikke slik at Nav kan få andres data eller sine egne bearbejdede data tilbake igjen fra SSB.

I intervjuer er det også flere som uttrykker ønske om at stat og kommune deler mer data mellom seg. Finansdepartementet viser i intervju til at flere kommuner eksempelvis har uttrykt ønske om å få inntektsdata fra skatteetaten automatisk inn i sine fagsystemer. Kommunene er avhengig av inntektsdata for å levere et tjuetalls tjenester, og med direkte data fra skatteetaten kan for eksempel kommunene enklere produsere automatiske vedtak for hvor mye en bruker skal betale for en barnehageplass.

Nav peker på at de har full oversikt over *statlige* sosialhjelpsytelser, men har ikke tilgang til informasjon om hvilke sosialhjelpsytelser brukerne får fra den enkelte kommunen. Statlige Nav må vente på den årlige KOSTRA-rapporteringen for å få tilgang til denne informasjonen.¹⁵⁵ Deling av åpne data mellom stat og kommune er uproblematisk. Deling av personsensitive data generelt og også mellom stat og kommune og mellom kommuner er et spørsmål som må avklares enten på et generelt grunnlag eller i enkeltsaker.¹⁵⁶ Statlige virksomheter og kommunene må imidlertid selv sørge for at de har nødvendig lovhjemmel for å få tilgang til aktuelle opplysninger fra andre virksomheter og/eller kommuner.

Digitaliserings- og forvaltningsdepartementet sier at det ikke er kjent med at det har vært større utfordringer med å få tilgang til andre virksomheters data i forbindelse med kunstig intelligens. Det kan imidlertid kreve lovendringer for å få tilgang til data eller å bruke dataene til andre formål. Digitaliserings- og forvaltningsdepartementet viser videre til at det kan være utfordrende å avveie minimeringsprinsippet i personvernregelverket mot behovet for å trene KI-modeller.¹⁵⁷ Prinsippet innebærer at organisasjoner ikke bør samle inn mer data enn det som er nødvendig for å oppfylle det angitte formålet, og de bør heller ikke oppbevare dataene lenger enn nødvendig. Det står imidlertid ifølge departementet i kontrast til arbeidet med kunstig intelligens, der tilgang på et størst mulig datagrunnlag ofte er en forutsetning for å lykkes.

¹⁵⁴ Digitaliseringsdirektoratet. (2023). Årsrapport 2022. <https://aarsrapport2022.digdir.no/sites/aarsrapport2021/files/2023-04/Digdirs%20%C3%A5rsrapport%202022.pdf>

¹⁵⁵ Verifisert referat fra intervju med Nav 8. november 2023.

¹⁵⁶ Verifisert referat fra intervjuer med Finansdepartementet 29. november 2023 og Arbeids- og integreringsdepartementet 18. desember 2023.

¹⁵⁷ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

5.2.2 Arbeidet med å tilgjengeliggjøre helsedata

I den nasjonale KI-strategien ble det vist til at regjeringen ville etablere en nasjonal løsning for å tilgjengeliggjøre helsedata til forskning og annen sekundærbruk. Denne løsningen, Helseanalyseplattformen, skulle gi mulighet for mer avanserte analyser av norske helsedata og legge grunnlaget for nye typer medisinsk og helsefaglig forskning.

Arbeidet med å utvikle Helseanalyseplattformen i en allmenn skyløsning ble imidlertid avvirket som følge av endringer i juridiske rammebetingelser etter Schrems II-dommen fra EU-domstolen i 2021. Ifølge Prop. 1 S (2023–2024) for Helse- og omsorgsdepartementet er det lagt en ny og alternativ strategi for å forenkle tilgangen på helsedata basert på mer effektiv utnyttelse av eksisterende data- og analysetjenester i universitets- og høyskolesektoren og samlet organisering av de nasjonale helseregistrene i Folkehelseinstituttet.

Gjennom svar på brevene og i intervju viser flere sykehus/helseforetak til at datatilgjengeligheten hindres av adskilte fagsystemer uten koblingsmulighet til andre datakilder. Det er blant annet mangelfull kommunikasjon mellom pasientjournalen og andre systemer som leger registrerer data i. Konsekvensen er at registrerte data om den enkelte pasienten ikke lagres eller gjøres tilgjengelig på ett sted.¹⁵⁸ Intern datatilgang og datautveksling med andre helseforetak blir dermed også et problem. Videre trekkes ikke-digitaliserte data frem som en stor hindring.¹⁵⁹ Det er også mangel på deling av data mellom spesialisthelsetjenesten og primærhelsetjenesten som kan påvirke mulighetene for å utvikle KI-systemer – inkludert tilgang på resultater fra blodprøver analysert hos private aktører. Problemene med å dele data på tvers av de ulike systemene forsterkes når det innføres nye medisinsk-tekniske løsninger. Da introduseres også ofte ny programvare som ikke uten videre kan kobles til eksisterende systemer.¹⁶⁰

Helse- og omsorgsdepartementet sier at i helsesammenheng brukes begrepet «primærbruk» om deling av data for helsehjelp og «sekundærbruk» om bruk av data til for eksempel forskning, styring og innovasjon. Departementet har fått tilbakemelding om at det er krevende å dele data for helsehjelp mellom virksomheter. Dessuten er det utfordringer med dobbeltrapportering av helsedata til sekundærbruk. Det betyr at helsepersonell bruker mer tid, og at de rapporterte dataene har dårligere kvalitet. Utlevering av helsedata fra de nasjonale helseregistrene går via Helsedataservice ved Folkehelseinstituttet. Fra 2024 er de nasjonale helseregistrene samlet i Folkehelseinstituttet, og et mål er at dette skal gi kortere utleveringstid på søknader om data og modernisering av helsedataområdet.¹⁶¹

Mangelfull tilgang på data og mangelfull kvalitet på dataene gjør det vanskelig å trene opp egne KI-systemer og å kunne validere og teste egne og innkjøpte systemer. Som vist tidligere har flere sykehus/helseforetak relativ høy aktivitet innenfor kunstig intelligens. I helsesektoren er det mange eksterne leverandører som tilbyr ferdige KI-systemer som hylleware. Helseforetakene påpeker at ved kjøp av ferdige KI-systemer kreves det omfattende testing og validering før modellen kan tas i bruk. Dermed må sykehuset validere om modellen fungerer korrekt ut fra sykehusets egne erfaringer og data. Grunnlagsdata og treningsdata som er brukt i utvikling og trening av selve KI-modellen, er ofte ikke basert på norske pasientdata. Derfor må innkjøpte KI-systemer gjennom omfattende testing på data for den befolkningsgruppen den skal behandle, for å kontrollere at KI-systemets treffsikkerhet i oppgaveløsningen samsvarer med det presisjonsnivået leverandøren oppgir. Flere av sykehusene/helseforetakene oppgir at validering av KI-modeller dermed kan innebære et omfattende arbeid og også en løpende kvalitetskontroll.

¹⁵⁸ Verifisert referat fra intervju med Akershus universitetssykehus HF 20. november 2023.

¹⁵⁹ Verifisert referat fra intervju med Akershus universitetssykehus HF 20. november 2023.

¹⁶⁰ Verifisert referat fra intervju med Akershus universitetssykehus HF 20. november 2023.

¹⁶¹ Brev fra Helse- og omsorgsdepartementet av 29. februar 2024.

5.3 Behov for infrastruktur og andre tekniske ressurser

I den nasjonale KI-strategien vises det til at utvikling og bruk av KI krever god kommunikasjonsinfrastruktur og tilgang til regnekraft. Arbeidet med kommunikasjonsinfrastruktur, og særlig 5G-nett, er et prioritert område. Tilgang til tilstrekkelig regnekraft skal sikres gjennom bruk av nasjonale og internasjonale ressurser for tungregning. På den måten er infrastruktur en annen viktig brikke for utvikling av kunstig intelligens.

I brev svarte flere statlige virksomheter og heleide selskaper at mangel på ulike former for ressurser er en utfordring i arbeidet med kunstig intelligens. Flere peker på mangel på maskinkraft og intern digital infrastruktur som et problem. Det gjelder for eksempel lagring av store datamengder, overføring og kommunikasjon av data og maskinenes analysekapasitet. I spørreundersøkelsen svarer i underkant av 50 prosent av informantene (av 104) at deres digitale infrastruktur og/eller IT-løsninger ikke er tilpasset KI. Konsekvensen for de aktuelle KI-prosjektene har vært forsinkelser i arbeidet, forringet kvalitet i løsningen og høyere kostnader.

Det kan være særlig kostbart å investere i maskiner som kan lagre og håndtere store datamengder. I spørreundersøkelsen svarer ca. 25 prosent av respondentene at finansiering har vært en utfordring. Særlig i helsesektoren trekkes det også frem at manglende infrastruktur mellom systemer på sykehusene er en utfordring i arbeidet med KI. I helsesektoren kan det også være et ressurs spørsmål om man kan kjøpe fri klinisk helsepersonell i arbeidet med å teste og utvikle systemer.

I intervju utdyper flere virksomheter hvor viktig det er å ha tilgangen på moderne dataplattformer for å utvikle og bruke KI. Flere har pågående utviklingsprogram som skal modernisere datavarehus og analyseplattformer og jobber med skyløsninger som vil gi nye muligheter. Per nå er mangel på slike løsninger en hindring i arbeidet med KI.^{162,163}

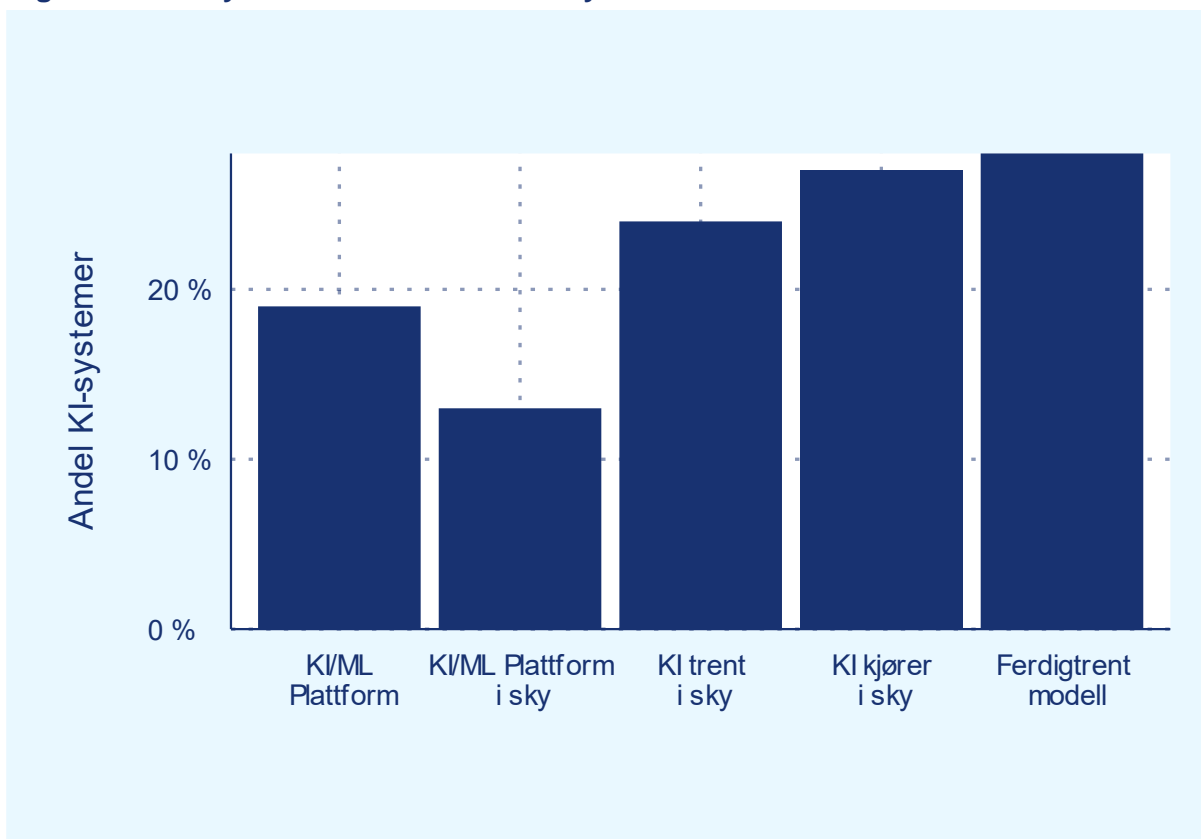
Vi ser også fra spørreundersøkelsen at det er relativt stor variasjon i den tekniske infrastrukturen, jf. Figur 12 nedenfor. For eksempel bruker 28 prosent av KI-systemene en ferdigtrent modell, og i underkant av hvert femte KI-system ble utviklet med hjelp av en kunstig intelligens eller maskinlæringsplattform. Totalt er det omtrent 25 prosent av KI-systemene som er utviklet og/eller kjøper i en sky. Kun 13 prosent benytter skybasert kunstig intelligens eller maskinlæringsplattform.

Det er klart at ulike systemer vil ha ulike krav til infrastruktur. Enkle, regelbaserte KI-systemer vil ikke nødvendigvis ha behov for en skybasert KI-plattform å kjøre på. Mer avanserte maskinlæringsmodeller er det vanskelig å utvikle og kjøre uten tilgang til et skymiljø: Cirka 55 prosent av respondentene i spørreundersøkelsen som bruker avansert maskinlæring, viser til at det er en utfordring at deres *digitale infrastruktur / IT-løsninger ikke er ikke tilpasset KI*. Blant respondentene uten avansert maskinlæring er det 41 prosent som opplever at infrastrukturen er en utfordring.

¹⁶² Verifisert referat fra intervju med skatteetaten 17. oktober 2023 og Statens pensjonskasse 1. desember 2023.

¹⁶³ Verifisert referat fra intervju med tolletaten 24. oktober 2023.

Figur 12 Verktøy/infrastruktur brukt i KI-systemene



Kilde: Riksrevisjonens spørreundersøkelse (n = 189 KI-systemer).

Flere av virksomhetene problematiserer at det ikke er en felles skyløsning i offentlig sektor – forskningsinstitusjonen NIBIO uttrykker det slik: For at de skal kunne utnytte KI til flere nasjonale oppgaver, er en koordinert nasjonal oppbygging av en felles infrastruktur for lagring og beregning en viktig forutsetning. Det er store synergier å hente gjennom mer samhandling på tvers når det gjelder infrastruktur¹⁶⁴. Statens pensjonskasse viser videre til at innkjøp av analyseplattformer er komplisert og en stor investering for den enkelte virksomheten, særlig mindre virksomheter. Offentlig sektor kunne ha effektivisert KI-arbeidet ved at infrastrukturen og analyseplattformene i større grad ble driftet sentralt. Det hadde også gitt et løft i den teknologiske modenheten i hele offentlig sektor, og nødvendig infrastruktur hadde blitt tilgjengelig for flere, ifølge Statens pensjonskasse. I dag er det mange offentlige virksomheter som ikke en gang har noen analyseplattform. Da er det langt frem til å begynne å utvikle og ikke minst å ta i bruk avansert teknologi som KI.¹⁶⁵ En annen virksomhet peker videre på at virksomheter i samme sektor som samhandler mye, vil kunne velge ulike skyløsninger (Nav og Statens pensjonskasse), noe som kan bidra til at datadeling vanskeliggjøres.¹⁶⁶

Nasjonalsikkerhetsmyndighets konseptvalgutredning for nasjonal skytjeneste peker på at det for noen datatyper og IT-systemer er sårbart å være avhengig av utenlandske skytjenester, og anbefaler overfor Justis- og beredskapsdepartementet etablering av en nasjonal skytjeneste for Norge.¹⁶⁷ Det er imidlertid uenighet om hva som er den beste fremgangsmåten for etablering av en nasjonal skyløsning, blant annet fordi flere statlige virksomheter allerede har investert i kommersielle skytjenester.¹⁶⁸

¹⁶⁴ Se også Skates erfaringsrapport for 2023.

¹⁶⁵ Verifisert referat fra intervju med Statens pensjonskasse 1. desember 2023.

¹⁶⁶ Verifisert referat fra intervju med Statens pensjonskasse 1. desember 2023.

¹⁶⁷ NSM (2023) *Konseptvalgutredning for nasjonal skytjeneste*. <https://nsm.no/regelverk-og-hjelp/rapporter/konseptvalgutredning-for-nasjonal-skytjeneste>

¹⁶⁸ Skate, erfaringsrapport 2023

I svarene på brevene trekker helsesektoren frem at det oppleves ressurskrevende å kjøpe medisinsk teknisk utstyr (MTU) og datautstyr. Andre barrierer er for eksempel at det trengs høykapasitetslinjer, stor lagrings- og prosesseringsplass i bildebehandlingssystemene og kraftig maskinvare for å kunne kjøre undersøkelser ved bruk av KI-systemer.

Helse- og omsorgsdepartementet viser til at Norsk helsenett SF er en nasjonal tjenesteleverandør av stamnettet, den digitale infrastrukturen som kobler alle sykehusene sammen. Gjennom stamnettet får helsepersonell tilgang til sentrale e-helseløsninger, slik som kjernejournal, samt tjenester fra Helfo og Nav. Videre er stamnettet et robust høykapasitetsnettverk spesialtilpasset helsesektorens behov for å kunne dele store datamengder på en sikker måte. Norsk helsenett bygger og administrerer Helsenettet, som er et medlemsnettverk for samhandling i helsetjenesten. Her kan aktørene i helse- og omsorgstjenestene dele pasient- og personopplysninger innenfor trygge og lovlige rammer.¹⁶⁹

Digitaliserings- og forvaltningsdepartementet viser i intervju til Sigma2 som en felles tungregningsinfrastruktur. Sigma2 tilbyr høytytende datamaskinressurser (HPC) og storskala datalagring til forskere i Norge. Departementet peker på at ambisjonen er at denne løsningen i større grad skal kunne brukes også av andre sektorer enn forsknings- og utdanningssektoren.¹⁷⁰

Norge bidrar med betydelige midler til EU gjennom medlemskap i EuroHPC JU (The European High Performance Computing Joint Undertaking), som blant annet har finansiert tungregningsmaskinen «LUMI» i Finland, der forskningsinstitutter og offentlige virksomheter kan søke om tilgang. Digitaliserings- og forvaltningsdepartementet jobber med dette gjennom EUs DIGITAL-programmet. DIGITAL-programmet er et investerings- og kapasitetsbyggingsprogram for digital omstilling og bruk av innovative digitale teknologier i samfunn og næringsliv. Gjennom EØS-avtalen deltar Norge på lik linje med EUs medlemsland. Digitaliserings- og forvaltningsdepartementet har tett dialog med Kunnskapsdepartementet om dette temaet siden Kunnskapsdepartementet har ansvaret Sigma2 og Digitaliseringsdepartementet har ansvaret for DIGITAL. Innovasjon Norge, Digitaliseringsdirektoratet og Forskningsrådet har ansvaret for å koordinere DIGITAL-programmet i Norge. Det har vært et ønske fra Digitaliserings- og forvaltningsdepartementet at det skal være en vektlegging av mulighetene virksomhetene, inkludert næringslivet, har til å tilgang til regnekraft gjennom DIGITAL-programmet.¹⁷¹

I april 2024 la regjeringen frem en strategi for norsk deltakelse i DIGITAL. Regjeringen ønsker at DIGITAL skal være et verktøy for den digitale omstillingen gjennom å utnytte mulighetene for finansiell støtte, tilgang til kunnskap, faglige nettverk og tilgang til kapasitet og infrastruktur.¹⁷²

5.4 Behov for kompetanse på KI

I Prop. 1 S (2020–2021) for Kommunal- og distriktsdepartementet vises det til at for å stimulere virksomheter til å ta i bruk KI er det behov for både spiss og bred kompetanse på kunstig intelligens. Den nasjonale KI-strategien fremholder at Norge skal ha høy kompetanse på grunnleggende IT-forskning og KI-forskning for å kunne forstå og dra nytte av endringer i teknologiutviklingen. Det krever at det finnes gode studietilbud som oppfyller de ulike sektorenes behov for både avansert kompetanse på kunstig intelligens og kompetanse på grunnleggende fag som statistikk, matematikk og informatikk.

Gjennom intervjuer, svar på brev og spørreundersøkelsen opplyser flere av virksomhetene at det er en utfordring med å sikre tilstrekkelig intern kompetanse som dataanalytiker og lignende. I spørreundersøkelsen svarer litt under halvparten av informantene (av 104) at de er avhengig av kompetanse fra innleide konsulenter eller leverandører for utvikling av KI. Konsekvensen har i hovedsak vært høyere kostnader og forsinkelser i arbeidet med å utvikle KI-systemer. Virksomhetene

¹⁶⁹ Brev fra Helse- og omsorgsdepartementet av 29. februar 24.

¹⁷⁰ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. mars 2024.

¹⁷¹ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. mars 2024.

¹⁷² Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. mars 2024.

oppgir i svar på brevet at det kan være krevende å få tilgang på og å bygge opp intern spisskompetanse på KI, og derfor må de gjerne leie inn konsulenter eller leverandører for å komme i gang med KI-utviklingen. At miljøene med KI-kompetanse er små, gjør også virksomhetene mer sårbare med hensyn til utvikling og drift av KI-systemer.

Flere av de intervjuede virksomhetene har etablert seksjoner eller avdelinger som spesialiserer seg på dataanalyse inkludert utvikling og bruk av kunstig intelligens. For eksempel har Nav en egen «AI-lab», tolltaten en dataanalyseavdeling og skatteetaten et eget kompetansesenter for datavitenskap bestående av ansatte med spisskompetanse på datautvikling, forretningsutvikling og prosjektledelse. Akershus universitetssykehus HF har en egen avdeling for medisinsk teknologi og e-helse med egne maskinlæringsspesialister. Statens pensjonskasse viser til at de har et godt internt kompetansemiljø for analyse og utvikling av KI. De ansetter kun kandidater med doktorgrad i fag som fysikk og statistikk, og det jobbes kontinuerlig med den interne kompetanseutviklingen.

De intervjuede virksomhetene trekker frem hvor viktig det er å ha juridisk kompetanse i arbeidet med KI. Tverrfaglig arbeid og god kompetanse står helt sentralt for å kunne utvikle gode KI-systemer. I intervju viser for eksempel Nav og skatteetaten til viktigheten av å sette sammen tverrfaglige grupper der ansatte i KI-fagmiljøet jobber tett sammen med fagfolk på de ulike tjenesteområdene samt jurister.

Videre er kompetanse blant ledere og beslutningstakere viktig for å lykkes med KI-prosjekter og realisere gevinster, siden det er ofte nødvendig med endringer i virksomhetens prosesser og god endringsledelse. Spørreundersøkelsen viser at mellom 35 og 40 prosent av respondentene har opplevd utfordringer i arbeidet med kunstig intelligens knyttet til manglende forankring internt, urealistiske forventninger av brukere eller ledelsen, eller motstand i organisasjonen for å endre eksisterende prosesser eller arbeidsstrukturer.

Finansdepartementet viser i intervju til at de fleste virksomhetene har tatt i bruk KI på en eller annen måte i produksjonsprosesser, men de trenger fortsatt å bygge kunnskap om og kompetanse på teknologien og hvordan den kan brukes på en god måte og implementeres i virksomhetenes arkitektur og produksjonsprosesser. Departementet viser videre til at alle underliggende virksomheter er utviklingsorienterte og bygger kompetanse, utforsker og finner ut hvilke muligheter og begrensninger som finnes i for eksempel ny teknologi. Finansdepartementet prøver å stimulere til dette.¹⁷³ Digitaliserings- og forvaltningsdepartementet viser i intervju til at utdanning og tilgang til kompetanse og spesialisert kompetanse på digitalisering og KI er viktig. Departementet viser til at det er Kunnskapsdepartementet som har ansvaret for kompetanse og utdanning, og det er en løpende dialog med Kunnskapsdepartementet om tilgangen på kompetanse, også utover den interdepartementale arbeidsgruppen for KI. Digitaliserings- og forvaltningsdepartementet viser i intervju til at temaet kompetanse er relevant for den interdepartementale KI-gruppen, men det er også en egen embetsgruppe for IT generelt, og departementet tar opp slike spørsmål i den gruppen.¹⁷⁴

Faktaboks 1 Interdepartemental gruppe for kunstig intelligens

Digitaliserings- og forvaltningsdepartementet har opprettet en interdepartemental gruppe for kunstig intelligens i 2018, som har fått flere medlemmer over tid og i dag er alle fagdepartementene representert. Hovedformålet er å ha dialog med de andre sektordepartementene og følge med på utviklingen av kunstig intelligens. Gruppen fungerer også som et nettverk hvor Digitaliserings- og forvaltningsdepartementet og andre fagdepartementer enklere kan kontakte hverandre og ha dialog om aktuelle problemstillinger. Gruppen møtes ved behov noen ganger i året og blir siden våren 2024 videreført som med revidert mandat.

Kilde: Digitaliserings- og forvaltningsdepartementet.

¹⁷³ Verifisert referat fra intervju med Finansdepartementet 29. november 2023.

¹⁷⁴ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. februar 2024.

Digitaliserings- og forvaltningsdepartementet ønsker ikke å utarbeide en kampanje eller andre tiltak som kun er rettet mot KI-kompetanse, men departementet mener det er behov for å lage noen strategier på digital kompetanse i forvaltningen generelt. Dette arbeidet er under planlegging. Departementet viser videre til at Digitaliseringsdirektoratet sammen med DFØ har laget en veiledning for mellomledere i staten for hvilken kompetanse de trenger for å drive digital transformasjon. I denne veiledningen kan eventuelt KI inkluderes.¹⁷⁵

Digitaliserings- og forvaltningsdepartementet har ifølge departementet alltid dialog med Kunnskapsdepartementet når det utvikler strategier, herunder i arbeidet med den nye digitaliseringsstrategien. Mangelen på kompetanse på IT, digitalisering og lignende har vært kjent i lengre tid. Universitetene og høyskolene har mye autonomi i hvordan de dimensjonerer studiene sine innenfor det oppdraget de er gitt. Digitaliserings- og forvaltningsdepartementet viser i intervju til Meld. St. 14 (2022–2023) *Utsyn over kompetansebehovet i Norge*, der regjeringen er tydelig på at ansvaret for å dimensjonere ressurser der behovet er størst, ligger hos utdanningsinstitusjonene.¹⁷⁶

Utdanningsløpet for IT og lignende tar tid, og dermed må det foregå kompetansehevingstiltak blant arbeidsstyrken i arbeidslivet. Kunnskapsdepartementet jobber med en ny stortingsmelding på dette området, ifølge Digitaliserings- og forvaltningsdepartementet.¹⁷⁷

Flere rapporter viser at behovet for IT-kompetanse er stort. På oppdrag fra Kompetansebehovsutvalget utarbeidet NIFU en undersøkelse om kompetansebehov i forbindelse med miljø- og klimautfordringer. Den konkluderte med at det forventes et stort behov for, og knapphet på, kompetanse innenfor IT og elektrofag på tvers av utdanningsnivåer. Rapporten peker på at det bør vurderes om noe av denne knappheten kan dekkes gjennom å utdanne flere på disse fagområdene.¹⁷⁸ I andre rapporter har man tallfestet behovet for flere med høyere IT-kompetanse frem mot 2030. Det vises til at det vil være behov for en økning fra ca. 40 000 ansatte med høyere IT-kompetanse i 2020 til mellom 55 000 og 75 000 ansatte i 2030.¹⁷⁹

En rapport på oppdrag fra Barne-, ungdoms- og familiedirektoratet som ble lansert i våren 2023 viser videre at det er stort behov for økning av kompetanse om risiko for diskriminering med kunstig intelligens.¹⁸⁰ Rapporten peker på behov for tverrfaglig kompetanse og tilgang på kompetansenettverk for å muliggjøre at virksomhetene kan ta i bruk kunstig intelligens med tilfredsstillende hensyn til risiko for diskriminering. Videre at dette må håndteres på et overordnet nivå i offentlig sektor siden en stor andel av offentlige virksomheter ikke har tilstrekkelig bred kompetanse i virksomheten.¹⁸¹

I Prop. 1 S (2023–2024) for Kunnskapsdepartementet vises det til at regjeringen har tydelige forventninger til at kunstig intelligens blir integrert i etablerte studieprogram der dette er relevant, og at regjeringen vil at utdanningsinstitusjonene vurderer hvordan personvern og etikk kan få en sentral plass i utdanningen innen kunstig intelligens. Kunnskapsdepartementet legger til grunn at institusjonene tar hensyn til disse forventningene i arbeidet med rammer for undervisningen.

Det har vært en økning i antall studieplasser innenfor IT de senere årene. I 2011 var det planlagt 1 480 studieplasser innenfor høyere IT-utdanning. I 2023 var tallet over 2 700. I samme periode har antallet søkere med IT som førstevalg økt fra 2 361 i 2011 til over 7 200 i 2023.¹⁸²

¹⁷⁵ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. februar 2024.

¹⁷⁶ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. februar 2024.

¹⁷⁷ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. februar 2024.

¹⁷⁸ Normann, H. E., Steen, M., Mäkitie, T., Klitkou, A., Børing, P., Solberg, E., Lund, H. B., Wardeberg, M. og Fossum, L. W. (2023). *Kompetanse for grønn omstilling: En gjennomgang av forskningslitteratur og arbeidslivets kompetansebehov knyttet til miljø- og klimautfordringer* (Rapport 2023:5). NIFU.

¹⁷⁹ Direktoratet for høyere utdanning og kompetanse. (2021, 24. september). *Behovet for IKT-kompetanse i Norge*. <https://hkdir.no/rapporter-undersokelser-og-statistikk/behovet-for-ikt-kompetanse-i-norge>. Se også Prop. 1 S (2023–2024) for Kunnskapsdepartementet.

¹⁸⁰ Rambøll Management Consulting og Vestlandsforskning, på oppdrag fra Barne-, ungdoms- og familiedirektoratet (Bufdir) (2022), *Bruk av kunstig intelligens i offentlig sektor og risiko for diskriminering*. Kunnskapsgrunnlag for arbeidet med å forebygge diskriminerende effekter ved bruk av kunstig intelligens i offentlig virksomhet, [VF-rapport nr. 7-2022](#)

¹⁸¹ Barne-, ungdoms- og familiedirektoratet (2023), *Bruk av kunstig intelligens i offentlig sektor og risiko for diskriminering*, avsnitt 5.2.4

¹⁸² Tall fra samordna optak.

5.5 Betydningen av språkressurser på norsk

Med utvikling av maskinlæringsmodeller basert på en ny arkitektur¹⁸³ i 2017, ble det mulig med større, kompleksere språkmodeller som åpnet for bruksområder som før ikke hadde vært tilgjengelig. Bruk av generativ kunstig intelligens basert på slike store språkmodeller har økt betydelig siden lanseringen av åpent tilgjengelig ChatGPT i november 2022¹⁸⁴. McKinsey & Company anslår potensialet med generativ kunstig intelligens til å være substansiell på tvers av næringslivet, og ser mange mulige bruksområder i offentlig sektor.¹⁸⁵ Grunnmodellene som disse KI-verktøy er imidlertid basert på er nesten utelukket utviklet av et fåtall globale, private selskaper uten tilknytning til norsk språk eller norske verdier.

For at norske innbyggere skal få mer avanserte KI-tjenester på sitt eget språk, er det avgjørende at det finnes gode språkressurser på begge skriftspråk og på samisk. Med bruk av kunstig intelligens kan man for eksempel effektivisere arbeidet med å oversette tekster, skrive tekster og med å analysere innholdet i store dokumentmengder. For å kunne utvikle slike løsninger for norske og samiske skriftspråk og dialekter, må imidlertid teknologien tilpasses disse språkene og de lokale forholdene. Språkteknologi, som verktøy for talegjenkjenning og språkforståelse, er derfor en viktig komponent i KI.

En sentral forutsetning for å kunne trene språkmodellene er tilgang på store mengder språkressurser. Ifølge den nasjonale KI-strategien, vil regjeringen legge til rette for at språkressurser samles inn og gjøres tilgjengelig.

Kultur- og likestillingsdepartementet har ansvaret for og styrer Nasjonalbiblioteket. Departementet opplyser i brev at Språkbanken ved Nasjonalbiblioteket tilbyr store datasett med norsk tekst og tale for å kunne utvikle god språkteknologi. Språkressursene i Språkbanken retter seg blant annet mot offentlige og private aktører som utvikler språkteknologiske produkter, blant annet i språkteknologikomponenter som inngår i KI.¹⁸⁶ I tildelingsbrevet fra Kultur- og likestillingsdepartementet for 2024 heter det blant annet at Nasjonalbiblioteket skal styrke sin posisjon som forskningsinstitusjon og leverandør av digital infrastruktur for digital humaniora og tilrettelagt språkkorpus til utvikling av norske KI-modeller. Delmålet er at Nasjonalbiblioteket skal produsere og tilby grunnlagsmateriale til bruk i språkteknologi, maskinlæring og kunstig intelligens.

Kultur- og likestillingsdepartementet oppgir også i brevet at det bør arbeides videre med å sikre språkdata til språkteknologi- og KI-formål. Tilgang på store grunnmodeller som støtter norske og samiske språk, kan bidra til utvikling av KI-verktøy som fungerer godt på norsk og samisk.¹⁸⁷ Det trengs store mengder data for å trene generelle språkmodeller (grunnmodeller), og det er behov for flere sett med fagspesifikke språkdata til å finjustere modellene slik at de fungerer på spesifikke fag- og forvaltningsområder. Tilgangen til treningsdata for ulike varianter av norsk (inkludert dialekter) er begrenset, og for fintrening av modeller i retning av det som trenges for chatbot-tjenester basert på store språkmodeller, er slike data nærmest fraværende, ifølge Nasjonalbiblioteket.¹⁸⁸

En utfordring ved tilgang på data til trening av generative språkmodeller er opphavsrett. Opphavsrett til åndsverk gir rettighetshavere en enerett til å fremstille eksemplarer av verket og gjøre verket tilgjengelig for allmennheten, jf. åndsverkloven § 3. I en digital sammenheng betyr dette blant annet at rettighetshaveren har enerett til å fremstille digitale kopier av sitt verk og laste det opp på internett.

¹⁸³ Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. og Polosukhin, I. (2017) *Attention is all you need*. [arXiv:1706.03762](https://arxiv.org/abs/1706.03762)

¹⁸⁴ Se for eksempel Teknologirådet (2024). *Gjennombruddet for generativ kunstig intelligens – en tidslinje*. [Nettartikkel](#)

¹⁸⁵ McKinsey & Company (2023), *Unlocking the potential of generative AI: Three key questions for government agencies*, [nettartikkel](#)

¹⁸⁶ Brev fra Kultur- og likestillingsdepartementet 19. februar 2024

¹⁸⁷ Brev fra Kultur- og likestillingsdepartementet 19. februar 2024

¹⁸⁸ Nasjonalbiblioteket svar på brev til Kultur- og likestillingsdepartementet av 31. august 2023 (elinnsyn).

Kultur- og likestillingsdepartementet og Nasjonalbiblioteket er i gang med å utrede problemstillinger knyttet til opphavsrett og trening av KI-modeller. Kultur- og likestillingsdepartementet har gitt Nasjonalbiblioteket i oppdrag gjennomføre et forskningsprosjekt som undersøker verdien av opphavsrettslig beskyttet materiale i trening av norske generative språkmodeller, og også vurdere om det er grunnlag for en kompensasjonsordning for norske rettighetshavere.¹⁸⁹

Digitaliserings- og forvaltningsdepartementet viser til at Digitaliseringsrundskrivet (21. desember 2022) gir konkrete anbefalinger om innhøsting og gjenbruk av offentlig produsert tekst til språkteknologiske formål. En av anbefalingene er at «[t]erminologi og begrepslister bør leveres til Felles begrepskatalog». Felles begrepskatalog ble utviklet i 2018 og er en del av Digitaliseringsdirektoratets Felles datakatalog. Formålet med Felles begrepskatalog er å legge til rette for deling av data gjennom å gjøre innholdet i dataene mer forståelig og skape en oversikt over dataenes betydning.¹⁹⁰

Videre utvikling og bruk av KI, blant annet utvikling av generelle språkmodeller, forutsetter bedre tilgang til regnekraft. Kultur- og likestillingsdepartementet samarbeider med andre relevante departementer om infrastruktur knyttet til kunstig intelligens.¹⁹¹ Departementet trekker frem at det er flere forhold som taler for å legge til rette for en felles norsk infrastruktur for KI. Til dette trengs det blant annet store mengder språkdata, utvikling av store språkmodeller og regnekraft for å trene språkmodeller på norsk og samisk. Dette er ressurser av et omfang som mindre offentlige og private enkeltaktører ikke kan skaffe til veie på egen hånd. Store internasjonale teknologiselskaper er ledende i utviklingen av KI, men det er ikke gitt at de tar hensyn til norsk språk og norske samfunnsforhold i sin utviklings- og forretningsvirksomhet. Fintrening av grunnmodeller til spesifikke formål og bruksområder bør i hovedsak kunne ivaretas av de enkelte virksomheter og sektorer.¹⁹²

Digitaliserings- og forvaltningsdepartementet oppgir i intervju at departementet løpende samarbeider med Digitaliseringsdirektoratet og Språkbanken ved Nasjonalbiblioteket om levering av norske korpus¹⁹³ til EU for bruk i sine systemer for automatisk oversettelse. Dette er også fulgt opp med en anbefaling i Digitaliseringsrundskrivet om at oversettelsesminner¹⁹⁴, som et resultat av kjøpte oversettelsestjenester, bør leveres til Nasjonalbibliotekets språkbank. Det er ikke alltid Nasjonalbiblioteket har fått tilsendt dette fra offentlige virksomheter.

Det er viktig at offentlige virksomheter som planlegger å ta i bruk nye KI-løsninger, tar hensyn til kravene i språkløvgivingen fra et tidlig tidspunkt i planleggingen. Alle offentlige organer har etter språkløven plikt til å bruke, utvikle og styrke bokmål, nynorsk og samiske språk. Ansvar for å bruke samiske språk er ytterligere regulert i sameløven. Dette kravet gjelder også kommunikasjon på digitale flater.¹⁹⁵

I intervju viser Digitaliserings- og forvaltningsdepartementet til at det har mye dialog om språkmodeller med Kunnskapsdepartementet og Kultur- og likestillingsdepartementet. I tillegg til arbeidet som gjennomføres av Nasjonalbiblioteket, viser departementet til at Universitetet i Oslo og NTNU jobber med å utvikle språkmodeller.¹⁹⁶

¹⁸⁹ Tildelingsbrev Nasjonalbiblioteket 2024

¹⁹⁰ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

¹⁹¹ Brev fra Kultur- og likestillingsdepartementet av 19. februar 2024.

¹⁹² Brev fra Kultur- og likestillingsdepartementet av 19. februar 2024.

¹⁹³ En avgrenset samling av utvalgte tekster.

¹⁹⁴ En database der setninger, avsnitt eller setningslignende enheter som tidligere er oversatt, er lagret for å hjelpe menneskelige oversettere.

¹⁹⁵ Brev fra Kultur- og likestillingsdepartementet av 19. februar 2024.

¹⁹⁶ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. mars 2024.

6 Statlige virksomheter tar hensyn til de etiske prinsippene i ulik grad i utvikling og bruk av kunstig intelligens

Hensynet til etiske prinsipper er undersøkt i deler av spørreundersøkelsen, gjennomgang av skriftlige retningslinjer og policyer knyttet til risikovurderinger av KI-systemer, og gjennom caseundersøkelser av fire utvalgte KI-systemer (jf. kapittel 2). Faktaopplysninger om de fire KI-systemene i caseundersøkelsen er verifisert gjennom separate caserapporter.

Mye av effektiviseringspotensialet med KI er basert på automatisering av arbeidsprosesser. Der kunstig intelligens brukes i forbindelse med automatisert saksbehandling, kan sentrale rettssikkerhetsprinsipper bli utfordret, og det er avgjørende med både tilstrekkelig kunnskap om forvaltningsrett i utvikling av systemene, og med teknologisk kompetanse i kontroll av saksbehandlingen og vedtak.¹⁹⁷ KI-systemer i kontakt med borgere, for eksempel basert på språkmodeller som gir tilpasset veiledning om saksbehandling, krever kompetanse i befolkningen, og avklart ansvar ved feil på grunn av misforståelser i kommunikasjonen med KI-systemet. KI-systemer som lærer av data kan videreføre og forsterke historiske skjevheter, eller endre ytelse eller praksis over tid. Videre kan det med avanserte maskinlæringsmodeller være utfordrende å få en god forklaring på hvorfor systemet handlet eller konkluderte slik det gjorde. Dermed kan det være vanskelig å sikre åpenhet og forklaringsevne og dermed en etterprøvbar saksbehandling, for å sikre klageretten.¹⁹⁸

Viktigheten av kontroll av KI-systemenes resultater kommer for eksempel frem av EUs KI-forordning, som krever at bruk av høyrisiko KI-systemer er tilrettelagt for effektiv overvåking av mennesker.¹⁹⁹

6.1 De utvalgte KI-systemene er mindre utviklingsprosjekter med lav risiko som brukes som beslutningsstøtte i saksbehandlingen

Skatteetaten, Statens pensjonskasse, Arbeidstilsynet og St. Olavs hospital HF har utviklet og tatt i bruk de utvalgte KI-systemene som inngår i caseundersøkelsen. Disse sorterer under tre ulike sektordepartement, har ulik ressurstilgang for utvikling og bruk av kunstig intelligens, og bruker KI til forskjellige formål. Det er også forskjeller i virksomhetenes erfaring, målt i antall KI-systemer under utvikling eller i bruk. Skatteetaten har erfaring fra utvikling eller bruk av flere enn 30 KI-systemer, mens Statens pensjonskasse har erfaring fra 6 KI-systemer og Arbeidstilsynet 2 KI-systemer.

Statens pensjonskasse og Arbeidstilsynet ønsket å utvikle KI-systemet delvis for å prøve ut mulighetene som kunstig intelligens gir. Utvikling av KI-systemet var helt nytt for disse virksomhetene. Arbeidstilsynet beskriver det som en læringsprosess, hvor de utforsker mulighetene med KI. Begge virksomhetene startet KI-utviklingen som mindre utviklingsprosjekter på avgrensede fag- og tjenesteområder som ga lavere risiko.

¹⁹⁷ Sivilombudsmannen. (2020). *Årsmelding for 2019. Dokument 4 (2019–2020)*.

¹⁹⁸ Sivilombudsmannen. (2020). *Årsmelding for 2019. Dokument 4 (2019–2020)*.

¹⁹⁹ KI-forordning, final draft (2024), artikkel 14

Tabell 2 Oversikt over sentrale kjennetegn ved de utvalgte KI-systemene

Sektor	Virksomhet	KI-system	Formåls-kategori ²⁰⁰	Beslutningsstøtte	I bruk siden	Utviklet selv	Kjører i sky	Personopplysninger
FIN	skatteetaten	fradragsmodellen	kontroll	ja	2022 (2014)	ja	ja	ja
AID	Arbeidstilsynet	prediksjonsindekstjenesten	kontroll	ja	2021 (2017)	ja	ja	nei
AID	Statens pensjonskasse	pensjonsutbetalinger	tildeling	ja	2021	ja	nei*	ja
HOD	St. Olavs hospital HF	fluorescensmikroskopet	tildeling	ja	2023	nei	nei	ja

Kilde: Riksrevisjonens caseundersøkelse av utvalgte KI-systemer

* Migrering til sky planlagt 2024

Skatteetaten, Statens pensjonskasse og Arbeidstilsynet startet utviklingen av KI-systemene på ulike tidspunkt. Utviklingen av KI-systemet hos Statens pensjonskasse startet i 2021, Arbeidstilsynet i 2017 og skatteetaten i 2013. KI-systemet hos skatteetaten og Arbeidstilsynet har vært videreutviklet med nye modeller i løpet av perioden, og Statens pensjonskasse planlegger en større utvikling til ny data- og analyseplattform i 2024.

Skatteetaten, Statens pensjonskasse og Arbeidstilsynet har selv utviklet KI-systemene, mens St. Olavs hospital HF har kjøpt inn medisinsk-teknisk utstyr som bruker KI fra en leverandør. Behov for det nye utstyret oppsto i 2020, og St. Olavs hospital HF er det første sykehuset i Norge som har kjøpt inn slikt utstyr med et innebygd KI-system i fagområdet.

De fire KI-systemene er å regne som beslutningsstøtte til saksbehandling/diagnostikk, hvor det er en eller flere saksbehandlere, kontrollører eller helsepersonell som treffer beslutningen. Det innebærer at KI-systemene er delvis automatiserte.

Skatteetatens og Arbeidstilsynets KI-systemer bidrar til *risikobasert* utplukk til henholdsvis kontroll og tilsyn. KI-systemenes prediksjoner effektiviserer arbeidet av saksbehandlere og kontrollører.

Skatteetatens KI-system *Fradragsmodell* predikerer risikoen for feil i skattefradrag og gjør et utplukk av skattepliktige til kontroll. Modellen er hovedsakelig trent på all informasjon om fradragene, informasjon fra skattemeldingen og informasjon om skatteyttere. Det behandles persondata ved bruk av KI-systemet. Det er saksbehandleren som treffer beslutningen om det skal gjøres kontroll. I tillegg er Fradragsmodellen planlagt å bidra til færre feil ved fradragsendringer og å bidra med analyser som kan forbedre regelverket.

Arbeidstilsynets KI-system *Prediksjonsindekstjenesten* brukes til å velge ut virksomheter/bedrifter for tilsyn gjennom å beregne sannsynligheten for funn av alvorlige brudd ved tilsyn hos tilsynsobjektet. KI-systemet brukes ikke på alle typer bedrifter, for eksempel ikke på jordbruksvirksomheter uten egne ansatte. Det er ulike føringer om å bruke Prediksjonsindekstjenesten innen ulike satsinger, ofte innenfor definerte målgrupper. Modellen er trent på resultater fra tidligere tilsyn. KI-systemet behandler ikke persondata. Tilsynsinspektørene får presentert, sammen med annen relevant

²⁰⁰ Formål av KI-systemer med konsekvenser for personer kan grovt klassifiseres i to kategorier basert på om konsekvensen er positiv (tildelingsformål) eller negativ (kontrollformål). Se f.eks. [LDOs veileder](#) for innebygd diskrimineringsvern, eller [Aequitas](#) fairness tree.

informasjon om tilsynsobjektet, en indikator som viser sannsynligheten for funn av alvorlige brudd, fordelt på fire kategorier fra lav til høy sannsynlighet.

KI-systemet *Pensjonsutbetalinger* i Statens pensjonskasse har som formål å «estimere pensjonsutbetalinger for å forhindre feilutbetalinger i etterlattepensjon». KI-systemet brukes til å beregne pensjonsytelser. Deretter sammenlignes ytelsene med saksbehandlernes fastsatte utbetalingsbeløp. KI-systemet brukes hittil kun for etterlattepensjon. KI-modellen tar utgangspunkt i pensjonsregelverket for å predikere et beløp for pensjonsytelsen for ulike medlemsgrupper i etterlattepensjon. KI-modellen er trent på data fra flere eksterne kilder og registre, deriblant lønns- og stillingsdata, Nav-data, inntektsopplysninger og personalia. Det predikerte beløpet brukes av saksbehandleren til å kontrollere og kvalitetssikre egne beregninger i etterkant av fastsettelsen av pensjonsbeløpet. Pensjonsregelverket for etterlattepensjon er kompleks, og derfor var det alltid to saksbehandlere på hver sak før KI-systemet ble tatt i bruk. KI-systemet har bidratt til at en saksbehandler kunne erstattes, men det er fortsatt alltid en menneskelig saksbehandler som tar beslutningen i hver sak.

KI-systemet St. Olavs hospital HF har kjøpt inn, er bygd inn i et fluorescensmikroskop, og bidrar til FISH-analyse (*fluorescence in situ hybridization*), som brukes i diagnostikk av kreft. KI-systemet klassifiserer celler i tre kategorier (positive, negative og udefinerbare) gjennom å påvise mutasjoner i enkeltgener. Andelen celler med mutasjon er et av flere analyseresultater som bidrar til diagnosen. Bruken av KI-systemet sparer tid sammenlignet med manuelt arbeid tidligere og forbedrer kvalitetssikringen og kvaliteten på analysene. Mikroskopet regnes som *in-vitro medical device* (IVD) og er CE-IVD sertifisert. Utstyret ble tilpasset og verifisert med data (vevsprøver) fra sykehuset. KI-systemet er kun en del av mikroskopet, og i anskaffelsesprosessen var det lagt mest vekt på andre faktorer.

6.2 Suksessfaktorer og utfordringer for utvikling av kunstig intelligens

I undersøkelsen av de utvalgte KI-systemene har vi kartlagt suksessfaktorer og utfordringer. Disse er basert på både erfaringene med utvikling av det enkelte KI-systemet og generelle erfaringer med KI for virksomheten.

Virksomhetene som utvikler KI-systemene selv, trekker frem følgende suksessfaktorer for utvikling av KI:

- begynne tidlig med et lavrisikoprojekt på et avgrenset fag-/tjenesteområde for å prøve seg frem, teste ut og bygge opp intern kompetanse innen KI
- etablere et tverrfaglig samarbeid med et team av utviklere, analytikere, fagfolk i linjen og jurister
- ha lovhemmel til å bruke egne data til analyseformål og utvikling av IT-systemer
- drive datadrevet innsiktsarbeid og prosessanalyse – analytikere kjenner til dataene, reglementet og saksbehandlingen som igjen reduserer faren for feil bruk av data
- bidra til en modningsprosess i faglinjen med mål om å utnytte mulighetene med KI

Tilsvarende har de erfart følgende utfordringer med KI utvikling og bruk:

- tilgang til kapasitet og kompetanse for å utvikle og forvalte KI-systemet
- tilgang til data, tilstrekkelig datakvalitet og/eller god nok IT-infrastruktur
- virksomhetens modenhet: utfordringer med å få virksomheten til å se mulighetsrommet med KI, samt forventningsavklaring om nye løsninger; muligheter og begrensninger
- synliggjøring av gevinster internt, men aller helst i forbedret måloppnåelse og effekt

- eierskap: løsninger som utvikles av en annen del av organisasjonen enn de som senere skal bruke den, noe som påvirker forventningsavklaringen og behovet for opplæring
- generelt enklere å håndtere etikk og KI på et overordnet nivå, men mer krevende i praksis

Ved innkjøp, tilpasning og validering av et leverandørsystem for bruk i sykehuset møter St. Olavs hospital HF andre utfordringer enn virksomheter som utvikler KI selv:

- En generell utfordring er at innenfor snevre fagområder kan det oppstå en monopolsituasjon for leverandører, noe som gir mindre fleksibilitet og øker kostnadene.
- En annen generell utfordring er at mange leverandører er amerikanske og har skyløsninger som det er krevende å bruke i Norge.
- Det har oppstått flere utfordringer i anskaffelsen av fluorescensmikroskop:
 - Det er usikkert om bruken av andre prober og tilpasninger er dekket av CE-sertifiseringen, da det ikke finnes retningslinjer for dette.
 - St. Olavs hospital HF er den første aktøren i Norge som anskaffer et slikt mikroskop med innebygd KI-system, og derfor finnes det ingen referanser for kravspesifikasjon eller validering av KI.
 - Det er liten kunnskap om KI i den medisinske fagavdelingen som stiller krav for konkurransen.
 - Det er krevende å integrere programvare og laboratedatasystem.
 - Siden KI-systemet ikke var hovedformålet med anskaffelsen, men del av det medisinske tekniske utstyret, kunne det ikke legges stor vekt på kriterier til KI i anskaffelsesprosessen, men andre hensyn ble avgjørende.

Før vi går nærmere inn i caseundersøkelsene, kartlegger vi rammebetingelsene for ansvarlig KI i virksomhetene. Den kartleggingen gjør vi ved å undersøke styringsdokumenter, som skriftlige retningslinjer brukt for KI, og i hvilken grad KI-prosjekter generelt følger de etiske prinsippene, som beskrevet av respondenter i spørreundersøkelsen.

6.3 Retningslinjer, policyer og arbeidet med ansvarlig KI ivaretar de etiske prinsippene i ulik grad

6.3.1 Styringsdokumenter brukt ved utviklingen av KI dekker i ulik grad de etiske prinsippene for ansvarlig KI

For å vurdere rammebetingelser for ansvarlig utvikling og bruk av KI spurte vi i spørreundersøkelsen om respondentenes bruk av skriftlige retningslinjer eller policyer på ulike områder relatert til ansvarlig KI. I det følgende omtaler vi det som styringsdokumenter. Det kan både være generelle styringsdokumenter og retningslinjer tilpasset til KI, så lenge de dekker områdene vi spurte om.

Digitaliseringsdirektoratet legger for eksempel til grunn at utvikling og bruk av KI ikke innebærer at statlige virksomheter må tenke helt nytt. Virksomhetenes etablerte rutiner og systemer for informasjonssikkerhet bør brukes som et utgangspunkt ved utvikling og bruk av KI. Samtidig er det en rekke risikoer og problemstillinger som særlig gjelder kunstig intelligens, som bør hensyntas i dette arbeidet.²⁰¹

Noen av de sentrale utfordringene med etiske prinsipper for ansvarlig KI står på

Digitaliseringsdirektoratets veiledningssider og i referansene der (jf. også kapittel 3):^{202,203}

²⁰¹ Digitaliseringsdirektoratet. (u.å.). *Hvordan vurderer jeg risiko ved bruk av kunstig intelligens?* Hentet 29. april 2024 fra <https://www.digdir.no/kunstig-intelligens/hvordan-vurderer-eg-risiko-ved-bruk-av-kunstig-intelligens/4537>

²⁰² Digitaliseringsdirektoratet. (u.å.). *Råd for ansvarlig utvikling og bruk av kunstig intelligens i offentlig sektor.* Hentet 24. mai 2024 fra <https://www.digdir.no/kunstig-intelligens/rad-ansvarlig-utvikling-og-bruk-av-kunstig-intelligens-i-offentlig-sektor/4272>

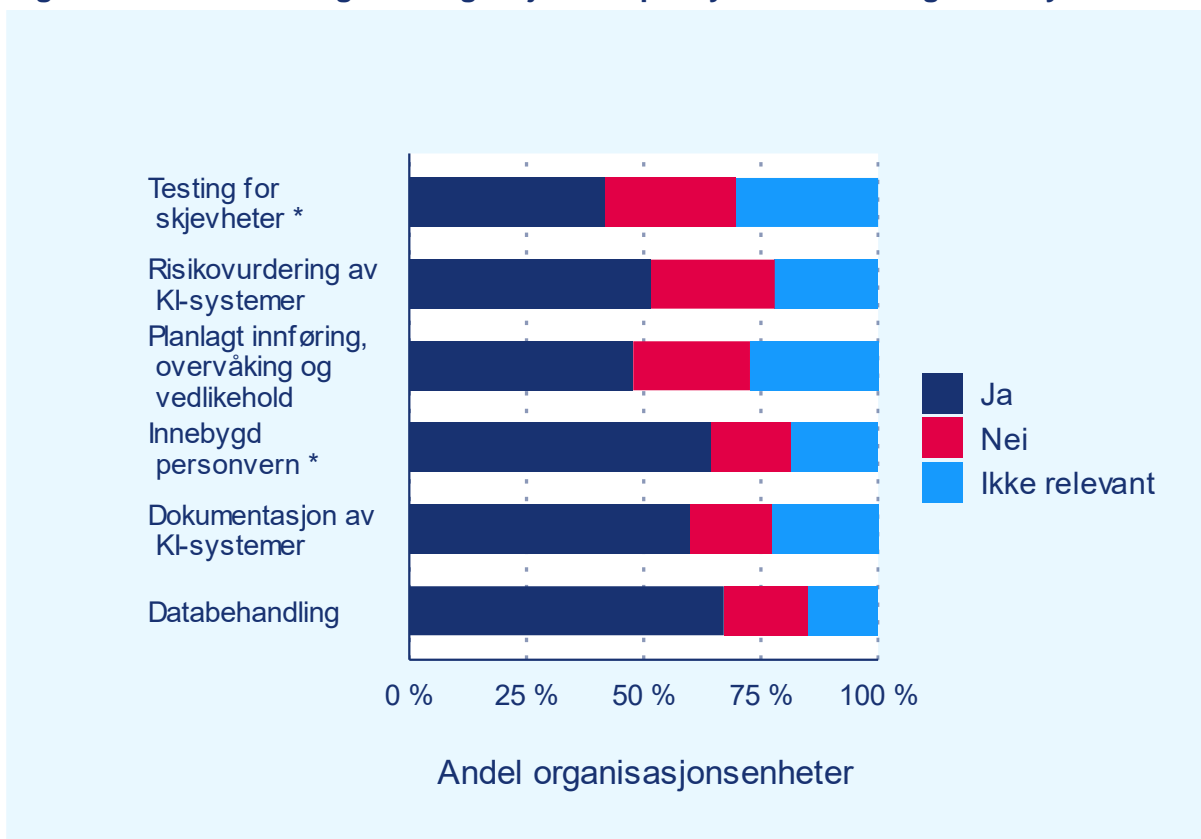
²⁰³ National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework AI RMF 1.0* (NIST AI 100-1). <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

- Personvern: Det kan være utfordrende å vurdere om trening av en maskinlæringsmodell er en behandling av personopplysninger som er i tråd med formålet for innhenting av disse opplysningene. Særlig gjelder dette hvis store mengder personopplysninger fra mange personer brukes til å trene en modell som skal brukes på *andre* personer. Dataminimeringsprinsippet blir utfordret. Videre kan retten til sletting av persondata bli vanskelig å etterleve dersom persondata er kodet inn i en maskinlæringsmodell.
- Likebehandling: Det er en risiko for diskriminering ved utvikling og bruk av visse KI-systemer. Hvis for eksempel dataene som brukes til å trene et KI-system, gjenspeiler uønskede samfunnsmessige fordommer, kan systemet utilsiktet opprettholde eller forsterke disse fordommene når det tar beslutninger eller gir prognoser.
- Åpenhet og forklarbarhet: Forvaltningsloven krever at enkeltvedtak skal kunne begrunnes, og at begrunnelsen viser til det aktuelle regelverket og de faktiske forholdene. Bruk av personopplysninger krever åpenhet og informasjon til dem opplysningene omhandler. For visse KI-systemer kan det være utfordrende å få en god forklaring på hvorfor systemet handlet eller konkluderte slik det gjorde.
- Sikkerhet: Informasjon kodet inn i algoritmer, som personopplysninger i treningsdata, kan være mulig å rekonstruere, og et KI-system kan dermed være en risiko med tanke på konfidensialitet. Et annet eksempel på en særegen risiko med KI-systemer er risikoen for at ondsinnede aktører får tilgang til treningsdata og kan endre disse på en måte som gjør at man får et bestemt utfall. Dette omtales vanligvis som *data poisoning*.
- Robusthet: At det mangler forklarbarhet for et utfall produsert av et KI-system, kan være utfordrende og innebære risiko med hensyn til likebehandlingen, ansvarligheten og påliteligheten til KI-systemet. Ved bruk over tid er det dessuten en risiko for at opprinnelige treningsdata ikke lenger passer til KI-systemets formål, noe som over tid mulig kan gi feil resultater. KI-systemer kan derfor kreve løpende overvåking og hyppig vedlikehold for å sikre at systemet gir resultater i tråd med formålet.

Figur 13 viser bruk av skriftlige retningslinjer eller policyer ved utvikling av KI-systemer med relevans for ansvarlig KI per organisasjonsenhet.²⁰⁴

²⁰⁴ En virksomhet kan ha flere organisasjonsenheter som jobber uavhengig av hverandre med KI. Flere respondenter fra samme virksomhet ble slått sammen til en organisasjonsenhet hvis det ikke gikk frem at respondentene hører til ulike organisasjonsenheter fra svar på brevet med spørsmål.

Figur 13 Bruk av skriftlige retningslinjer eller policyer ved utvikling av KI-systemer



Kilde: Riksrevisjonens spørreundersøkelse (n = 104 respondenter, aggregert til 82 organisasjonsenheter).

* Svar for retningslinjer om innebygd personvern og testing for skjevheter i KI-systemets utfall er fremstilt bare for virksomheter med minst ett KI-system som benytter persondata.

Figuren viser at retningslinjer og policyer, som brukes ved utvikling av KI-systemer, i ulik grad dekker forhold ved ansvarlig-KI. Over 60 prosent av organisasjonsenhetene oppgir å ha skriftlige retningslinjer eller policyer for databehandling og innebygd personvern, mens under halvparten oppgir å ha styringsdokumenter for planlagt innføring, overvåking og vedlikehold og testing for skjevheter av KI-systemer. Blant de som bruker persondata i KI-systemene, svarer 64 prosent at de har skriftlige retningslinjer eller policyer for innebygd personvern. Videre viser Figur 13 at styringsdokumenter for testing av skjevheter knyttet til utfall er minst brukt: 42 prosent av respondentene med minst ett KI-system som benytter persondata, sier de har slike retningslinjer eller policyer.

Noen respondenter henviser til at styringsdokumenter er under utvikling, eller at disse forholdene i praksis blir diskutert i prosjektteamet selv om det ikke finnes noen retningslinjer. Videre utdyper noen respondenter at ikke alle forhold er relevante fordi KI-systemene de bruker, er enkle eller bare brukes til interne prosesser.²⁰⁵ Slike forhold er reflektert i svaralternativet «ikke relevant».

Retningslinjer eller policyer for dokumentasjon av KI-systemer

Det er viktig å dokumentere hvordan et KI-system utvikles og implementeres.²⁰⁶ Det gjør systemet mer etterprøvbart og åpent ved at utviklingsprosessen kan bli forstått og modeller i systemet reproduisert.

²⁰⁵ Spørreundersøkelse, svar i fritekst.

²⁰⁶ European Commission, Directorate-General for Communications Networks, Content and Technology, (2019). *Ethics guidelines for trustworthy AI*, Publications Office. <https://data.europa.eu/doi/10.2759/346720>

Data-profiler (*data profile*)²⁰⁷ og modell-kort (*model cards*)²⁰⁸ er anbefalt av OECD²⁰⁹ for å dokumentere KI-systemer basert på maskinlæring, og angir en standard for minimumsgraden av innhold i dokumentasjonen. Videre bør dokumentasjonskrav vurderes opp mot risikoen i KI-systemet og mot brukerne av dokumentasjonen.

60 prosent av organisasjonsenhetene svarer i spørreundersøkelsen at de bruker skriftlige retningslinjer eller policyer for dokumentasjon av KI-systemer. Flere av respondentene utdyper i svar i fritekst at de svarer med bakgrunn i faktisk dokumentasjon istedenfor retningslinjer eller policyer, og at dokumentasjonen varierer mellom KI-systemene.²¹⁰ Flertallet av respondentene (flere enn 80 prosent) som oppga at de bruker styringsdokumenter til dokumentasjon, svarer at standarden deres for dokumentasjon helt eller delvis dekker

- databehandling og kilder
- modellvalg
- oppdatering og vedlikehold

Under halvparten svarer at standardene for dokumentasjon fullstendig dekker

- forklaring av KI-systemenes utfall
- valg av funksjons- og ytelsesnivå for KI-systemet²¹¹

Styringsdokumenter for risikovurdering av KI-systemer

Som nevnt i avsnittet ovenfor kan KI-systemer innebære risikoer som ikke nødvendigvis er dekket av tidligere standarder for risikovurderinger. Derfor har det blitt lansert en rekke nye standarder og anbefalinger for risikovurderinger av KI internasjonalt²¹².

KI-spesifikk risiko skyldes ofte kompleksiteten til algoritmene eller læring av mønstre fra data. Derfor er risikoen størst for KI-systemer som bruker maskinlæring. I spørreundersøkelsen oppga respondenter fra 23 organisasjonsenheter at de utvikler eller bruker minst ett avansert KI-system²¹³, og at de bruker skriftlige retningslinjer eller policyer for risikovurdering av KI-systemer. Vi mottok dokumentasjon av retningslinjene fra respondenter fra 17 virksomheter i tillegg til de fire virksomhetene i caseundersøkelsen.

Dokumentgjennomgang viser at alle som sendte sine relaterte styringsdokumenter, bruker generelle systemer for risikovurdering av KI-systemer, og omtrent halvparten har enten en overordnet strategi eller policy for ansvarlig bruk av KI i virksomheten eller jobber med å utvikle slike policyer eller retningslinjer.

I fem tilfeller, der man har laget en mal for risikoanalyse, omtales primært standard IT-sikkerhet og risiko knyttet til informasjonssikkerhet. Dersom man bruker en mal for personvern vurderinger og risikoanalyser med tanke på personopplysninger, omtales typisk risiko knyttet til personvernprinsippene som dataminimering eller formålsbegrensning.

Noen virksomheter har presisert at tjenester som ChatGPT ikke har unntak fra at taushetsbelagt informasjon ikke kan deles med eksterne tjenester, og at resultater fra slike systemer må

²⁰⁷ Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H. & Crawford, K. (2018). *Datasheets for Datasets*. <https://arxiv.org/abs/1803.09010>. Data profiles er en mal for dokumentasjon av datasett brukt i utvikling av maskinlæringsbaserte KI-systemer, som inkluderer informasjon som datakilder samt representasjon av ulike persongrupper og mulige skjevheter.

²⁰⁸ Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji., I. D. & Gebru, T. (2018). *Model Cards for Model Reporting*. <https://arxiv.org/abs/1810.03993>. Model cards er en mal for dokumentasjon av maskinlæringsbaserte KI-systemer, opprinnelig foreslått for standard informasjon som burde leveres av utviklere sammen med modeller som blir publisert til bruk av andre.

²⁰⁹ González, F., Ortiz, T., Sánchez Ávalos, R. (2021). *Responsible use of AI for public policy: Data science toolkit*. OECD, IDB. <https://oecd.ai/en/ai-public-policy-data-science-toolkit>

²¹⁰ Spørreundersøkelse, svar i fritekst.

²¹¹ Spørreundersøkelsen, spørsmål 44.

²¹² For eksempel ISO/IEC 23894:2020 – Risk management of AI systems og NIST Artificial Intelligence Risk Management Framework (AI RMF), NIST AI 100-1.

²¹³ Unntatt rene forskningsprosjekter, eller leverandørsystemer uten tilpasning eller validering.

kvalitetssikres. KI-spesifikke risikoer for informasjonssikkerheten og personvernet, som risikoen for lekkasje av persondata som er kodet inn i en maskinlæringsmodell, er ikke omtalt i noe av den tilsendte dokumentasjonen av systemer eller policyer for risikovurderinger. Risikoen på undersøkelsestidspunktet er imidlertid begrenset siden mange avanserte KI-systemer fortsatt er under utvikling, er bygd på ferdigtrente leverandørsystemer eller bare er brukt internt.

Enkelte virksomheter omtaler i tillegg risikoen ved robustheten til KI-systemer til bruk i automatiserte prosesser, bruk av tiltak som *model cards* (jf. avsnitt om retningslinjer eller policyer for dokumentasjon) eller overvåking og vedlikehold mot risiko for *modell-drift*²¹⁴.

Transparens og likebehandling er stort sett ikke dekket i styringsdokumentene, utover relaterte personvernprinsipper som er omtalt i maler for personvern-vurderinger. Det er få virksomheter som behandler risikoen for skjevheter og diskriminering, risikoen som følge av manglende forklarbarhet av kompliserte modeller, eller åpenhetskrav i automatiserte beslutninger.

6.3.2 Virksomhetenes arbeid med KI-utvikling ivaretar ansvarlig KI i ulik grad

I spørreundersøkelsen ble respondentene bedt om å svare på flere påstander om hvordan de jobber med KI-systemene, med svaralternativer «ja», «nei» eller «ikke relevant». Påstandene måler forhold som er relevante for de etiske prinsippene for ansvarlig KI, som sikkerhet og robusthet, forklarbarhet og åpenhet. Svarene²¹⁵ viser at enkelte av prinsippene for ansvarlig KI ivaretas mer enn andre.

Om lag 60 prosent av organisasjonsenhetene svarer *ja* til påstander om arbeidet med robusthet, for eksempel:

- *Vi har tilgjengelig dokumentasjon av endringer i KI systemet.*
- *KI-systemene våre har mekanismer for å håndtere feil eller uventet input.*
- *Vi tester KI-systemene våre kontinuerlig for å sikre at systemene er stabile over tid.*

Omtrent 25 prosent svarer at disse utsagnene ikke er relevante for deres bruk av KI.

Mellom 35 og 40 prosent av organisasjonsenhetene svarer ja til de følgende påstandene om sikkerhet:

- *KI-systemene våre testes for motstandsdyktighet mot manipulasjon og overlagt feilbruk.*
- *KI-systemene våre er testet for svakheter mot cyberangrep eller datainnbrudd.*

Mellom 40 og 45 prosent svarer at disse påstandene ikke er relevante for deres bruk av KI. Med om lag 15 prosent er andelen organisasjonsenheter som svarer *nei* på påstandene, lavest for disse utsagnene knyttet til sikkerhet og robusthet.

Derimot er andel *nei*-svar størst på utsagn knyttet til forklarbarhet (26 prosent) eller åpenhet (om lag 40 prosent):

- *KI-systemet gir begrunnelser for sine beslutninger eller resultater som er tilpasset brukerens kunnskapsnivå.*

Denne påstanden om forklarbarhet blir bekreftet av 40 prosent av organisasjonsenhetene, mens for en tredjedel er den ikke relevant.

Påstander om åpenhet får lavest andel bekreftende svar. Mellom 25 og 30 prosent av organisasjonsenhetene svarer ja til disse påstandene:

²¹⁴ Endring av modellens ytelse over tid

²¹⁵ Spørreundersøkelsen, spm. 41.

- *Informasjon om bruk av KI ligger åpent tilgjengelig for eksterne brukere eller berørte personer.*
- *Informasjon om hvordan KI systemet fungerer og treffer beslutning eller gir resultater, ligger åpent tilgjengelig for eksterne brukere eller berørte personer.*

Omtrent 30 prosent av organisasjonsenhetene svarer at dette ikke er relevant for deres KI bruk.

Spørreundersøkelsen inneholdt også spørsmål om likebehandling og testing av skjevheter i KI-systemets resultater. Bare respondenter med ansvar for KI-systemer som kan berøre enkeltpersoner direkte på en substansiell måte (jf. figur 10 i kapittel 4), ble bedt om å svare på om de har testet for skjevheter mellom ulike persongrupper i KI-systemets utfall.²¹⁶ Dette gjelder kun 16 prosent av KI-systemene i spørreundersøkelsen. For 58 prosent av disse KI-systemene er svaret nei på dette spørsmålet, og for 29 prosent svarer respondentene at de vet ikke. For kun 13 prosent av KI-systemene med resultater som kan berøre enkeltpersoner substansielt, angir respondentene at de har testet for skjevheter i KI-systemenes utfall.

Virksomhetene som svarte nei til at de testet skjevheter, har ulike forklaringer på dette:²¹⁷

- har ikke tilstrekkelige data
- ikke relevant for RPA-prosesser eller på grunn av konsekvenser for virksomheter og ikke privatpersoner
- har ikke lov til å bruke nødvendige personopplysninger for test av skjevheter
- er i en tidlig fase av KI-utviklingen
- kontrollerer utfall og retter det alltid opp i etterkant

6.3.3 KI-prosjektene bruker mest generelle retningslinjer og policyer som i ulik grad regulerer ansvarlig KI

KI-systemenes robusthet og den grunnleggende IT-sikkerheten, som tilgangsstyring, er relevant for alle KI-systemer. Særskilt KI-riisiko, som manipulering av inndata til trening av neste modellversjon, er ikke relevant i caseundersøkelsen, fordi de fire KI-systemene brukes bare av interne ansatte og videre ikke er helautomatiserte. Transparens inklusive forklarbarhet er relevant for alle de fire systemene dersom de brukes i saksbehandling, til utvalg av virksomheter til tilsyn, eller til diagnostikk. Rettferdighet er også relevant for alle fire KI-systemer: Tre av systemene brukes til prosesser som kan ha betydning for enkeltpersoner. Arbeidstilsynets KI-system brukes på virksomheter, og dermed er diskriminering etter likestillings- og diskrimineringsloven ikke relevant, men usaklig forskjellsbehandling (jf. avsnitt 6.4). Personvern er relevant for tre av de fire KI-systemene som bruker persondata (skatteetatens fradragmodell, Statens pensjonskasses system for pensjonsutbetalinger og St. Olavs hospital HFs fluorescensmikroskop).

Caseundersøkelsen viser at i alle de fire KI-prosjektene ble KI-systemene utviklet eller kjøpt inn i tråd med virksomhetenes generelle styringsrammeverk og retningslinjer. Gjennomgangen av rammeverk og retningslinjer viser at disse omtaler og stiller krav til noen av de etiske prinsippene for ansvarlig KI. Kravene og omtalen handler først og fremst om de etiske prinsippene om sikkerhet og robusthet og om personvern. Prinsippene om likebehandling og transparens blir i liten grad omtalt og inngår ikke som krav.

Arbeidstilsynet har lagt sine IT-policyer til grunn da det utviklet KI-systemet *Prediksjonsindekstjenesten*. Disse policyene omfatter blant annet retningslinjer for informasjonssikkerhet, klassifisering av informasjonsressurser og sikring av nettverk og infrastruktur.

²¹⁶ Spørreundersøkelsen, spørsmål 39.

²¹⁷ Spørreundersøkelsen, spørsmål 40.

Arbeidstilsynet regulerer også personvern hensyn i en egen retningslinje for innebygd personvern og personvernerklæring. Forhold ved likebehandling omtales ikke og inngår ikke som krav.

Virksomhetene i caseundersøkelsen har, med unntak av skatteetaten, ikke etablert en egen KI-policy. Skatteetaten godkjente i 2023 en intern KI-policy²¹⁸ for hele virksomheten, som definerer fem prinsipper for etatens utvikling og bruk av kunstig intelligens: *ansvarlighet, åpenhet, likebehandling, teknisk robusthet, sikkerhet og beredskap, og datastyring og personvern*. Dette dekker de etiske prinsippene for ansvarlig utvikling av kunstig intelligens. Bruken av prinsippene i skatteetaten skal ses i sammenheng med automatiseringsgraden av oppgaven, hvor direkte bruken av KI treffer mottakerne, og risikoen. I KI-prosjektet som inngår i caseundersøkelsen, er arbeidet med å operasjonalisere disse prinsippene i oppstartsfasen, og det er ikke ennå klart hvordan de kan praktisk gjennomføres i prosjektene.

6.4 Etiske prinsipper om personvern, sikkerhet og robusthet ivaretas i større grad enn prinsippene om transparens og likebehandling

I denne delen fremstiller vi styringen av utviklingen av KI-systemene og hvordan de etiske prinsippene for ansvarlig KI ivaretas. På bakgrunn av rammeverket²¹⁹ og spørsmålene der sammenfatter vi styringen i fem hovedkategorier:

- styring og ansvar ved utvikling eller innkjøp, tilpasning og validering av KI-systemet
- personvern
- likebehandling
- åpenhet
- sikkerhet og robusthet

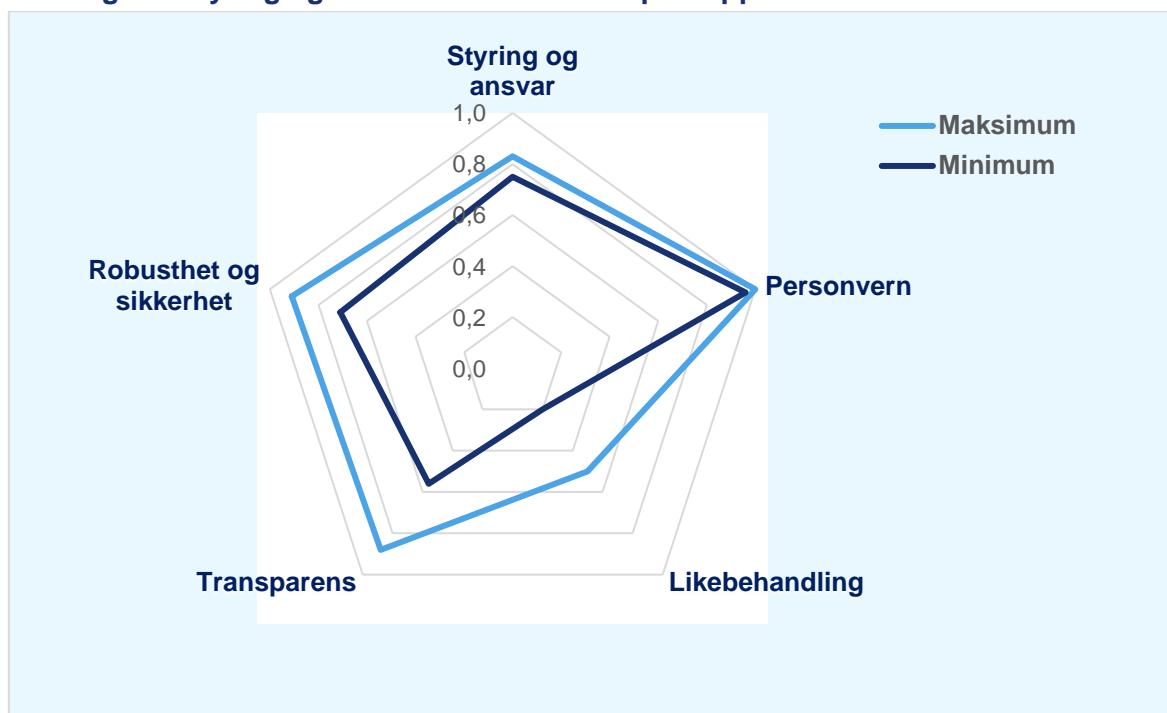
Vi vurderer modenhetsnivået til KI-systemet innenfor de ulike etiske prinsippene ut fra hvor mange krav i de underliggende revisjonsspørsmålene i rammeverket som er oppfylt. For hver kategori er det beregnet en gjennomsnittsscore (med en skala fra 0,0 til 1,0) av kravene og forholdene vi har vurdert i rammeverket for hvert case. Scoren gjenspeiler KI-prosjektene modenhets- eller aktivitetsnivå på de ulike områdene.

Vurderingen av modenhetsnivå er basert på dokumentasjonen vi har mottatt, og opplysninger fra prosjektteamene. Vi har ikke gjennomgått kode eller utført tekniske tester. Et lavt modenhetsnivå innenfor likebehandling betyr derfor ikke at KI-systemet forskjellsbehandler, men at KI-prosjektet har dokumentert en lav andel av aktiviteter som skal sikre likebehandling, ut fra vårt rammeverk.

²¹⁸ Skatteetaten (2023, 20 november). *Policy for utvikling og bruk av KI i Skatteetaten*. <https://www.skatteetaten.no/globalassets/om-skatteetaten/om-oss/ki/policy-ki-i-skatteetaten-1-1.pdf>

²¹⁹ Rammeverket består av en liste med revisjonsspørsmål knyttet til hvordan de etiske prinsippene som utgjør ansvarlig KI er ivaretatt i KI-prosjektene styring, bruk av data og modellens utvikling, implementering og overvåking. Jf. kapittel 2.2 og vedlegg 3

Figur 14 Modenhetsnivå av selvutviklede KI-systemene i caseundersøkelser om ansvarlig KI - styring og ivaretagelse av etiske prinsipper



Kilde: Riksrevisjonens caseundersøkelse av utvalgte KI-systemer (n = 3).

Figur 14 viser minimum og maksimum modenhetsnivå til de tre KI-prosjektene som virksomhetene har utviklet selv, i de fem vurderte kategoriene. Resultatene fra undersøkelsen viser at KI-systemene scorer høyt på kategoriene *styring og ansvar*, *personvern* og *robusthet og sikkerhet*. KI-systemene scorer lavere og med større variasjon på kategoriene *transparens* og *likebehandling*.

St. Olav hospital HF's KI-system, som er innebygd i medisinsk-teknisk utstyr som sykehuset har kjøpt inn, er ikke med i den samme oversiktsfiguren, siden grunnlaget for å vurdere modenhetsnivået for de ulike etiske prinsippene er for forskjellig fra selvutviklede systemer. Utstyret er CE-sertifisert med samsvarserklæring om at det er i henhold til relevant regelverk²²⁰, og derfor er noen av de kriteriene vi vurderer i rammeverket, utenfor sykehusets ansvar, men ligger hos det *meldte organet* som har vurdert det slik at utstyret oppfyller kravene i regelverket (samsvarsvurdering). Hvorvidt disse kravene dekker særskilt risiko med KI, er ikke St. Olavs hospital HF's ansvar å sikre, utover å passe på at utstyret har CE-merking og en samsvarserklæring. Derfor er ikke dette en del av vår undersøkelse.²²¹ Vi har basert vurderingen vår av håndteringen av de etiske prinsippene i forbindelse med fluorescensmikroskopet på tilpasnings- og valideringsprosessene i sykehuset, og dette er beskrevet under avsnittene om hvert etiske prinsipp nedenfor. Hovedtrekkene ligner imidlertid de selvutviklede KI-systemene: Styring og ansvar, personvern, robusthet og sikkerhet er i større grad jobbet med enn transparens og likebehandling.

6.4.1 KI-systemene er tydelig forankret i virksomhetene og kan vise til positive resultater, men dokumentasjon av KI-systemet har svakheter

Kategorien styring og ansvar, jf. figur 14, fikk en høy score ved vurdering av KI-systemene. Undersøkelsen vår av styring og ansvar for de tre KI-systemene som virksomhetene har utviklet selv, viser at

²²⁰ IVDR: In-vitro diagnostic medical device regulation, Regulation (EU) 2017/746.

²²¹ CE-sertifisering av utstyr med høyrisiko KI kommer til å bli oppdatert med KI-forordningen.

- KI-systemene er tydelig forankret i virksomhetenes strategi- og utviklingsarbeid
- utviklingen av KI-systemene skjer etter faseindelte planer og milepæler
- det er etablert en tydelig målstruktur for utvikling av KI-systemene
- KI-systemene gir positive resultater og gevinster
- risikovurderingene av KI er generelle og er ikke gjennomført for det enkelte KI-systemet
- KI-systemene er mindre utviklingsprosjekter med lav risiko
- risikoene som vurderes ved ansvarlig KI er robusthet og sikkerhet og personvern
- dokumentasjonen av KI-systemets utvikling og bruk har svakheter og mangler

Utviklingen av KI-systemene legger til grunn faseindelte prosjekt- og utviklingsplaner med milepæler, men livssyklusplanlegging for KI-systemet gjennom alle livsfaser er ikke en del av styringen. For eksempel har Statens pensjonskasse brukt et produktutviklingsteam ved utvikling av KI-systemet fremfor en mer tradisjonell IT-prosjektorganisering. Her er planene lagt for kortere utviklingsforløp med hyppigere leveranser, og mer prøving, feiling og korrigeringer underveis. Siden flere av KI-systemene er organisert som mindre utviklingsprosjekt med få utviklingsressurser, mener virksomhetene at det er mindre behov for tradisjonell prosjektplanlegging.

KI-systemene har en tydelig forankring i virksomheten de utvikles i, ved at de er tydelig koblet til strategi- og utviklingsplaner, og at de brukes innenfor virksomhetens kjerneoppgaver. Det er etablert tydelige formål med KI-prosjektene, og disse er tatt ned i mer konkrete effekt-/gevinstmål. Likevel er det ikke definert kvalitets- og ytelsesmål for KI-modellen i målstrukturen, som grunnlag for å vurdere når KI-systemet skal tas i bruk.

KI-systemets måloppnåelse og resultater er dokumentert gjennom analyser etter at systemet er tatt i bruk. Resultatene av KI-systemet er også vurdert i lys av erfaringer med bruk av systemet i saksbehandlingen. Med unntak av Fradragsmodellen i skatteetaten er ikke KI-systemene systematisk evaluert. Virksomhetene vurderer at hovedformålet med KI-systemet er oppnådd og kan dokumentere blant annet følgende positive resultater:

Tabell 3 Hovedresultater og gevinster ved utvalgte KI-systemer

Arbeidstilsynet	Skatteetaten	Statens pensjonskasse	St. Olavs hospital HF
Mer treffsikre og effektive tilsyn	Mer treffsikre og lønnsomme kontroller	Høyere kvalitet i utbetalingsporteføljen og bedre ressursbruk	Tidsbesparelser, høyere kvalitet på analyse og kvalitetssikring

Kilde: Riksrevisjonens caseundersøkelse av utvalgte KI-systemer

For eksempel har treffsikkerheten²²² til skatteetatens kontroller ved bruk av KI-systemet økt, og kontrollene er mer lønnsomme målt i gjennomsnittlig antall kroner avdekket ved kontrollene.

Risikovurderingene som er gjennomført i forbindelse med utviklingen av KI-systemene, tar utgangspunkt i skybaserte data- og analyseplattformer og generell behandling av persondataene i virksomheten. Det er ikke gjennomført risikovurderinger av selve KI-systemet og algoritmene som brukes.

Som tidligere omtalt har virksomhetenes formål med KI-utviklingen delvis vært å prøve ut og få erfaring med KI. Derfor er alle de selvutviklede KI-systemene organisert som mindre utviklingsprosjekter på et område med lav risiko. KI-systemene brukes som beslutningsstøtte i saksbehandlingen, og det er alltid en saksbehandler eller tilsynsinspektør som treffer beslutningen. Resultatene til KI-systemet i fluorescensmikroskopet ved St. Olavs hospital HF blir verifisert av

²²² Med treffsikkerheten til kontroller menes her hvor stor andel av de kontrollerte fradragsendringene som modellen predikerte som uriktige, og som faktisk var uriktige.

bioingeniør og patolog, og analysen er bare en av flere bidrag til diagnosen. Virksomhetene vurderer derfor risikoen ved bruk av KI-systemene som lav.

Risikoen forbundet med ansvarlig utvikling av KI er ulik blant KI-systemene. For eksempel behandler skatteetaten og Statens pensjonskasse persondata i sine KI-system, mens Arbeidstilsynet ikke bruker persondata. Risiko relatert til de etiske prinsippene om robusthet og sikkerhet og på personvernet er vurdert ved utviklingen av KI-systemene. For eksempel er det vurdert å være risiko med informasjonssikkerheten ved alle KI-systemene. Videre har skatteetaten og Statens pensjonskasse vurdert risiko, i ulikt omfang, for at KI-systemene går ut over personvernet. De etiske prinsippene transparens og likebehandling er ikke vurdert på grunn av KI-systemenes innretning og risikoprofil.

Dokumentasjonen av KI-systemene viser svakheter og mangler. Enkelte KI-systemer mangler oppdatert informasjon og inneholder informasjon som ikke lenger er riktig. For eksempel blir informasjon om KI-systemets måltall og resultater ikke oppdatert fra tidligere versjoner av KI-systemet. Videre mangler det informasjon om hvilke veivalg som er tatt ved utvikling av KI-modellen, og hvorfor en KI-modell er valgt fremfor andre. Dialogen om videreutvikling av KI-systemet er også i varierende grad dokumentert. Virksomhetene som har tatt i bruk skyløsninger, oppgir at deler av dokumentasjon av KI-systemet, som prestasjon og informasjon om drift, har blitt bedre ved bruk av skyløsninger. Virksomhetene forklarer svakheter og mangler i dokumentasjonen av KI-systemene med at dette er mindre utviklingsprosjekter med få utviklingsressurser. Videre er deler av dokumentasjonen, som modellutprøving og valg, dokumentert i lokale applikasjoner og fagsystemer som det er krevende å hente ut informasjonen fra og etterprøve utenfor utviklingsteamet.

I forbindelse med caseundersøkelsen peker Statens pensjonskasse på at det er nødvendig med felles maler som kan brukes i KI-utvikling, blant annet basert på risiko knyttet til KI, inkludert det man *ikke* må gjøre. Ifølge virksomheten burde veiledningsmaterieell sammenstilt av Digitaliseringsdirektoratet operasjonaliseres og inkludere felles maler for vurderinger av om et KI-system burde gjennomføre en vurdering av personvernkonsekvenser, eller hvilke typer dokumentasjon om KI-systemet som virksomheter burde vurdere. På denne måten kunne innbyggernes rettssikkerhet øke ved at rettigheter og friheter ivaretas på lik måte i staten. I tillegg kunne arbeidet med KI i virksomhetene effektiviseres betydelig og åpne for mer innovasjon med KI i små og mellomstore virksomheter.

Situasjonen for leverandørsystemet er mer kompleks fordi KI-systemet er kun en del av mikroskopet. Sammenligning av ulike tilbud og utvalgsriterier i anskaffelsen legger mest vekt på kriterier som ikke berører KI-systemet. Dokumentasjonen fra leverandør er rettet mot brukere eller teknisk vedlikehold av mikroskopet og inneholder liten informasjon om KI-systemet.

6.4.2 Personvernprinsippene er ivaretatt ved utvikling av KI-systemene

Kategorien personvern omfatter blant annet hvordan KI-systemene behandler og bruker persondata, rettslig grunnlag for behandling av persondata, vurdering av personvernkonsekvenser og dataminimering. Disse forholdene inngår som en del av krav til innebygd personvern.²²³

KI-systemene ved skatteetaten, Statens pensjonskasse og St. Olavs hospital HF behandler og bruker persondata. KI-systemene har fått en høy score for kategorien personvern, slik figur 14 viste. Den høye scoren skyldes flere forhold:

- Saksbehandlerne bruker KI-systemene som beslutningsstøtte.
- Virksomhetene har utviklet egne maler for vurdering av personvernkonsekvenser og vurdert personvernprinsippene og bruk av personopplysninger for KI-systemene (jf. avsnitt 6.1.1).

²²³ Datatilsynet. (2023, 27. juli). *Innebygd personvern og personvern som standard*. <https://www.datatilsynet.no/rettigheter-og-pliker/virksomhetenes-pliker/innebygd-personvern-og-personvern-som-standard/hva-er-innebygd-personvern/>

- KI-systemene ivaretar lovligheten ved behandling og bruk av personopplysninger gjennom hjemmel i egne særlover.
- Dataminimeringsprinsippet er ivaretatt ved utviklingen av KI-systemene.

Kravet til innebygd personvern omfatter å ta i bruk tilstrekkelige tekniske og organisatoriske tiltak og å ivareta personvernprinsipper og de registrertes rettigheter og friheter på en effektiv måte. Resultatene ovenfor viser at utviklingen av KI-systemene følger kravene til innebygd personvern.

Ved utvikling av KI-systemene vurderes rolle og betydning i saksbehandlingsprosessen og i relasjon til saksbehandler/helsepersonell. KI-systemene behandler og bruker personopplysninger for å gi beslutningsstøtte. Dokumentasjonen tydeliggjør at det er saksbehandler/helsepersonell som bruker eget faglig skjønn og tar den endelige beslutningen.

Personvernforordningen artikkel 35 nr. 3 bokstav a stiller krav om en vurdering av personvernkonsekvenser i tilfelle av automatisert modellering av personlige aspekter av fysiske personer «som danner grunnlag for avgjørelser som har rettsvirkning for den fysiske personen eller på lignende måte i betydelig grad påvirker den fysiske personen»²²⁴. Virksomhetene som bruker persondata, har utviklet egne maler som hjelper med vurderingen av personvernkonsekvenser og med å bestemme om det er behov for å gjennomføre en slik vurdering.²²⁵ Personvernprinsippene²²⁶ og hvordan disse er ivaretatt, samt hvilke personopplysninger som kan brukes i KI-modellen, er vurdert for KI-systemene. Skatteetaten og Statens pensjonskasse vurderer begge prosessen med vurderingen av personvernkonsekvenser som såpass ressurskrevende at det oppleves som utfordring i KI-prosjektene.

Behandling av personopplysninger krever at virksomheten har et rettslig grunnlag (lovlighetsprinsippet).²²⁷ Lovligheten ved behandling og bruk av personopplysninger er vurdert for KI-systemene som virksomhetene har utviklet selv med persondata, fradragsmodellen i skatteetaten og Statens pensjonskasses system for pensjonsutbetalinger. Vurderingen viser at virksomhetene har rettslig grunnlag for å analysere, utvikle og teste IT-systemer og for noen også profilering. Virksomhetenes rettslige grunnlag er gitt ved særlovsbestemmelser, for eksempel skatteforvaltningsloven²²⁸ og Statens pensjonskasseloven.²²⁹ Virksomhetene mener dette var en suksessfaktor for utviklingen av KI-systemene.

Kun personopplysninger som er nødvendige for behandlingen, skal behandles. Ifølge dataminimeringsprinsippet skal innsamlingen og behandlingen av persondata være begrenset til kun det som er nødvendig.²³⁰ Dette kan være utfordrende i utvikling av KI-modeller, siden disse lærer av dataene hvilke opplysninger som er relevant og nødvendige til formålet i modelltreningen. Dataminimering er vurdert og praktisert i utviklingen av KI-systemene blant annet gjennom testing av effekten av dataene som inngår i KI-modellen. For eksempel vektlegger Statens pensjonskasse dataminimering i sitt utviklingsarbeid og vurderer løpende hvilke variabler det er behov for i deres behandling. Både skatteetaten og Statens pensjonskasse vektlegger også egen kunnskap og erfaring

²²⁴ Personvernforordningen artikkel 35 nr. 3 bokstav a: «en systematisk og omfattende vurdering av personlige aspekter ved fysiske personer som er basert på automatisert behandling, herunder profilering, og som danner grunnlag for avgjørelser som har rettsvirkning for den fysiske personen eller på lignende måte i betydelig grad påvirker den fysiske personen».

²²⁵ Datatilsynet. (2023, 27. juli). Vurdering av personvernkonsekvenser (DPIA). <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/vurdering-av-personvernkonsekvenser/>

²²⁶ Beskrevet bl.a. i [Datatilsynets veiledning om innebygd personvern og personvern som standard](#): Lovlighet, rettferdighet, åpenhet, formålsbegrensning, dataminimering, riktighet, lagringsbegrensning, integritet og konfidensialitet, og ansvarlighet.

²²⁷ Datatilsynet. (2023, 27. juli). *Innebygd personvern og personvern som standard*. <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/innebygd-personvern-og-personvern-som-standard/>

²²⁸ Jf. skatteforvaltningsloven § 5-11. *Sammenstilling, profilering og automatiserte avgjørelser* og § 5-12. *Bruk av personopplysninger ved utvikling og testing av IT-systemer*. Skatteforvaltningsloven. (2016). *Lov om skatteforvaltning* (LOV-2016-05-27-14). Lovdata. https://lovdata.no/dokument/NL/lov/2016-05-27-14/KAPITTEL_5#KAPITTEL_5. Kapittel 5 Alminnelige saksbehandlingsregler.

²²⁹ Statens pensjonskasseloven. (1949). *Lov om Statens pensjonskasse* (LOV-1949-07-28-26). Lovdata https://lovdata.no/dokument/NL/lov/1949-07-28-26/KAPITTEL_9#KAPITTEL_9. Kapittel 8. Forskjellige bestemmelser, § 45 b.

²³⁰ Datatilsynet. (2019, 16. juli). *Dataminimering*. <https://www.datatilsynet.no/rettigheter-og-plikter/personvernprinsippene/grunnleggende-personvernprinsipper/dataminimering/>

med behandling av dataene som inngår i KI-modellen, samt fagkunnskapen i saksbehandlingen hvor KI brukes.

6.4.3 Likebehandling vurderes på et overordnet nivå, og testing av skjevheter varierer blant KI-systemene

Likebehandling kan tolkes på ulike måter, og handler ofte om hvordan *personer* behandles. Kravene til saklighet, forsvarlighet og proporsjonalitet i offentlig forvaltning gjelder imidlertid hele saksbehandlingen.²³¹ Selv om det ikke er noen risiko for diskriminering av personer etter likestillings- og diskrimineringsloven, kan det likevel være risiko for at virksomheter utvikler og bruker KI-systemer som gir usaklig forskjellsbehandling og at KI-modellen ikke er tilstrekkelig robust og treffsikker.

I analysen av kategorien likebehandling har vi blant annet vurdert dokumenterte tester av ulike former for skjevheter (bias) i både data som inngår i modellutviklingen, og utfallet av KI-systemet, i tillegg til vern mot diskriminering²³². Resultatene for kategorien likebehandling, som er vist i figur 14, viser at modenhetsnivået er lavere enn i andre kategorier og varierer mer mellom KI-systemene. Undersøkelsen viser at likebehandling har ulik betydning i de ulike KI-systemene og derfor håndteres forskjellig i KI-systemene.

Likebehandling vurderes på overordnet nivå i to av virksomhetene. Skatteetatens nylig lanserte policy for utvikling og bruk av KI²³³ omtaler rettferdighet som overordnet prinsipp, men arbeidet med å tolke og operasjonalisere prinsippet i det undersøkte KI-systemet er i startfasen. Statens pensjonskasse vurderer rettferdighet på overordnet nivå som personvernprinsipp i en personvernkonsekvensvurdering.

De tre selvutviklede KI-prosjektene i caseundersøkelsen sannsynliggjør at inndataene er representative, men ingen av casene har dokumentert relaterte tester av inndata eller av KI-systemets utfall. Treningsdataene som leverandørsystemet bygger på, er ukjent. Prøver for tilpasning av mikroskopet ble valgt ut basert på faglige vurderinger slik at ulik signalkvalitet kan testes, men representasjon av ulike persongrupper er ikke dokumentert.

Statens pensjonskasse viser til at medlemsmassen og datagrunnlaget er dominert av kvinner, og bruker derfor ikke variabler som navn, alder, kjønn og sivilstatus i modellen. Erfaringer og forskning viser imidlertid at det ikke er nok å fjerne diskriminerende variabler som kjønn, alder mv. for å unngå skjevheter i utfallet av maskinlæringsmodeller.²³⁴

KI-systemer som bruker persondata, gjennomfører noen tester for usaklig forskjellsbehandling. For eksempel tester skatteetatens modell for ulik ytelse ut fra ulike typer fradrag, men ikke for ulike persongrupper ut fra persondatavariabler. I valideringsprosessen av mikroskopet har sykehuset testet om det kan være skjevheter som går ut over nøyaktigheten av klassifisering av celler i ulike organgrupper eller krefttyper. Sykehuset vurderer at det ikke er relevant å teste for skjevheter ut fra ulike persongrupper ved analyse på cellenivå, som i tillegg kvalitetssikres av medisinsk personell.

KI-prosjektene som virksomhetene utvikler selv, jobber med å sikre at KI-systemenes utfall er forklarbare (jf. avsnitt om transparens), noe som på sikt kan bidra til en forståelse av hvordan hver

²³¹ Jf. ulovfestede prinsipper i norsk forvaltningsrett, [NOU 2019:5](#) kap. 11.7

²³² [Likestillings- og diskrimineringsombudets veileder for innebygd diskrimineringsvern](#) forklarer at rettferdighet kan testes ut fra om KI-systemet fungerer like godt for ulike grupper, som et av fem steg i innebygd diskrimineringsvern. Et annen steg er test og eventuell korrigering av mulige skjevheter i datagrunnlaget modellen trenes på. For vurdering om diskriminering etter likestillings- og diskrimineringsloven, er dette persongrupper definert via diskrimineringsvariabler. Mer generelt kan både personer og virksomheter grupperes etter variabler som ikke burde ha betydning av avgjørelsen for å teste usaklig forskjellsbehandling.

²³³ Skatteetaten. (2023). Policy for utvikling og bruk av KI i Skatteetaten (versjon 1.1). <https://www.skatteetaten.no/globalassets/om-skatteetaten/om-oss/ki/policy-ki-i-skatteetaten-1-1.pdf>

²³⁴ Se f.eks.: Likestillings- og diskrimineringsombudet. (2023). Innebygd diskrimineringsvern. https://ldo.no/globalassets/ldo_2019/bilder-til-nye-nettsider/ki/ldo-innebygd-diskrimineringsvern.pdf; Ruf, B. & Detyniecki, M. (2020). Active Fairness Instead of Unawareness. <https://arxiv.org/pdf/2009.06251.pdf>; og Verma, S. & Rubin, J. (2018). Fairness Definitions Explained. 2018 ACM/IEEE International Workshop on Software Fairness. <https://fairware.cs.umass.edu/papers/Verma.pdf>

funksjon påvirker prediksjonen, og kan avdekke skjulte mønstre eller feil i datasettet og modellen, inkludert eventuell usaklig eller ulovlig forskjellsbehandling.

KI-prosjektene viser til at en menneskelig beslutningstaker fungerer som kontrollinstans. Alle virksomhetene i caseundersøkelsene vurderer at KI-systemene ivaretar likhetsprinsippet på en god måte. De legger til grunn at grupper og individer behandles likt ved at de er gjenstand for det samme KI-systemet og beregningene. Videre sikres likebehandling ved at saksbehandlere alltid er involvert og kan gjennomgå resultatet fra KI-systemene og avdekke eventuelle feil. Virksomhetene vurderer at mulige skjevheter i utfallet av KI-systemet ville få små konsekvenser siden KI-systemet ikke tar autonome beslutninger. Arbeidstilsynet viser videre til at modellen brukes innenfor utvalgte målgrupper av *like* virksomheter. Når virksomheter velges ut for kontroll, gjøres dette innenfor disse målgruppene. Derfor vil ikke kontrollutvalget være skjevt, selv om det skulle være skjevheter i modellen.

6.4.4 KI-systemene er i varierende grad tilrettelagt for åpenhet og forklarbarhet

Transparens består av åpenhet og forklarbarhet. Åpenhet innenfor kunstig intelligens handler om å gi relevant informasjon til interessenter, enten de er interne eller eksterne brukere. Forklarbarheten til KI-systemer handler om at man kan forklare hvordan en KI-modell tar en beslutning, og hvorfor den har kommet frem til et resultat.²³⁵ Analyse av kategorien åpenhet består blant annet av vurderinger av hvordan KI-modellens beslutninger og resultater kan forklares, dokumentasjon av utvikling, veivalg og kvaliteten på KI-modellen og om det er gitt informasjon om bruken av KI modellen til interne og eksterne brukere.

Resultatene for kategorien transparens viser at det er stor variasjon blant KI-systemene. Arbeidstilsynet har forankret transparens som et delmål i konseptforslaget, mens fradragsmodellen og KI-systemet for pensjonsutbetalinger som bruker persondata, anser transparens primært som et personvernprinsipp. Undersøkelsen av transparens ved KI-systemene viser at

- det er åpenhet internt i virksomheten om KI-systemenes funksjonsmåte, ytelse og kvalitet
- KI-systemene varierer når det gjelder å dokumentere utvikling av og veivalgene for KI-modellen
- det kan i varierende grad forklares hvordan KI-modellen tar en beslutning, og hvorfor den har kommet frem til et resultat
- det gis overordnet informasjon til eksterne brukere om generell bruk av KI i virksomhetene

KI-systemene er tydelig beskrevet i prosjektdokumentasjonen, deriblant hvordan KI-systemet fungerer, og hvilke data og algoritmer som brukes. Det varierer hvor godt prosjektene har dokumentert utviklingsprosessen og grunnlaget for beslutninger om for eksempel en algoritmetype brukt i KI-systemet.

Avveininger mellom forklarbarheten, treffsikkerheten og ytelsen til KI-modellene er lite dokumentert. Arbeidstilsynets KI-system, Prediksjonsindekstjenesten (PIT), ivaretar balansen mellom forklarbarhet og ytelse i hele utviklingsforløpet. PIT brukte først en enkel og forklarbar KI-modell, og legger til rette for god forklarbarhet når tjenesten går over til en mer komplisert modell i 2023 gjennom beregning av SHAP-verdier^{236,237}. Disse kan brukes for å estimere betydning av variabler for KI-systemets utfall i enkelttilfeller, basert på en forklaringsmodell utledet fra spillteori²³⁸. Også skatteetaten beregner SHAP-verdier i forbindelse med fradragsmodellen og har i tillegg begynt med å teste andre forklaringsmetodikker. Statens pensjonskasse bruker en modelltype som i utgangspunktet er delvis

²³⁵ Digitaliseringsdirektoratet. (u.å.). *Åpenhet og kunstig intelligens*. Hentet 11. april 2024 fra <https://www.digdir.no/kunstig-intelligens/apenhet-og-kunstig-intelligens/4581>

²³⁶ Lundberg, S. M. & Lee, S. I. (2017). A unified approach to interpreting model predictions. I I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan & R. Garnett (Red.), *Advances in Neural Information Processing Systems* 30.

²³⁷ Digitaliseringsdirektoratet. (u.å.). *Åpenhet og kunstig intelligens*. Hentet 11. april 2024 fra <https://www.digdir.no/kunstig-intelligens/apenhet-og-kunstig-intelligens/4581>; SHAP (SHapley Additive exPlanations)

²³⁸ Variablene som inngår i KI-systemet, er sett på som spillere i et kooperativt spill, og KI-systemets utfall er resultat av spillet.

forklarbar, men transparens avhenger av modellparametere og faktisk forklaringssevne med valgte parametere, og dette er ikke dokumentert.

For leverandørsystemet i fluorescensmikroskopet er det ikke kjent hvilken type algoritme²³⁹ KI-systemet bruker. Ingen informasjon om utviklingsprosessen er tilgjengelig, og det er ingen måte å forklare KI-systemets resultater på. Anskaffelsen inneholdt et ønske om «gjennomsiktighet» i programvaren, men dette ble ikke fulgt opp videre i anskaffelses- eller valideringsprosessen.

To av virksomhetene som ble undersøkt i caseundersøkelsen, har lagt ut informasjon til eksterne brukere om at de bruker kunstig intelligens i sin saksbehandling. For eksempel har skatteetaten lagt ut generell informasjon på sine nettsider om at de bruker kunstig intelligens blant annet i arbeidet med skattemeldingen.²⁴⁰ Åpenhet om KI-systemene til eksterne brukere vurderes annerledes av virksomhetene. Virksomhetene mener det ikke er aktuelt å gi informasjon til eksterne om bruk av KI-systemet og hvordan det fungerer, blant annet fordi formålet med bruken og konsekvensene for brukerne er små. Skatteetaten mener at forklarbarheten til KI-systemet til skattebetalere ikke er et lovkrav for dem, og at det kan være i konflikt med deres kontrollformål.

6.4.5 Robusthet og sikkerhet er i stor grad ivaretatt i KI-systemene

Kategorien robusthet og sikkerhet har blitt analysert blant annet ut fra vurderinger av beskyttelse mot feil, manipulering, rutiner for logging og overvåkning, bruker- og tilgangsstyring, og rammer for databehandling.

Undersøkelsen av robusthet og sikkerhet for de fire KI-systemene viser at

- kravene til IT-sikkerheten er ivaretatt, både i policy og praksis
- dokumentasjonen av datamodellen og dataflyten er bedret ved overgang til skybaserte løsninger
- kvaliteten på data sjekkes og vurderes

Ut fra dokumentasjonen fremstår det som at KI-prosjektene har etablert en prosedyre som ivaretar de grunnleggende kravene til IT-sikkerhet, som tilgangsstyring, logging, endringshåndtering og back-up for KI-systemet. Praksis synes i stor grad å være dokumentert i relevante policyer, rutinebeskrivelser og mer driftsrettede arbeidsdokumenter.

KI-systemene har ved overgang til skyløsninger forbedret dokumentasjonen av datamodellen og dataflyten, samt sporbarhet og overvåking. KI-systemets resultater blir overvåket, enten i tett samarbeid med brukerne av modellen eller med automatiserte tester av ytelse mot måltall. I tillegg gis det varsler om resultatene, og informasjon om kjøring av dataprosessering blir logget. En slik overvåking av KI-systemets resultater forebygger at KI-systemets ytelse endrer seg ubemerket over tid, noe som er en kjent risiko i KI-systemer som lærer av data. Kvaliteten på data som brukes i KI-systemene sjekkes og vurderes av virksomhetene. De selvutviklede KI-systemene bruker stort sett egne data som kommer fra kjernevirksomheten, slik at kontroll av datakvaliteten ligger delvis utenfor selve KI-systemet.

Regelmessige oppdateringer er planlagt, og prosjektteamene jobber delvis med oppdaterte treningsdata som reduserer risikoen for at modeller selvforsterkes. Vedlikeholds- og oppdateringsavtalen for leverandørsystemet i mikroskopet inkluderer også KI-systemet. Mikroskopet er videre tilpasset lokale prøver av ulik kvalitet, men tilpasnings- og valideringsprosessen er ikke tydelig separert og basert på de samme prøvene.

²³⁹ Til denne typen anvendelse er det mest sannsynlig med en avansert «sort boks» modell for bildegjenkjenning og klassifisering

²⁴⁰ Skatteetaten. (2023, 5. desember). *Slik kan vi bruke kunstig intelligens på en sikker og ansvarlig måte.*

<https://www.skatteetaten.no/presse/nyhetsrommet/slik-kan-vi-bruke-kunstig-intelligens-pa-en-sikker-og-ansvarlig-mate/>

6.4.6 Forskjeller i hvordan ansvarlig KI ivaretas ved anskaffelser og selvutvikling av KI-systemer

KI-systemet for pensjonsutbetalinger i Statens pensjonskasse, fradragsmodellen i skatteetaten og prediksjonsindekstjenesten i Arbeidstilsynet er selvutviklet av virksomheten, mens KI-systemet i fluorescensmikroskopet i St. Olavs hospital HF er del av en anskaffelse av medisinsk-teknisk utstyr, med innebygd KI som en del av løsningen. Kravene til ansvarlig bruk av kunstig intelligens er de samme, men undersøkelsen viser noen forskjeller mellom hvordan selvutviklede KI-systemer og leverandørsystemer ivaretar ansvarlig KI:

- Prosessen for utvikling og anskaffelse av KI er forskjellig, og dette påvirker rolle- og ansvarsfordelingen og hvordan det arbeides med å ivareta ansvarlig KI.
- Kravstilling i anskaffelse og dokumentasjon fra leverandør er avgjørende.

Virksomhetene som selv har utviklet KI-systemer, har begynt i det små med mindre utviklingsprosjekter med lav risiko for å høste erfaringer. Noen har over tid gradvis videreutviklet og forbedret KI-systemet og KI-modellen som ligger til grunn. Som del av KI-utviklingen har disse også arbeidet med å konkretisere de etiske prinsippene for ansvarlig KI og vurdere relevansen for sitt KI-systemet. Det innebærer å gjøre egne analyser av for eksempel vurdering av personvernkonsekvenser og finne ut hvordan man kan måle ytelsen til KI-systemene. Erfaringen med selv å utvikle KI-systemene og ivareta ansvarlig KI kan gi virksomheten bedre kunnskap og læring.

Ved anskaffelser er prosessen annerledes. Da er det leverandøren som utvikler KI-systemet og finner løsninger for å ivareta ansvarlig KI. En uavhengig tredjepart, et meldt organ²⁴¹, undersøker at det medisinsk-tekniske utstyret er i samsvar med gjeldende regulering, og utsteder et sertifikat som produsenten trenger for å kunne CE-merke produktet. Rollen til kunden er å stille relevante krav som ivaretar de etiske prinsippene for ansvarlig KI i leverandørsystemene, og å vurdere hvordan tilbudene fra leverandørene oppfyller kravene. Mye av ivaretakelsen av ansvarlig KI handler da om dokumentasjonen fra leverandørene og kundens evne til å vurdere om prinsippene er ivaretatt tilstrekkelig i leverandørsystemet. Caseundersøkelsen av anskaffelsen av fluorescensmikroskopet ved St. Olavs hospital HF viser at det er stilt enkelte krav til ytelsen, kvaliteten og informasjonssikkerheten til KI-systemet. Anskaffelsen inneholdt et ønske om at programvaren skulle ha «gjennomsiktighet», men dette ble ikke fulgt opp videre i anskaffelses- eller valideringsprosessen. Utover dette er det ikke stilt krav til KI-systemet om åpenhet og forklarbarhet eller likebehandling. Dette fører blant annet til at virksomheten ikke har informasjon fra leverandøren om hvilke algoritmer som er brukt, og hvilke treningsdata KI-modellen baserer seg på. Videre er det lite dokumentasjon av det innebygde KI-systemet utover kravene som er stilt om CE-merking og samsvarserklæring²⁴² av det medisinsk-tekniske utstyret. Ifølge St. Olavs hospital HF skyldes dette at KI kun utgjorde en liten del av det utstyret som skulle anskaffes.

²⁴¹ Slike meldte organer finnes i ulike bransjer. For medisinsk utstyr i Norge har DMP ansvar for utpeking og tilsyn med meldt organ.

²⁴² Samsvarserklæringen er produsentens skriftlige erklæring om at utstyret oppfyller kravene som gjelder for utstyret i regelverket. Jf. Direktoratet for medisinske produkter. (2021, 25. mai). *Samsvarserklæring*. <https://www.dmp.no/medisinsk-utstyr/utvikling-og-produksjon/samsvarsvurdering/samsvarserklaring>

7 Tilrettelegger Digitaliserings- og forvaltningsdepartementets samordning for ansvarlig bruk av KI i statlige virksomheter?

7.1 Digitaliserings- og forvaltningsdepartementets ansvar og sektorprinsippet

Digitaliserings- og forvaltningsdepartementet har det overordnede ansvaret for å tilrettelegge og samordne regjeringens IT-politikk og å være en pådriver for IT-utviklingen i offentlig sektor. Departementet skal blant annet identifisere felles utfordringer og vurdere behovet for å iverksette tverrsektorielle tiltak og vurdere aktuelle virkemidler. En del av denne jobben er å utvikle strategier for digitalisering, inkludert kunstig intelligens, og utarbeide meldinger til Stortinget på området.²⁴³ Digitaliserings- og forvaltningsdepartementet har også ansvaret for å implementere EUs nye KI-forordning i norsk rett.

Bruk av kunstig intelligens i staten er et tverrsektorielt tema. Men sektorprinsippet står imidlertid sterkt i norsk forvaltning, og Digitaliserings- og forvaltningsdepartementet henviser også til sektorprinsippet. De legger til grunn i sin rolle- og ansvarsutøvelse at fagdepartementene er ansvarlig for måloppnåelse og resultater innenfor digitalisering og kunstig intelligens i sine sektorer. Digitaliserings- og forvaltningsdepartementet viser til i intervju at departementet er avhengig av andre for å nå egne mål. Digitaliserings- og forvaltningsdepartementet har ingen fullmakter overfor sektordepartementene, og er ifølge departementet i hovedsak henvist til å ta i bruk såkalte pedagogiske virkemidler. Flere av virkemidlene og tiltakene for å gjennomføre IT-politikken sorterer også under sektordepartementene. For eksempel ligger ansvaret for IT-infrastruktur for forskning under Kunnskapsdepartementet.²⁴⁴

I kapittel 5 omtales de mest sentrale utfordringene statlige virksomheter og heleide selskaper møter i sitt arbeid med å utvikle og ta i bruk kunstig intelligens. I dette kapitlet vil vi se på aktivitetene som departementet gjennomfører for å tilrettelegge og samordne, og hvordan de har fylt pådriverrollen gitt virkemidlene de har til rådighet. Vi vil også se på hva Digitaliserings- og forvaltningsdepartementets underliggende virksomheter har gjort. Særlig viktige her er Digitaliseringsdirektoratet og Datatilsynet.

7.2 Regjeringens strategi for kunstig intelligens og oppfølgingen av denne

Regjeringens strategi for kunstig intelligens ble offentliggjort i januar 2020 med daværende Kommunal- og distriktsdepartementet (nå Digitaliserings- og forvaltningsdepartementet) som ansvarlig departement. Denne strategien er et sentralt virkemiddel for bruk av kunstig intelligens i offentlig sektor. Strategien er omtalt blant annet i Prop. 1 S (2020–2021) for Kommunal- og distriktsdepartementet. Der viser departementet til at strategien angir retningen for Norges satsing på kunstig intelligens, og at det er nødvendig å følge utviklingen av KI tett. Departementet sier også at de vil følge opp regjeringens intensjon med å utvikle retningslinjer og etiske prinsipper for bruk av kunstig intelligens, i samarbeid med de nordiske landene og EU.²⁴⁵ Digitaliserings- og forvaltningsdepartementet presiserer i intervju at den nasjonale KI-strategien ikke legger opp til en satsing på kunstig intelligens, og at KI ikke er et mål i seg selv.

²⁴³ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

²⁴⁴ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

²⁴⁵ Prop. 1S (2019–2020) Kommunal- og distriktsdepartementet, s. 115.

Norge var noe senere ute med en nasjonal KI-strategi enn de andre nordiske landene. Finland fikk en KI-strategi i 2017, mens Sverige fulgte i 2018 og Danmark i 2019.²⁴⁶

KI-strategien inneholder konkrete tiltak og aktiviteter som regjeringen sier den vil gjennomføre. Tiltakene er fordelt på en rekke områder, som dataforvaltning, regelverk, infrastruktur, forskning og kompetanse. Det er også et eget kapittel om «ansvarlig og pålitelig kunstig intelligens». Det varierer hvor konkrete tiltakene er. De fleste tiltakene er konkrete i form av formuleringer som «legge frem ...», og «etablere ...» og lignende. Andre tiltak er mindre forpliktende, for eksempel: «Regjeringen vil *vurdere* virkemidler som kan gjøre det enklere for bransjer å dele data» (vår utheving).

Digitaliserings- og forvaltningsdepartementet og Digitaliseringsdirektoratet viser i intervju til at det ikke er gjennomført noen formell evaluering av KI-strategien. Direktoratet viser til at strategien kom uten handlingsplan, noe som gjør det vanskelig å evaluere, ifølge Digitaliseringsdirektoratet. Direktoratet har derfor heller ikke evaluert gjennomføringen av de enkelte punktene i strategien hvor regjeringen sier de vil gjøre noe.

Departementet viser i intervju til at det imidlertid holder løpende oversikt over hvilke tiltak som er iverksatt, og hva som er status for disse. Riksrevisjonen har fått se oversikten. Departementet vurderer at de fleste tiltakene er gjennomført eller påbegynt. Enkelte tiltak er også vurdert til å ikke lenger være aktuelle. Status per november 2023 viser at blant alle de 51 tiltakene/aktivitetene omtalt i KI-strategien har 21 tiltak fått status som gjennomført, 12 har status som «startet», mens resten er angitt enten som «løpende oppgaver» eller som at de blir fulgt opp i andre strategier/prosesser. Regjeringen har dermed gjennomført en rekke tiltak, der flere er rettet mot utfordringene omtalt i kapittel 5.

Oversikten viser at Digitaliserings- og forvaltningsdepartementet og Digitaliseringsdepartementet er tillagt ansvaret for over halvparten av tiltakene. Flere av tiltakene Digitaliserings- og forvaltningsdepartementet har ansvaret for, blir fulgt opp gjennom arbeid med andre dokumenter, for eksempel ble det utarbeidet en ny stortingsmelding om digitalisering (jf. Meld. St. 22 (2020–2021) *Data som ressurs — Datadrevet økonomi og innovasjon*). Kunnskapsdepartementet har også ansvaret for flere oppgaver, mens for eksempel Helse- og omsorgsdepartementet og Justis- og beredskapsdepartementet har ansvaret for enkelte av tiltakene.

7.2.1 Fremskaffelse av styringsinformasjon

Digitaliserings- og forvaltningsdepartementet har ikke kartlagt bruken av kunstig intelligens i offentlig sektor. Digitaliserings- og forvaltningsdepartementet viser til at Digitaliseringsdirektoratet og Norwegian Artificial Intelligence Research Consortium (NORA.ai)²⁴⁷ samarbeider om å holde oversikt over bruken og utviklingen av KI i Norge. NORA.ai har gjennom tilgjengelig informasjon og direkte kontakt med ulike aktører innhentet opplysninger om KI-prosjekter i offentlig sektor i Norge. Dette er ikke en komplett oversikt over alle norske KI-initiativer. Digitaliseringsdirektoratet og NORA.ai samarbeider nå om denne listen, og den bygges ut fortløpende. Digitaliseringsdirektoratet så behovet for å få oversikt over KI-bruken i offentlig sektor som et ledd i utviklingen av veiledningsmaterialet. Digitaliserings- og forvaltningsdepartementet peker på at oppdraget til Digitaliseringsdirektoratet blant annet er å følge med på utviklingen.²⁴⁸ Det er imidlertid ikke lagt opp til noen systematisk oppdatering av denne kartleggingen. Derfor finnes det ingen målinger av hvordan bruken av kunstig intelligens i offentlig sektor har utviklet seg. Kartleggingen ble ikke gjennomført på oppdrag av Digitaliserings- og forvaltningsdepartementet og har heller ikke blitt brukt til å evaluere tilstanden til kunstig intelligens på

²⁴⁶ OECD. (2023). *The State of Implementation of the OECD AI principles four years on* (OECD Artificial Intelligence Papers, October 2023 No. 3). OECD Publishing.

²⁴⁷ NORA er et norsk samarbeid innen AI, maskinlæring og robotikk mellom åtte universiteter, fem høyskoler og fem forskningsinstitutter. NORA arbeider for å styrke norsk forskning, utdanning og innovasjon på disse feltene.

²⁴⁸ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

noen måte. Verken Digitaliseringsdirektoratet eller Digitaliserings- og forvaltningsdepartementet har satt i verk noen tiltak som følge av kartleggingsmaterialet.²⁴⁹

Digitaliserings- og forvaltningsdepartementet viser i intervju til at det ikke sitter med et systematisk kunnskapsgrunnlag om hvor potensialet for KI er størst, og viser tilbake til at det er den enkelte sektoren som er ansvarlig for å vurdere dette potensialet og utvikle og ta i bruk KI i egen sektor.

Digitaliserings- og forvaltningsdepartementet ønsket å få en oversikt over investeringsnivå for KI i Norge, fortrinnsvis som en del av EUs oversikt, eller i det minste i form av sammenlignbare tall.²⁵⁰ Departementet var i 2021 i kontakt med SSB om dette, men det lyktes ikke å få noen konkrete svar fra SSB, og saken ble ikke videreført.

7.2.2 En gjennomgang av regelverket

I kapittel 5 ble det vist at flere statlige virksomheter og heleide selskaper mener at regelverk og personvern er komplisert i arbeidet med kunstig intelligens.

I KI-strategien fra 2020 trekkes det frem flere aspekter ved hvorfor det trengs et digitaliseringsvennlig regelverk for utvikling og bruk av KI. I strategien vises det også til flere tiltak for å sikre et digitaliseringsvennlig regelverk. Et av tiltakene var å gjennomgå og vurdere regelverk som er til hinder for hensiktsmessig og ønsket bruk av kunstig intelligens i offentlig og privat sektor. I henhold til Digitaliserings- og forvaltningsdepartementets oversikt skulle dette tiltaket følges opp av departementet selv.

Digitaliserings- og forvaltningsdepartementet kategoriserer dette fortsatt som «startet» i 2023, og Digitaliserings- og forvaltningsdepartementet kommenterer at planen for oppfølging av tiltaket er å «kartlegge hvilke sektorer som har størst utfordringer, og etablere samarbeid om kartleggingen med aktuelt sektordepartement».

Digitaliserings- og forvaltningsdepartementet²⁵¹ viser i intervju til at departementet hittil ikke har tatt initiativ til å utvikle ytterligere regelverksrettede virkemidler, og legger til grunn at det er sektordepartementene og virksomhetene som må vurdere behovet for dette. Digitaliserings- og forvaltningsdepartementet mener også at strategien ikke legger opp til at man skal gjennomgå og vurdere regelverket i en bred kontekst. I dag har ikke departementet noen oversikt over hjemmelsgrunnlaget for å behandle personopplysninger i de ulike sektorene. Som forberedelse til KI-forordningen har regjeringen imidlertid satt ned en hurtigarbeidende arbeidsgruppe for å vurdere behov for nasjonal regulering utover KI-forordningen (se neste avsnitt). Digitaliserings- og forvaltningsdepartementet opplyser om at denne juridiske arbeidsgruppen er videreført for å arbeide med andre aktuelle tverrgående problemstillinger.

Digitaliserings- og forvaltningsdepartementet viser i intervju til at det kan se ut som kompleksiteten og vanskelighetsgraden i personvernregelverket brukes som en unnskyldning til å ikke gjøre de interne vurderingene som er nødvendige å gjøre for å kunne ta i bruk kunstig intelligens.

Digitaliserings- og forvaltningsdepartementet mener at dersom det skal gjøres noe med hjemmelsproblematikken knyttet til databruk eller kunstig intelligens, må dette gjøres i særlovene i de ulike sektorene. Departementet viser til at personopplysningsloven stiller krav til at hjemler må være presise og tydelige, og at det derfor er vanskelig å lage generelle hjemler som er klare nok.

²⁴⁹ Verifisert referat fra intervju med Digitaliserings- og forvaltningsdepartementet 14. mars 2024.

²⁵⁰ Korrespondanse med Riksrevisjonen, 9. mars 2023.

²⁵¹ Daværende Kommunal- og distriktsdepartementet, 14. september 2023.

Departementet viser til at erfaringer med hvordan problemstillinger med KI og data kan løses i regelverket, kan deles i den interdepartementale arbeidsgruppen for KI.²⁵²

Digitaliserings- og forvaltningsdepartementet viser i intervju til at det nå ser at kompetanse- og veiledningsbehovet er en felles problemstilling. Sektorlovgivningen er veldig forskjellig, og det blir en avveining i hvor langt departementet skal gå inn i den konkrete sektoren, og hvor relevante den ene sektorens problemstillinger er for de andre sektorene. Digitaliserings- og forvaltningsdepartementet undersøker slike spørsmål, og noe av dette er temaet for den interdepartementale arbeidsgruppen for KI. Departementet viser til at det alle departementer er representert i denne arbeidsgruppen, som har som formål å skape bedre koordinering og informasjonsutveksling på KI-området.

Innføringen av EUs KI-forordning og regulering av KI

Som omtalt ovenfor er det Digitaliserings- og forvaltningsdepartementet som har ansvaret for å innlemme EUs KI-forordning i norsk rett. I denne forbindelse satte regjeringen ned en hurtigarbeidende arbeidsgruppe som fikk mandat til å utarbeide en plan for dette og å vurdere nasjonale behov for regulering (sivil sektor) utover EUs forordning. Arbeidet ble ferdigstilt som planlagt i november 2023. Selv om det er Digitaliserings- og forvaltningsdepartementet som har ansvaret for dette arbeidet, var det Justis- og beredskapsdepartementet som ledet arbeidsgruppen. Arbeidsgruppen besto ellers av representanter fra fem andre departementer, inkludert Digitaliserings- og forvaltningsdepartementet.²⁵³

Arbeidsgruppen var særlig opptatt av å hente inn innspill fra andre offentlige instanser og fra eksperter på området både i akademia og i privat sektor, særlig om spørsmålet om behovet for nasjonale regler. Det generelle inntrykket var at de inviterte kompetansemiljøene hadde mange av de samme innspillene og i hovedsak var det bred enighet blant medlemmene i arbeidsgruppen om de overordnende anbefalingene og forslagene til tiltak som ble lagt frem.²⁵⁴

Et spørsmål som arbeidsgruppen diskuterte, var tilgangen på norske data. For å utvikle og bruke KI trenger man ofte tilgang på allerede eksisterende data. Adgangen til å bruke eksisterende data til å trene opp KI-modeller må skje i henhold til det generelle personvernregelverket og et sektorspesifikt regelverk. I dag er det generelt liten tilgang til å trene KI-modeller basert på eksisterende data. Lovavdelingen i Justis- og beredskapsdepartementet har ikke fått noen henvendelser om at dette spørsmålet bør vurderes på et generelt grunnlag.²⁵⁵

Et annet spørsmål som rapporten tar opp, er risiko for diskriminering med kunstig intelligens. Rapporten peker på at forbudene mot diskriminering i likestillings- og diskrimineringsloven og i Grunnloven er teknologinøytrale og vil følgelig også gjelde for kunstig intelligens. Rapporten viser videre til at KI-forordningen stiller flere krav til høyrisiko KI-systemer, som blant annet skal minimere risiko for diskriminering. Handlingsrommet for nasjonal regulering er større dersom man ønsker å regulere KI-systemer med lavere risiko, ifølge rapporten. Rapporten konkluderte i november 2023 med at det ikke er identifisert behov for ytterligere regulering i norsk lov som følge EUs KI-forordning.²⁵⁶ I mai 2024 lanserte Likestillings- og diskrimineringsombudet en utredning som konkluderte med at det er behov for en gjennomgang av likestillings- og diskrimineringsloven for å sikre at algoritmisk diskriminering kan fanges opp på en bedre måte enn i dagens lovverk.²⁵⁷

Den hurtigarbeidende arbeidsgruppen sier i sin rapport at det er et vel så stort behov for veiledning og større forståelse av hvordan allerede eksisterende regelverk skal anvendes for KI, som behov for mer

²⁵² Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

²⁵³ Verifisert referat fra intervju med Justis- og beredskapsdepartementet 5. februar 2024.

²⁵⁴ Verifisert referat fra intervju med Justis- og beredskapsdepartementet 5. februar 2024.

²⁵⁵ Verifisert referat fra intervju med Justis- og beredskapsdepartementet 5. februar 2024.

²⁵⁶ Regulering av kunstig intelligens, arbeidsgrupperapport s. 28

²⁵⁷ V. B. Strand (2024), «Algoritmer, kunstig intelligens og diskriminering. En analyse av likestillings- og diskrimineringslovens muligheter og begrensninger»

nasjonal regulering. Arbeidsgruppen påpeker at usikkerhet om regelverket kan få negative konsekvenser:

Tvil rundt gjeldende regelverk kan være hemmende for nye prosjekter og forskning på feltet, og kan fungere som en brems for utviklingen.²⁵⁸

Gruppen forventer at veiledningsbehovet blir enda større når KI-forordningen trer i kraft. Arbeidsgruppen peker også på at det åpenbart er behov for en kraftig kompetanseheving på KI i alle sektorer for at ikke tvil og usikkerhet skal hindre god bruk av KI.

Arbeidet med ny forvaltningslov

Enkelte av de intervjuede virksomhetene viser til arbeidet med ny forvaltningslov, og at dette kan påvirke spørsmålet om i hvilken grad reguleringer på personvernområdet skal skje gjennom et generelt regelverk, som i forvaltningsloven, og hva som skal reguleres gjennom sektorlovene. Arbeids- og integreringsdepartementet og Nav viser begge i intervju til at spørsmålet om hvordan den nye forvaltningsloven utformes, vil påvirke arbeidet med å utvikle egne sektorlovgivninger.²⁵⁹

Justis- og beredskapsdepartementet har ansvaret for forvaltningsloven. Forvaltningslovutvalget leverte sin utredning, NOU 2019: 5, til Justis- og beredskapsdepartementet i mars 2019. Utredningen peker på utfordringer ved helautomatisert saksbehandling og kommunikasjon samt utfordringer med personvern i offentlig digital tjenesteproduksjon.²⁶⁰

Justis- og beredskapsdepartementet påpeker i intervju at forholdet mellom et generelt regelverk og et sektorspesifikt regelverk vil være et aktuelt spørsmål på flere områder, også for kunstig intelligens. Forvaltningsloven skal være en generell lov som inneholder generelle regler om behandlingsmåten i offentlig forvaltning. Spesifikke forhold utover det generelle skal primært dekkes av særlover. Man kan innlemme flere bestemmelser i forvaltningsloven, som hjemmel til behandling av persondata, men det vil uansett ofte være behov for spesifikke bestemmelser innenfor ulike sektorer som best håndteres i særlovgivningen. Hvis det åpnes generelt opp for bruk av persondata i forvaltningsloven, kan det kreve at særlovene må begrense bestemmelsene i forvaltningsloven på enkelte områder. Særlovene kan inneholde flere bestemmelser om kunstig intelligens, og til en viss grad kan noe ivaretas i forvaltningsloven. Det vil imidlertid alltid være et visst spenn mellom hva som reguleres generelt gjennom forvaltningsloven, og hva som skal reguleres i særlovgivningen innenfor den enkelte sektoren.²⁶¹

Det er også nevnt i KI-strategien at «det skal stilles krav om gjennomsiktighet og etterrettelighet i nye digitale løsninger for offentlig forvaltning der KI er en del av løsningen».²⁶² Om planer for oppfølging sier Digitaliserings- og forvaltningsdepartementet at dette vil «måtte komme inn som en del av arbeidet med ny forvaltningslov.»

7.2.3 Datatilsynets regulatoriske sandkasse for å forstå kravene til personvern bedre

Et annet av tiltakene i den nasjonale KI-strategien var å etablere en regulatorisk sandkasse på personvernområdet. Den regulatoriske sandkassen var ment å sikre bedre forståelse av kravene til personvern. Et annet formål var å bidra til å korte ned tiden fra utvikling og testing til faktisk utrulling av KI-løsninger ved å spre kunnskap om behandling av personopplysninger i forbindelse med KI-

²⁵⁸ Regulering av kunstig intelligens, arbeidsgrupperapport s. 39.

²⁵⁹ Verifisert referat fra intervju med Nav 8. november 2023 og med Arbeids- og inkluderingsdepartementet 18. desember 2023.

²⁶⁰ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

²⁶¹ Verifisert referat fra intervju med Justis- og beredskapsdepartementet 5. februar 2024.

²⁶² Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*. <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/> s. 6.

systemer. Løsningene som settes i drift etter å ha deltatt i sandkassen, kan dermed fungere som foregangseksempler.

Datatilsynet fikk i oppgave av Digitaliserings- og forvaltningsdepartementet²⁶³ å opprette den regulatoriske sandkassen, og ordningen kom på plass i november 2020. Norge var tidlig ute med å etablere en slik sandkasse på KI-området²⁶⁴, og lignende initiativer har blitt igangsatt i en rekke land. KI-sandkassen ble evaluert i 2023.²⁶⁵

Datatilsynet viser i intervju til at den regulatoriske sandkassen er ett av flere virkemidler tilsynet har til disposisjon for å nå sine overordnede mål om å sikre et godt personvern for alle og medvirke til at enkeltpersoners rettigheter ikke blir krenket gjennom bruk av opplysninger som kan knyttes til dem.

Prosjektene kan søke om å delta i sandkassen, og Datatilsynet gir deltakerne dialogbasert veiledning om personvernregelverket. Datatilsynet viser i intervju til at de best kan vurdere hvordan de kan bistå med veiledning, ut fra søknadene og i arbeidet med prosjektene i sandkassen. Datatilsynet legger vekt på at prosjektene skal være samfunnsnyttige, kunne være nyttige for langt flere enn de virksomhetene som deltar i sandkassen, skal vurdere nye aspekter knyttet til personvernet og skal prøve ut personvernvennlige løsninger. Frem til utgangen av 2023 var sandkassen kun rettet mot KI-prosjekter. Fra 2024 har mandatet til sandkassen blitt utvidet til å gjelde personvernvennlig innovasjon og digitalisering.

Faktaboks 2 Regulatorisk sandkasse

En regulatorisk sandkasse er et kontrollert testmiljø for virksomheter som vil eksperimentere med nye produkter, teknologier og tjenester under oppfølging av myndighetene. På denne måten vil virksomhetene få økt forståelse for de regulatoriske kravene (kravene som stilles i lovgivningen), og myndighetene vil få økt forståelse for nye teknologiske løsninger. Det tette samarbeidet mellom aktører og myndigheter i sandkassen vil bidra til at risiko og problemstillinger lettere vil kunne identifiseres og løsninger skisseres.

Datatilsynets regulatoriske sandkasse gir råd om kravene personvernreglene stiller til utvikling og bruk av kunstig intelligens som tar i bruk personopplysninger.

Kilde: Datatilsynet. (u.å.). *Sandkassesiden*. Hentet 12. februar 2024 fra <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/>

Aktivitet i sandkassen

Sandkassen har per mars 2024 hatt fem søknadsrunder (sist i desember 2023).

Tabell 4 oppsummerer antallet søknader og antallet prosjekter som har blitt tatt opp i ordningen.

²⁶³ Datatilsynet er i dag underlagt Digitaliserings- og forvaltningsdepartementet.

²⁶⁴ Storbritannia var det første landet i verden som opprettet en slik sandkasse, våren 2020, jf. Datatilsynet. (2020). *Starter regulatorisk sandkasse for utvikling av ansvarlig kunstig intelligens*. <https://www.datatilsynet.no/aktuelt/aktuelle-nyheter-2020/regulatorisk-sandkasse-for-utvikling-av-ansvarlig-kunstig-intelligens/>

²⁶⁵ Agenda Kaupang. (2023). *Evaluering av Datatilsynets sandkasse for kunstig intelligens*. Datatilsynet. <https://www.datatilsynet.no/contentassets/41e268e72f7c48d6b0a177156a815c5b/rapport-datatilsynet--evaluering-av-sandkassa.pdf>

Tabell 4 Søknader og deltakelse i den regulatoriske sandkassen, 2020–2024

	Søkere totalt	Privat sektor	Offentlig sektor	Antall prosjekter tatt opp
2021 ²⁶⁶	46	29	17	8
2022	11	8	3	3
2023 ²⁶⁷	26	18	8	2
Totalt	83	55	28	13

Kilder: Datatilsynets årsrapporter og nyhetsmeldinger.

Tabellen viser at det er vesentlig færre søkere i 2022 og 2023 enn det var i 2021. I 2023 var det to søknadsrunder, og i den første søknadsrunden var det kun seks prosjekter som søkte. Datatilsynet viser i intervju til at de ønsker at flere skal søke om å få delta i sandkassen, men at det av ressurshensyn ikke har promotert ordningen andre steder enn på egne nettsider. Totalt har 83 offentlige og private virksomheter søkt om å delta i sandkassen. 13 prosjekter har blitt tatt opp, hvorav 7 er fra offentlig sektor. Per mars 2024 er det gjennomført 11 prosjekter. Av disse har 5 vært prosjekter i staten – henholdsvis fra Nav, Akershus universitetssykehus HF, Helse Bergen HF, Norges vassdrags- og energidirektorat og PolitiHøgskolen.

Overordnede erfaringer med sandkassen

I 2023 ble den regulatoriske sandkassen eksternt evaluert.²⁶⁸ Evalueringen viste at sandkassen er et godt tiltak, men at det også er noen utfordringer. Evalueringen peker blant annet på

- betydningen av å få flere til å søke om å få delta i ordningen
- at Datatilsynet bør se på hvordan erfaring og kunnskap kan deles mer bredt
- at Datatilsynet i større grad bør ivareta teknologiske spørsmål i prosjektene

Datatilsynet trekker frem i intervju at det særlig ser nytten av rådene som blir utarbeidet i det enkelte sandkasseprosjekt, av læringen prosjektene gir Datatilsynet, og av at erfaringer og læringen fra det enkelte prosjekt tilgjengeliggjøres for det offentlige. Datatilsynet opplyser at den eksterne evalueringsrapporten i stor grad omfatter funn og vurderinger som er godt kjent for Datatilsynet. Datatilsynet sier videre at det legger stor vekt på rådene i rapporten.

Gjennom intervju med Nav og Akershus universitetssykehus HF har vi sett nærmere på erfaringene deres med sandkassen.

Navs prosjekt i sandkassen

I 2021 deltok Nav i Datatilsynets regulatoriske sandkasse med en KI-modell som skulle predikere utviklingen av sykefraværet og dermed kunne forutse hvilke sykemeldte brukere som ville ha behov for oppfølging gjennom et dialogmøte.²⁶⁹ Konklusjonen fra Datatilsynet var at Nav hadde hjemmel til å bruke en slik KI-modell, men det var *usikkert* om Nav hadde et rettslig grunnlag for å bruke personopplysninger til å utvikle algoritmene (utviklingsfasen).²⁷⁰ Det betyr at man ikke nødvendigvis kan bruke personopplysninger om *noen* personer, innhentet til et bestemt formål, til å utvikle en modell som skal brukes til å forutsi et utfall for *andre* personer. Problemstillingen er ikke spesifikk for Nav. Den vil gjelde alle som vil bruke personopplysninger til å trene en KI-modell, med mindre de har en

²⁶⁶ I 2021 ble det gjennomført to søknadsrunder.

²⁶⁷ I 2023 ble det gjennomført to søknadsrunder. Per mars 2024 har det ikke blitt valgt ut prosjekter etter den siste søknadsrunden i 2023.

²⁶⁸ Agenda Kaupang. (2023). *Evaluering av Datatilsynets sandkasse for kunstig intelligens*. Datatilsynet.

<https://www.datatilsynet.no/contentassets/41e268e72f7c48d6b0a177156a815c5b/rapport-datatilsynet--evaluering-av-sandkassa.pdf>

²⁶⁹ Dialogmøte 1 er en samtale mellom den sykemeldte, arbeidsgiver og eventuelt sykemelder. Samtalen skal gjennomføres innen 7 uker i et hvert sykefraværsløp. Innen uke 26 skal Nav-veileder kalle inn den sykemeldte og arbeidsgiver til dialogmøte 2 der hensikten er at den sykemeldte, arbeidsgiver og eventuelt sykemelder sammen med Nav skal finne muligheter og løsninger som kan bidra til at den sykemeldte kommer raskt tilbake i arbeid. Møtene skal i utgangspunktet gjennomføres for alle sykemeldte, men det kan gjøres unntak dersom et slikt møte åpenbart er unødvendig.

²⁷⁰ Datatilsynet. (2022). *NAV – Sluttrapport fra sandkasseprosjektet med NAV*. <https://www.datatilsynet.no/regelverk-og-verktoy/sandkasse-for-kunstig-intelligens/ferdige-prosjekter-og-rapporter/nav-sluttrapport/>

tydelig hjemmel i egen særlov. Som omtalt i kapittel 5 er mange usikre på hva som vil være lov eller ikke lov i henhold til personvernforordningen. De eksplisitte rådene Datatilsynets sandkasse gir til Nav, vil kun gjelde Navs prosjekt.

Et eksempel på en særlovshjemmel finner vi hos Statens pensjonskasse, som sier at Statens pensjonskasse kan behandle innhentede personopplysninger «for å utvikle og teste IT-systemer».²⁷¹ Alle som ikke har en slik særlovshjemmel, må derfor, i tråd med sektorprinsippet, avklare dette på egen hånd.

Den endelige konklusjonen fra Navs deltakelse i sandkassen var at Nav ikke hadde grunnlag for å gå videre med å utvikle denne KI-modellen. Nav opplyser i intervju at deres juridiske hovedfokus i sandkasseprosjektet var knyttet til bruken av KI. Det rettslige grunnlaget for å utvikle algoritmene ble i liten grad drøftet mens prosjektet pågikk. På bakgrunn av denne konklusjonen viser Nav i intervju til at det har vært en dialog mellom Nav og Arbeids- og inkluderingsdepartementet om det er et behov for et spesifikt rettslig grunnlag for at Nav skal kunne bruke egne og eventuelt andre sektors data til å utvikle og bruke KI. Per i dag (mai 2024) er spørsmålet imidlertid uavklart. Nav må derfor være forsiktige i tilnærmingen til hvordan de skal utvikle og eventuelt bruke kunstig intelligens. Selv om Nav har et uavklart juridisk handlingsrom, har de per i dag likevel tilstrekkelige hjemler til å utvikle KI slik de jobber med dette i dag. Basert på erfaring og arbeid med eksisterende og nye prosjekter må Nav tydeliggjøre det juridiske handlingsrommet. Til dette trenger ikke Nav per i dag noen lovendring, blant annet fordi KI-prosjektene har vært konsentrert innenfor enkeltformål.²⁷² Per i dag har Nav imidlertid ingen produksjonssatte løsninger som baserer seg på egenutviklet KI.²⁷³

I sandkasserapporten til Nav-prosjektet kommer Datatilsynet med noen betraktninger om veien videre:

Lovene som hjemler behandling av personopplysninger, er sjeldent utformet på en måte som åpner for at personopplysningene kan brukes til maskinlæring i utvikling av kunstig intelligens. Det er viktig at lovgiver i tiden fremover legger til rette for utvikling av KI i offentlig sektor innenfor ansvarlige rammer. Dersom NAV skal utvikle modellen videre, vil det være nødvendig med et klart og tydelig supplerende rettsgrunnlag gjennom lovgivning.

Nav opplyser i intervju at Arbeids- og inkluderingsdepartementet ikke har gitt noen føringer til Nav om å gjennomføre noen konkrete KI-prosjekter. Det er tett dialog mellom departementet og Nav om regelverket, men Nav mener det ville vært positivt om Arbeids- og inkluderingsdepartementet i større grad hadde understøttet Navs arbeid med regelverksforståelse og dialog med andre departementer, som Digitaliserings- og forvaltningsdepartementet og Justis- og beredskapsdepartementet.²⁷⁴ Arbeids- og inkluderingsdepartementet på sin side viser i intervju til at det først og fremst er Nav selv som må avklare sitt behov for rettslig grunnlag.

Akershus universitetssykehus HFs prosjekt i sandkassen

Akershus universitetssykehus HF viser i intervju til at det opplevde det som positivt å delta i sandkassen. Sykehuset deltok våren 2022 med et prosjekt om utvikling av KI for å analysere EKG²⁷⁵. Sykehuset ønsket å diskutere algoritmeskjevheter og hvordan de kunne sikre at algoritmen ga rettfærdige prediksjoner basert på EKG-målinger. I intervju viser Akershus universitetssykehus HF til at læringen fra sandkasseprosjektet er overførbart til andre prosjekter. Prosjektet løftet frem betydningen av å vurdere diskriminering (bias) og lignende i KI-prosjekter. En veileder (jf. veileder av 8. november 2023) fra Ligestillings og diskrimineringsombudet om hvordan utviklerteam kan vurdere risiko for

²⁷¹ Statens pensjonskasselov. (1949). Lov om Statens pensjonskasse (LOV-1949-07-28-26), § 45 bokstav b sjette ledd.

²⁷² Verifisert referat fra intervju med Nav 8. november 2023.

²⁷³ Verifisert referat fra intervju med Arbeids- og inkluderingsdepartementet 18. desember 2023.

²⁷⁴ Verifisert referat fra intervju med Nav 8. november 2023.

²⁷⁵ Datatilsynet (2023), Sluttrapport fra sandkasseprosjektet til Ahus (EKG AI)

diskriminering i KI-systemer, er blant annet basert på noe av læringen fra dette prosjektet. Sykehuset må imidlertid alltid vurdere muligheten for diskriminering i hvert enkelt prosjekt.

Akershus universitetssykehus HF viser videre til at prosjektet ga sykehuset et tydelig innblikk i hvordan diskriminering blir definert i Norge. Definisjonen består flere faktorer, som kjønn, alder og etnisitet. I helsesektoren har man imidlertid kun tilgang til to av disse kategoriene: alder og kjønn ved fødsel. Man har ikke tilgang til informasjon om kjønnsidentitet, etnisitet og så videre. Før man tar stilling til om en KI-modell er diskriminerende, må man derfor først få på plass en analyse av om helsevesenet generelt er diskriminerende. Da må helsevesenet ha data ut fra definisjonen av diskriminering, men det har ikke helsevesenet i dag. Dette er et spørsmål som må tas på et overordnet nivå, og er ikke ennå diskutert eller avklart.

Digitalisering- og forvaltningsdepartementets vurdering av sandkassen

På spørsmål om den regulatoriske sandkassen har fungert etter hensikten, viser Digitaliserings- og forvaltningsdepartementet til at departementet hadde bedt Datatilsynet om gjennomføring av den eksterne evalueringen, men at departementet ikke selv har undersøkt dette. Departementet viser til at kunnskapen fra enkeltprosjektene ligger tilgjengelig på sandkassens nettsider, og at sandkassen er samfinansiert med flere andre departementer, slik at alle departementene er kjent med at sandkassen finnes. Et forbedringsområde som ble identifisert i den eksterne evalueringen av sandkassen i 2023, var at de må jobbe enda bedre for å spre kunnskapen, erfaringene og kompetansen de bygger i sandkassen. Datatilsynet holder mange foredrag i både nasjonale og internasjonale fora om sandkassen og den jobben de gjør. Datatilsynet sier selv tydelig at det må sørge for at kompetansen det opparbeidet seg i sandkassen, må ut på en enda bedre måte enn det har klart til nå.

7.3 Digitaliserings- og forvaltningsdepartementets pådriverrolle og tiltak utover KI-strategien

I intervju viser Digitaliserings- og forvaltningsdepartementet til at det har en pådriverrolle overfor andre departementer på digitaliseringsområdet, også når det gjelder KI. Departementet viser til at det er avhengig av andre for å nå egne mål, og at det i hovedsak må bruke pedagogiske virkemidler som retningslinjer eller rundskriv og oppfølging av dem. Videre viser departementet til det er opptatt av å identifisere tverrgående behov og å få plass felles løsninger på tvers av sektorene. Departementet arbeider også for å sikre bedre koordinering der det ser behov for dette.

Digitaliserings- og forvaltningsdepartementet viser til at det er politisk målsetting om å ta bruk KI til effektivisering i offentlig sektor og for verdiskaping i samfunnet, og at departementet kan tilrettelegge for at det skjer på en forsvarlig og etisk måte. Sektorprinsippet gjelder også på KI-området, noe som betyr at den enkelte sektoren er ansvarlig for utvikling og bruk av KI innenfor eget ansvarsområde. Departementet viser til at det ofte er koblet på prosessene på KI-området i andre departementer, og vurderer da om de har tatt hensyn til faktorer som Digitaliserings- og forvaltningsdepartementet anser som sentrale – enten prosessene gjelder kunstig intelligens, digitalisering eller andre relevante områder. Pådriverrollen handler ifølge departementet også om å legge til rette, blant annet ved veiledningen gjennom Digitaliseringsdirektoratet og tiltak som Datatilsynets sandkasse.

Digitaliseringsdirektoratets innspill til ny digitaliseringsstrategi, som også skal omhandle KI, peker direktoratet på relativt store utfordringer med den sterke sektoreringen vi har i dag:

Norge har gode forutsetninger for å klare å endre oss slik at vi tar i bruk de mulighetene som ligger i digitalisering. Men det vil kreve strukturelle endringer av hvordan forvaltningen organiserer arbeidet med digitalisering. (...) En av de viktigste hindringene er, etter Digdirs erfaring, utfordringene med å organisere tverrgående samarbeid i et sektorstyrt samfunn. Dette gjelder for alle nivåer av forvaltningen, og på tvers av offentlig og privat sektor.²⁷⁶

Ser vi på andre tiltak enn KI-strategien, opprettet Digitaliserings- og forvaltningsdepartementet en interdepartemental gruppe for KI i 2018. Arbeidsgruppen har ifølge departementet i det siste diskutert arbeidet med digitaliseringsstrategien, KI-forordning, rapporten om regulering av kunstig intelligens fra den regjeringsoppnevnte arbeidsgruppen ledet av Justis- og beredskapsdepartementet. Den interdepartementale arbeidsgruppen brukes ellers til å oppdatere hverandre på hva som skjer i de ulike sektorene. Gruppen fungerer som et kollegium når den er samlet, og fungerer som et nettverk der man har kontaktpersoner i de ulike departementene som kan brukes bilateralt. Departementet har godt utbytte av dette nettverket. Digitaliserings- og forvaltningsdepartementet planlegger å bruke denne gruppen i forbindelse med politikktutviklingen som pågår nå, og som skal inn i digitaliseringsstrategien. Ifølge departementet blir utfordringer med utvikling eller bruk av KI hos underliggende virksomheter/etater som regel ikke tatt opp til diskusjon i denne gruppen.

Digitaliserings- og forvaltningsdepartementet viser til at alle departementer er representert i den interdepartementale arbeidsgruppen, og at gruppen fra 2018 har hatt ulike funksjoner over tid og videreføres nå med revidert mandat fra våren 2024. Departementet har gode erfaringer fra arbeidet i denne gruppen.

Digitaliserings- og forvaltningsdepartementet sier i intervju at den nasjonale KI-strategien ikke la opp til noen satsing på kunstig intelligens, og at KI heller ikke er et mål i seg selv. Departementet viser til sektorprinsippet og legger til grunn i sin rolle- og ansvarsutøvelse at fagdepartementene er ansvarlige for måloppnåelsen og resultatene for digitalisering og kunstig intelligens i sine sektorer. Det er opp til hvert enkelt sektordepartement å vurdere KI som virkemiddel på sitt område. Videre påpeker departementet at bruken av KI i offentlig sektor skal være behovsstyrt.²⁷⁷

7.3.1 Digitaliseringsdirektoratets arbeid med kunstig intelligens

Blant de aktuelle organisatoriske virkemidler som Digitaliserings- og forvaltningsdepartementet har til rådighet, er Digitaliseringsdirektoratet det viktigste. Digitaliseringsdirektoratet skal være regjeringens fremste verktøy for effektiv og samordnet digitalisering av offentlig sektor og av samfunnet for øvrig. Digitaliseringsdirektoratets dialog med Digitaliserings- og forvaltningsdepartementet skjer på ulike nivåer, både gjennom den formelle styringsdialogen og i en løpende faglig dialog. På undersøkelsestidspunktet har Digitaliseringsdirektoratet ca. to årsverk som jobber med KI.

Digitaliseringsdirektoratet uttaler i intervju at de har hatt KI på agendaen siden 2019. Digitaliseringsdirektoratet har videre en overordnet rolle i å implementere den nasjonale strategien for KI og skal være premissgiver, igangsetter, pådriver og tilrettelegger for KI i det offentlige.

Digitaliseringsdirektoratets prioriteringer er styrt av tildelingsbrevet fra Digitaliserings- og forvaltningsdepartementet. «Kunstig intelligens» er bare omtalt én gang i tildelingsbrevene i perioden 2019–2023. Den ene omtalen er et konkret oppdrag fra departementet om et tiltak fra KI-strategien om å «utvikle veiledning for ansvarlig bruk av kunstig intelligens i offentlig sektor».²⁷⁸ Direktoratet har ikke noen annet direkte oppfølgingsansvar for KI-strategien enn dette ene oppdraget. Denne veiledningen er tilgjengelig på Digitaliserings- og forvaltningsdepartementets nettsider.

²⁷⁶ Digitaliseringsdirektoratet. (2023). *Innspill til ny nasjonal digitaliseringsstrategi*.

²⁷⁷ Verifisert referat fra intervju med Kommunal- og distriktsdepartementet 14. september 2023.

²⁷⁸ Statsbudsjettet 2022 – supplerende tildelingsbrev nr. 1, mars 2022

Veiledningsarbeidet er ifølge direktoratet sentralt for å sikre ansvarlig bruk av KI. Arbeidet må imidlertid deles inn i ulike deler for at veiledningsmaterialet skal kunne treffe godt, og for å sikre at det ikke utvikles ulike og konkurrerende tilnærminger. Direktoratet viser til at det første tiltaket for å kunne utarbeide veiledning om ansvarlig KI-bruk er å få oversikt over KI-systemene i staten. Oversikten bidrar til et kunnskapsgrunnlag for å målrette tiltak, ifølge direktoratet.

Digitaliseringsdirektoratet opplyser i intervju at det ikke har noen kvantitativ analyse av bruken av veiledningsmaterialet om KI. Direktoratet baserer seg på at deler av veiledningsmaterialet omtales mye, og at mange tar kontakt for å få nærmere opplysninger om deler av materialet. Veiledning om generativ KI er særlig etterspurt.

Digitaliseringsdirektoratet opplyser i intervju at det ser behov for mer samordnet tolkning rundt KI, der man fanger opp tverrfaglige spørsmål, også om regelverk. Det er behov for en mer felles tilnærming til risiko, ifølge direktoratet. Direktoratet viser til at det kan være ulike risikoer som må håndteres ved bruk av KI i ulike etater. Offentlige etater er på ulike utviklingsnivå i utviklingen og bruken av KI, men det er likevel overordnede problemstillinger som må favnes i en overordnet veiledning.

Som en del av veiledningen opprettet Digitaliseringsdirektoratet også et diskusjonsforum i den såkalte Datalandsbyen²⁷⁹ for å diskutere KI i offentlig sektor. Her har det vært liten aktivitet. Av seks diskusjonstråder er det fire som kun har ett innlegg (innlegget som ble postet ved opprettelsen av tråden), og siden 6. juni 2023 er det kun 18 innlegg totalt som handler om KI. Direktoratet opplyser i intervju at så langt har direktoratet fått mindre tilbakemeldinger enn forventet gjennom Datalandsbyen. Digitaliseringsdirektoratet ser et stort behov for mer veiledning og sier de vil prøve å fortsette dette arbeidet innenfor den eksisterende ressursrammen, men mener at det er svært kompetanse- og ressurskrevende å vurdere konkrete spørsmål om utvikling og bruk av KI.²⁸⁰

Om Digitaliseringsdirektoratets rolle i oppfølging av KI-strategien generelt sier Digitaliseringsdirektoratet at de ikke har mottatt et oppdrag fra Digitaliserings- og forvaltningsdepartementet om å følge opp noen av de 51 tiltakene i den nasjonale KI-strategien, bortsett fra å utvikle den nevnte veiledningen.

Andre oppgaver som er viktig for digitalisering generelt, er også viktig for kunstig intelligens, for eksempel deling av data og tilgang på data av god kvalitet – som omtalt ovenfor.

Direktoratet viser til i intervju at det er et stort behov for tilsyn på KI området. Tilsynsarbeidet må etter beste evne tilpasses til fremtiden, selv om utviklingen går raskt og det er usikkerhet. Likevel er det viktig å se hvilke problemer som kan komme, og hvilke behov for tilsyn som vil komme.

En gjennomgang av årsrapportene for Digitaliseringsdirektoratet i perioden 2019–2022 viser at kunstig intelligens blir omtalt mye mer i 2022 sammenlignet med 2019 og 2020. For eksempel omtalte årsrapporten for 2020 temaet slik:

Regelverket er en utfordring. Vi trenger et digitaliseringsvennlig regelverk, klart og forståelig, med enhetlig bruk av begreper. Det må også legges til rette for mer automatisert saksbehandling og fornuftig bruk av kunstig intelligens.

Omtalen av KI for 2020 handlet altså ikke om noe mer konkret Digitaliseringsdirektoratet hadde foretatt seg. Fra 2022 blir kunstig intelligens omtalt mye mer, og flere områder som direktoratet rapporterer på, kan indirekte få følger for KI uten at KI er eksplisitt definert som mål. Mange av

²⁷⁹ Datalandsbyen er ifølge Digitaliseringsdirektoratet et nettforum der man skal kunne diskutere alt rundt deling og bruk av data, jf. <https://datalandsbyen.norge.no/>

²⁸⁰ Verifisert referat fra intervju med Digitaliseringsdirektoratet 20. september 2023.

oppgavene til Digitaliseringsdirektoratet er satt i sammenheng med utviklingen av kunstig intelligens, og direktoratet har definert ansvarlig bruk av kunstig intelligens som ett av fire hovedmål for sitt arbeid.

På spørsmål om andre aktiviteter lister Digitaliseringsdirektoratet opp følgende:

- Direktoratet satt i 2019 i gang en offentlig ph.d.²⁸¹ som skulle belyse bruken av KI i offentlig sektor. Dette arbeidet planlegges ferdigstilt våren 2024.
- Digitaliseringsdirektoratet organiserte et innspillmøte i forbindelse med utarbeidelsen av den nasjonale KI-strategien.
- Direktoratet har tidligere deltatt i CAHAI med én medarbeider i arbeidsgruppen. Arbeidet i denne gruppen er overført til Justis- og beredskapsdepartementet.²⁸²
- Direktoratet sendte et hørings svar som bidrag til det norske posisjonsnotatet til EUs KI-forordning (jf. forslag til forordning fra Kommisjonen om et harmonisert regelverk for kunstig intelligens og om endring av visse EU-rettsakter).
- Direktoratet har begynt å forberede hvordan KI-forordning kan implementeres i Norge.
- Direktoratet har laget en oversikt over bruken av KI i offentlig sektor i samarbeid med NORA²⁸³.
- Digitaliseringsdirektoratet administrerer «Fagforum for kunstig intelligens i staten».

Fagforum for kunstig intelligens var et tiltak som kom nedenfra, og ble etablert av KI-interesserte medarbeidere fra blant annet Nav og skatteetaten. Det var ikke et tiltak fra Digitaliserings- og forvaltningsdepartementet eller Digitaliseringsdirektoratets side, selv om Digitaliseringsdirektoratet nå har tatt over administreringen av fagforumet. Fagforumet er også nevnt i departementets oppfølging av den nasjonale KI-strategien, i forbindelse med at regjeringen vil «legge til rette for samarbeid og utveksling av beste praksis for KI i både statlig og kommunal forvaltning». Digitaliserings- og forvaltningsdepartementet viser til at slik tilrettelegging skjer delvis gjennom Fagforum for kunstig intelligens, men at det «må gis som oppdrag til Digitaliseringsdirektoratet dersom det skal gjøres mer systematisk». Digitaliserings- og forvaltningsdepartementet har ikke gitt et slikt oppdrag til Digitaliseringsdirektoratet. Digitaliserings- og forvaltningsdepartementet opplyste imidlertid i intervju at det vurderer å slutte med «oppdragsstyring» av Digitaliseringsdirektoratet fordi departementet mener det er en lite effektiv måte å styre på.

Videre hadde Digitaliseringsdirektoratet en sak for Skate²⁸⁴ om «KI-tiltak» i juni 2023, der blant annet Skate konkluderer med at «veiledning om de regulatoriske aspektene er viktig, det er behov for mer kraft i det juridiske arbeidet.» Skate besluttet videre som følger²⁸⁵:

- Skate ber Digitaliseringsdirektoratet arbeide videre med skisserte tiltak og sikre involvering av relevante virksomheter i og utenfor Skate.
- Skatevirksomhetene stiller med relevante ressurser i samarbeidet.
- Skate vil initiere en diskusjon om sammenhengen mellom KI og deling av data knyttet til ny digitaliseringsstrategi. Offentlig sektor må ha en offensiv tilnærming til bruk av ny teknologi.

Til slutt kan det nevnes at Digitaliseringsdirektoratet har tatt initiativ til en nordisk-baltisk arbeidsgruppe, «Nobareg», om håndtering av blant annet KI-regelverk og hvordan tilsyn på KI-området kan innrettes. Prosjektet er finansiert av Nordisk ministerråd og ledes av Digitaliseringsdirektoratet.

²⁸¹ H. Broomfield, ph.d.-tema «Government by Algorithm – Protecting Democratic Values in the Deployment of Artificial Intelligence by the Public Sector», finansiert 50 prosent av Digitaliseringsdirektoratet og 50 prosent av Forskningsrådet.

²⁸² CAHAI - Ad hoc Committee on Artificial Intelligence - Artificial Intelligence (coe.int)

²⁸³ Nora = Norwegian Artificial Intelligence Research Consortium.

²⁸⁴ Skate er et organ for et utvalg toppledere i staten, som fungerer som et strategisk samarbeidsråd og rådgivende organ til Digitaliseringsdirektoratet og digitaliseringsministeren.

²⁸⁵ Skate-møte 14. juni 2023 – referat <https://www.digdir.no/skate/skate-mote-14-juni-2023-referat/4690>

8 Vurderinger

8.1 Statlige virksomheter og heleide selskap utnytter mulighetene ved kunstig intelligens ulikt, og kunstig intelligens er fortsatt lite i bruk

Kunstig intelligens er et viktig verktøy for utvikling av en bærekraftig, effektiv og brukerrettet offentlig forvaltning. Potensialet og behovet for å ta i bruk kunstig intelligens i offentlig sektor er tatt opp i en rekke meldinger og innstillinger, herunder Meld. St. 27 (2015–2016) *Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet* og Meld. St. 30 (2019–2020) *En innovativ offentlig sektor — Kultur, ledelse og kompetanse og strategier*.²⁸⁶ Stortinget viser videre til betydningen av at utviklingen og bruk av kunstig intelligens må skje innenfor ansvarlige rammer der blant annet personvernet ivaretas.²⁸⁷

Undersøkelsen viser at det er stor variasjon i hvilken grad statlige virksomheter, heleide selskaper og helseforetak har tatt i bruk kunstig intelligens. Blant de over 200 undersøkte virksomhetene er det under halvparten som har erfaring med utvikling og/eller bruk av kunstig intelligens. I tillegg har i underkant av 20 prosent planer om å ta i bruk kunstig intelligens i fremtiden. I denne undersøkelsen har vi lagt til grunn den brede definisjonen av KI, som i den nasjonale KI-strategien, og derfor er det mange prosjekter og systemer som kan rubiseres som KI uten at det er noe annet enn en enkel programvarerobot, som ikke faller under KI-definisjonen i f.eks. KI-forordningen²⁸⁸. Over halvparten av KI-systemene i undersøkelsen bruker kun slike regelbaserte eller enkle statistiske modeller for å tolke data. Med en strengere definisjon av hva som kan defineres som et KI-system, med krav til mer avanserte modeller, ville antallet oppgitte KI-systemer ha vært langt lavere.

Blant virksomhetene som har erfaring med kunstig intelligens, viser undersøkelsen at et flertall (56 prosent) har erfaring med kun mellom ett og tre KI-prosjekter. I overkant av 10 prosent av virksomhetene har erfaring med ti eller flere KI-prosjekter.

Blant virksomhetene som utvikler og bruker kunstig intelligens, har helseforetakene klart flest prosjekter. Det er også helseforetakene som i størst grad bruker avanserte modeller. Deretter følger kunnskapssektoren og noen av virksomhetene under Finansdepartementet. Under Nærings- og fiskeridepartementet er det også flere KI-prosjekter blant både underliggende virksomheter og heleide selskaper. Nærmere 70 prosent av de innrapporterte KI-prosjektene ligger under disse fire departementene.

Noen av ulikhetene mellom sektorene kan etter vår vurdering tilskrives at sektorene har ulike behov og grunnleggende forutsetninger for å utvikle og bruke KI. I helsesektoren er det uttalt at kunstig intelligens kan bidra til mer effektiv ressursbruk og til et bærekraftig helsevesen, jf. for eksempel Meld. St. 7 (2019–2020) *Nasjonal helse- og sykehusplan 2020–2023* og oppdragsbrevet for 2024 til de regionale helseforetakene. Helsektoren har i utgangspunktet i tillegg noen grunnleggende forutsetninger for å kunne utvikle og å ta i bruk kunstig intelligens. Det finnes eksempelvis store mengder data blant annet gjennom helseregistrene og etablerte forskningsmiljøer. Undersøkelsen viser videre at Helse- og omsorgsdepartementet har gitt føringer overfor helseforetakene om å ta et aktivt grep om å utvikle og bruke kunstig intelligens i egen sektor.

²⁸⁶ Kommunal- og moderniseringsdepartementet. (2019). *En digital offentlig sektor. Digitaliseringsstrategi for offentlig sektor 2019–2025*.

²⁸⁷ Innst. 191 S (2020–2021).

²⁸⁸ European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act), [P9_TA\(2024\)0138](#), fortalet punkt (12)

Etter vår vurdering er det i de andre undersøkte sektorene mer opp til den enkelte virksomhet å eventuelt ta i bruk kunstig intelligens. De intervjuede departementene viser til prinsippene om mål- og resultatstyring, og følgelig at virksomhetene selv må velge og bruke de virkemidlene og den teknologien som gir de beste løsningene for å sikre en effektiv oppgaveløsning.

Undersøkelsen viser at de undersøkte departementene, som Helse- og omsorgsdepartementet og Finansdepartementet på den ene siden og Arbeids- og inkluderingsdepartementet på den andre, har noe ulik tilnærming til utarbeidelse av hjemler som tydeliggjør hvorvidt underliggende virksomheter har et juridisk behandlingsgrunnlag for behandling av personopplysninger i KI-systemer.

Uavhengig av hvor avanserte KI-systemene er, viser undersøkelsen at virksomhetenes mål med å utvikle og ta i bruk KI-systemer i stor grad er å effektivisere og forbedre egen oppgaveløsning, inkludert å frigjøre tid. Undersøkelsen viser videre at de fleste virksomhetene mener de i stor grad lykkes med å nå målene de har satt seg for å utvikle og å ta i bruk kunstig intelligens.

Samlet sett ser omfanget av utvikling og bruk av kunstig intelligens i staten ut til å være lite så langt. Noen få sektorer og virksomheter har imidlertid begynt med utforskning og satt noen systemer i produksjon. Blant helseforetakene og for eksempel i skatteetaten og Statens pensjonskasse finnes det eksempler på bruk av kunstig intelligens som andre virksomheter kan lære av. Samtidig har Nav som forvalter omtrent en tredel av statsbudsjettet, ennå ingen produksjonssatte løsninger basert på egenutviklet kunstig intelligens. Dette på tross av at Nav begynte tidlig med utviklingen og har hatt en egen enhet for kunstig intelligens siden 2017. Etersom det er få eksempler på at det har blitt utviklet og tatt i bruk KI-systemer, er det få eksempler hvor KI i stor grad har bidratt til effektivisering og en mer produktiv offentlig sektor. Etter vår vurdering ser det ut til å være et stort potensial i offentlig sektor til å effektivisere gjennom digitalisering og bruk av kunstig intelligens som så langt ikke er benyttet.

8.2 Viktige forutsetninger for å ta KI i bruk i stor skala i statlige virksomheter er ikke på plass

Undersøkelsen viser at det er flere barrierer for utvikling og bruk av kunstig intelligens i staten. Digitaliserings- og forvaltningsdepartementet anser teknologien som relativt ny og mener at det vil ta tid å utvikle og ta i bruk kunstig intelligens. Første versjon av skatteetatens KI-baserte fradragmodell ble imidlertid satt i produksjon allerede for ti år siden (i 2014). Undersøkelsen viser at det er en rekke faktorer som ser ut til å gjøre arbeidet med å ta i bruk kunstig intelligens i staten vanskelig for mange virksomheter.

8.2.1 Regulatoriske uklarheter blir ikke avklart

I Norge har vi generelt et teknologinøytralt regelverk. Offentlig sektor står dermed fritt til å velge hvilken teknologi som passer best til sin oppgaveløsning, og regelverket skal ikke hindre bruk av teknologi. Et teknologinøytralt regelverk kan dermed også bidra til innovasjon. I Norge finnes det heller ingen spesifikk regulering av kunstig intelligens.

Undersøkelsen viser imidlertid at både mangel på tydelige lovhjemler og usikkerhet rundt tolkning av lover og regelverk er faktorer som i større grad ser ut til å hindre virksomheter i å utvikle og å ta i bruk kunstig intelligens. Det er særlig stor usikkerhet rundt behandlingen av personopplysninger i utvikling og bruk av KI-systemer. Noen departementer, som Finansdepartementet og Helse- og omsorgsdepartementet, har arbeidet gjennom sektorlovverket med å gi underliggende virksomheter et tydeligere juridisk grunnlag til å kunne behandle personopplysninger. Undersøkelsen viser også at det er innenfor disse sektorene det er flest KI-prosjekter. Vi merker oss imidlertid at Nav og Statens pensjonskasse, som begge er underlagt Arbeids- og inkluderingsdepartementet, har ulike

hjemmelsgrunnlag for å behandle personopplysninger i sine respektive lover. Statens pensjonskasse har et tydelig rettslig grunnlag for å behandle personopplysninger, mens for Navs del er dette uavklart.

Undersøkelsen viser at virksomhetene ofte synes det er vanskelig å håndtere spørsmål om personvern i det enkelte prosjektet, ikke minst fordi personvern må være ivaretatt i alle deler av et KI-prosjekt, også i utviklingsfasen. Selv om virksomhetene har et generelt hjemmelsgrunnlag for å behandle personopplysninger, må dette vurderes spesifikt i det enkelte KI-prosjektet. Her ser det ut til å være mye usikkerhet. Noen velger å løse dette ved å prøve seg frem med mindre prosjekter for å bygge erfaring. Etter vår vurdering er det en risiko for at usikkerheten forsinker og hindrer virksomheter i å ta i bruk de mulighetene som kunstig intelligens gir.

I Dokument 3:8 (2023–2024) *Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen* ble det vist til at manglende og sene avklaringer av juridiske problemstillinger hindrer deling og gjenbruk av data.

Sektordepartementene og virksomhetene har selv ansvaret for å vurdere eget hjemmelsgrunnlag og behovet for eventuelle endringer i eget regelverk for å tydeliggjøre grunnlaget for å behandle personopplysninger. Men det er også en rekke ganske like juridiske problemstillinger som mange virksomheter møter på. Virksomhetene har i stor grad like spørsmål om hjemmelsgrunnlag og tolkningen av regelverket, og da særlig spørsmål om behandlingen av personopplysninger. For å bistå virksomhetene i slike spørsmål har Digitaliserings- og forvaltningsdepartementet arbeidet med veiledning, særlig gjennom Datatilsynet, men også gjennom Digitaliseringsdirektoratet. Digitaliseringsdirektoratet opplyser i intervju at det ser behov for en mer samordnet tolkning av regelverk for utvikling og bruk av KI. Gitt de utfordringene et stort antall virksomheter møter i arbeidet med personopplysninger, ser behovet for veiledning ut til å være stort. Vår vurdering er at det er særlig viktig å sikre en mest mulig enhetlig forståelse av hvordan man skal legge regelverket til grunn i utvikling og bruk av kunstig intelligens. Det samme har Digitaliseringsdirektoratet og noen store statlige etater påpekt.

Opprettelsen av den regulatoriske sandkassen hos Datatilsynet var et sentralt tiltak for å bistå blant annet offentlige virksomheter i å øke forståelsen av de regulatoriske kravene som stilles på personvernområdet. Bedre kunnskap vil kunne korte ned tiden fra utvikling og testing til faktisk utrulling av KI-løsninger. Løsningene som settes i drift etter å ha deltatt i sandkassen, skulle fungere som foregangseksempler. På overordnet nivå ser sandkassen ut til å være et godt tiltak. Det gir læring for de som deltar, og for Datatilsynet. Undersøkelsen viser imidlertid at det er kun fem prosjekter i statlig regi som har deltatt i sandkassen så langt. Det har vært relativt få søkere til ordningen, og Datatilsynet har heller ikke kapasitet til å håndtere flere prosjekter. Erfaringene med å delta spriker. Blant de to undersøkte prosjektene var ett prosjekt negativt fordi det ble ikke gitt noe svar på det prosjektet ønsket å avklare. Det andre prosjektet syntes deltakelsen var positiv. Samtidig oppstod det spørsmål som ikke kunne løses innenfor prosjektets rammer. Selv om deltakelse kan gi læring for enkeltprosjektene, virker det som det er en utfordring for sandkassen å formidle denne læringen videre og sikre at den blir brukt av andre KI-prosjekter til videre utvikling.

I behandlingen av Meld. St. 30 (2019–2020) viste kommunal- og forvaltningskomiteen til at den var svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, inkludert et digitaliseringsvennlig regelverk. I den nasjonale strategien for kunstig intelligens vises det også til at det skal gjøres en gjennomgang og vurdere regelverk som er til hinder for hensiktsmessig og ønsket bruk av kunstig intelligens i offentlig og privat sektor. Undersøkelsen viser at en slik gjennomgang ikke har blitt gjort. Digitaliserings- og forvaltningsdepartementet legger til grunn at det er sektordepartementene og virksomhetene som må vurdere behovet for dette. Selv om hver sektor har ansvaret for eget regelverk, fremstår det som utydelig hvem som skal sikre at hindringer i regelverket blir gjennomgått. En gjennomgang av regelverket på tvers kunne etter vår

vurdering på en bedre måte ha fått frem behovet for felles veiledning på tvers av virksomhetene og sektorene.

8.2.2 Mangelfull infrastruktur og tilgang på data av god kvalitet

En viktig forutsetning for å kunne utvikle KI-systemer er som regel tilgang på store datamengder av god kvalitet – ofte fra både egen virksomhet og andre. Dataene må kunne brukes til å trene opp maskiner før et system kan settes inn i produksjon. Tilgangen på data gjelder også behovet for å validere innkjøpte modeller og for å kunne kalibrere og kontinuerlig overvåke eksisterende modeller. Potensialet for å effektivisere gjennom kunstig intelligens ligger i mulighetene til å utnytte både eksisterende og nye data.

Data av god kvalitet er over flere år løftet frem som særlig viktig, blant annet gjennom prinsippet om orden i eget hus. Gjennom Digitaliseringsrundskrivet er statlige virksomheter pålagt å sikre god datakvalitet, og dette er løftet frem i den nasjonale strategien for kunstig intelligens.

Undersøkelsen viser at mangelfull tilgang på data av god kvalitet i noen tilfeller er til hinder for å utvikle og bruke kunstig intelligens. Rundt halvparten av de som svarte på spørreundersøkelsen, opplyser at dårlig tilgang på data og lav datakvalitet har vært en utfordring i arbeidet med å utvikle og å ta i bruk kunstig intelligens. Utfordringer med tilgangen på data av god kvalitet kan ha flere årsaker. Noen virksomheter mangler gode, interne data. Andre virksomheter har mye data, men de er ikke nødvendigvis digitale eller tilgjengelige i et felles system. Hvorvidt tilgang på data er et problem, kommer også an på hvilket ambisjonsnivå en virksomhet har. Kompliserte modeller som skal bruke flere kilder, som bilder og tekst, i tillegg til strukturerte registerdata, vil kreve et langt større datagrunnlag enn for eksempel en enkel regnskapsrobot. Flere virksomheter har ikke god nok infrastruktur til verken å lagre eller håndtere store datamengder til analyse. Særlig for små virksomheter kan det være kostbart å investere i ulike digitale løsninger som kreves for å utvikle og bruke kunstig intelligens. Vi merker oss at flere etterlyser felles digitale løsninger og muligheter for å få tilgang på skyløsninger.

Manglende datadeling på tvers av virksomheter kan være en annen årsak til at tilgangen på data er mangelfull. Skal datadelingen fungere, må den enkelte offentlige virksomheten legge egne data til rette på en god måte slik at det blir enkelt å dele med andre, og slik at andre kan bruke data på en hensiktsmessig måte. Digitaliserings- og forvaltningsdepartementet viser til at den enkelte virksomheten selv må ta ansvar for at dataene er av god kvalitet. Felles datakatalog er etablert med åpne data og metadata om lukkede datasett for å legge til rette for å dele data, mens tilgang på lukkede data som for eksempel personopplysninger krever andre løsninger. Etableringen av Felles datakatalog og andre tiltak i regi av Digitaliseringsdirektoratet har bidratt positivt til mer deling av åpne data. Det er imidlertid anslagsvis kun en fjerdedel av de virksomhetene som er omfattet av Digitaliseringsrundskrivet, som oppfyller kravene om deling og videre bruk av informasjon på Felles datakatalog. Her vil vi også vise til utfordringer med deling av data i offentlig sektor og manglende registrering i Felles datakatalog, som ble påvist i Riksrevisjonens Dokument 3:8 (2023–2024).

Samlet sett er det store utfordringer i arbeidet med kunstig intelligens som følge av mangelfull tilgang på data av god kvalitet, ofte sammen med svake systemer for å håndtere data. Sektorene har et ansvar for å sikre god datakvalitet og at åpne data er av en kvalitet som enkelt kan deles med andre sektorer. Slike krav er gitt blant annet gjennom Digitaliseringsrundskrivet om orden i eget hus og krav om deling av data i offentlig sektor. Gitt utfordringene med mangelfull tilgang på data av god kvalitet er det etter vår vurdering ikke tilstrekkelig oppfølging fra sektormyndighetenes side av at kravene til deling av data og særlig til orden i eget hus blir etterlevd i virksomhetene. Dette vil kunne ha konsekvenser for offentlig sektors muligheter for å utvikle og utnytte potensialet i kunstig intelligens.

8.2.3 Stort behov for kompetanse

Undersøkelsen viser at en av de største utfordringene med å utvikle og bruke kunstig intelligens er mangel på kompetanse. Det gjelder ofte IT-relatert kompetanse, men det kan også være behov for juridisk kompetanse og en bedre samhandling mellom ulike kompetansemiljøer internt i en virksomhet. I noen store virksomheter som skatteetaten, Nav og i helseforetakene, og også i mellomstore virksomheter som Statens pensjonskasse har det vært satset på å etablere større miljøer med kompetanse til å arbeide med kunstig intelligens. Flere virksomheter har imidlertid ikke kapasitet eller ressurser til å ansette flere med nødvendig digital kompetanse. Mange virksomheter er derfor avhengig av å leie inn konsulenter. I anskaffelse av både konsulenttjenester og leverandørsystemer er det behov for bestillerkompetanse for å kunne stille riktige krav til henholdsvis konsulentene og leverandørsystem. Det er også viktig å sikre tilstrekkelig kompetanseoverføring og dokumentasjon fra eksterne konsulenter og leverandører.

I forbindelse med utfordringene med kompetanse merker vi oss at Meld. St. 14 (2022–2023) *Utsyn over kompetansebehovet i Norge* viser at myndighetene og universitetene og høyskolene har et delt ansvar for å dimensjonere utdanningstilbudet – til dels ut fra særlige arbeidskraftsbehov. Vi merker oss også at behovet for IT-kompetanse er antatt å øke i årene frem til 2030.

8.3 De etiske prinsippene for ansvarlig bruk av KI ivaretas i ulik grad. Kontrollmekanismer som sikrer ansvarlig bruk av kunstig intelligens må være på plass

Ny teknologi, som kunstig intelligens, har et stort potensial for å øke produktiviteten og kvaliteten i offentlig sektor nå og fremover. Men for at tilliten til offentlige systemer og saksbehandling skal bevares, er det likevel nødvendig å regulere bruken av kunstig intelligens i offentlig sektor slik at teknologien blir brukt på best mulig måte innenfor trygge og ansvarlige rammer. Utgangspunktet for dette er en ansvarlig og pålitelig utvikling og bruk av kunstig intelligens forankret i etiske prinsipper og krav til personvern, likebehandling, sikkerhet og robusthet og forklarbarhet og åpenhet.²⁸⁹ Det er viktig å ha kontrollmekanismer på plass for å ivareta disse prinsippene og rettsikkerheten.

8.3.1 Virksomhetene har kun delvis etablert systemer for sikring av ansvarlig bruk av KI

Undersøkelsen viser at virksomhetenes systemer for å sikre ansvarlig bruk av KI ivaretar de etiske prinsippene i ulik grad. Dette gjelder både for styringsdokumentene som retningslinjer eller policyer knyttet til KI, og for virksomhetenes rapportering av hvordan de arbeider med ansvarlig utvikling av KI. Styringsdokumentene for risikovurdering omtaler også i ulik grad de særskilte risikoene som kunstig intelligens fører med seg.

Styringsdokumentene dekker i større grad kravene til sikkerhet, robusthet og personvern, og i mindre grad kravene til likebehandling, forklarbarhet og åpenhet. Over 60 prosent av organisasjonsenhetene svarer i spørreundersøkelsen at de har benyttet skriftlige retningslinjer eller policyer for databehandling og innebygd personvern ved utvikling av KI. Under halvparten oppgir at de har benyttet retningslinjer for planlagt innføring, overvåkning og vedlikehold og testing for skjevheter av KI-systemer. Videre svarer flertallet av organisasjonsenhetene som bruker standarder for dokumentasjon av KI-systemer, at disse helt eller delvis dekker databehandling og kilder, modellvalg og oppdatering og vedlikehold. Under halvparten svarer at standardene stiller krav til dokumentasjon av forklaring av KI-systemets utfall og valg av funksjons- og ytelsesnivå.

²⁸⁹ Jf. kapittel 3

Gjennomgangen av generelle styringsdokumenter for utviklingen av de fire KI-systemene vi har gått i dybden på viser det samme. Disse omtaler og stiller krav i større grad til sikkerhet, robusthet og personvern og i mindre grad til likebehandling og åpenhet.

I spørreundersøkelsen spurte vi virksomhetene hvordan de arbeider med ansvarlig kunstig intelligens for sine KI-systemer. Undersøkelsen viser at mellom 35 og 60 prosent svarte ja til påstander om sikkerhet og robusthet, for eksempel at KI-systemene var testet for svakheter ved cyberangrep og testing av stabilitet i KI-systemene over tid. En mindre andel, mellom 25 og 30 prosent, svarte ja til påstander om åpenhet, for eksempel at informasjon om bruk av KI og hvordan KI-systemet fungerer og treffer beslutninger ligger åpent for eksterne brukere. Videre har kun 13 prosent av KI-systemene som direkte berører enkeltpersoner substansielt, blitt testet for skjevheter i KI-systemenes utfall.

Det er en rekke risikoer og problemstillinger ved ansvarlig bruk av KI som bør tas hensyn til når man utvikler og bruker KI. Tradisjonelle rammeverk og styringsdokumenter for risikovurderinger av IT-systemer kan derfor være mindre egnet til å avdekke KI-risiko, med mindre de er særskilt tilpasset slike risikoer. Gjennomgangen av styringsdokumenter for risikovurderinger fra 21 virksomheter viser at de stort sett er generelle. Enkelte virksomheter har sendt med maler for risikoanalyse eller personvern vurderinger, som dekker standard IT-sikkerhet og informasjonssikkerhet samt risikoene for personvern forbundet med blant annet dataminimering og formålsbegrensning. Rettferdighet, forklarbarhet og åpenhet er i liten grad dekket i styringsdokumentene for risikovurderinger. Når det ikke finnes retningslinjer eller policyer for slike spesifikke KI-risikoer og utfordringer, blir håndteringen av dette mer avhengig av kompetansen til fagpersonene som utvikler eller kjøper inn KI-systemene.

Statlige virksomheters fremskritt i hvordan de jobber med å ivareta ansvarlig bruk av kunstig intelligens er ulik. Enkelte har formulert overordnede prinsipper for ansvarlig kunstig intelligens i virksomheten, mens andre har innført mer konkrete retningslinjer knyttet til noen av de etiske prinsippene.

Ut fra det vi har funnet i spørreundersøkelsen og gjennomgangen av styringsdokumentene, mener vi at virksomhetene kun delvis har etablert systemer som sikrer ansvarlig bruk av kunstig intelligens. Det er en svakhet som kan føre til at deler av den særskilte risikoen ved utvikling og bruk av KI ikke fanges opp eller vurderes. Manglene ved virksomhetenes styringssystem kan øke risikoen for at det ikke arbeides systematisk med prinsipper for likebehandling, forklarbarhet og åpenhet. Dette illustrerer behovet for å heve kompetansen og veiledningen av virksomhetene for noen av de etiske prinsippene for ansvarlig bruk av kunstig intelligens.

8.3.2 Ved utvikling og bruk av KI ivaretas hensynet til sikkerhet, robusthet og personvern i større grad enn transparens og likebehandling

De fire utvalgte KI-systemene i undersøkelsen er mindre utviklingsprosjekter med lav risiko som brukes som beslutningsstøtte i saksbehandlingen. Undersøkelsen av disse KI-systemene viser at de ulike hensynene til ansvarlig bruk av kunstig intelligens ivaretas i ulik grad.

Undersøkelsen viser at KI-prosjektene er tydelig forankret i virksomheten og i virksomhetens målstruktur, og det vises til positive resultater. Samtidig er det svakheter og mangler i dokumentasjonen av KI-systemene. Dette gjør det for eksempel vanskelig å få kjennskap til hvilke veivalg som er tatt i utvikling av KI-modellen, og hvorfor en KI-modell er valgt fremfor andre.

Videre er manglene ved virksomhetenes systemer for sikring av ansvarlig bruk av KI også til stede i utviklingen av de utvalgte KI-systemene. De fire utvalgte KI-systemene har gjennomført flest aktiviteter rettet mot sikkerhet, robusthet og personvern og færre aktiviteter rettet mot likebehandling, forklarbarhet og åpenhet. Variasjon blant KI-prosjektene er størst i områdene transparens og likebehandling. Utviklingen og bruken av KI-systemene viser blant annet det følgende:

- Lovligheten ved behandling og bruk av personopplysninger og prinsippet om innebygd personvern ivaretas.
- Rettferdighet vurderes overordnet i utviklingen av KI-systemene. KI-prosjektene legger til grunn at likebehandling ivaretas ved at alle er gjenstand for samme KI-system og beregninger, og saksbehandlere alltid er involvert og treffer beslutningen. Risiko for diskriminering eller usaklig forskjellsbehandling er imidlertid også til stede når KI-systemer er brukt som beslutningsstøtte, ikke bare i autonome systemer. Det er i liten grad gjennomført tester for skjevheter i KI-systemets inndata og resultat.
- Det er åpenhet internt om KI-systemenes funksjon, ytelse og kvalitet, og det deles generell informasjon om KI-bruk i virksomheten med eksterne. KI-systemenes beslutninger og resultater kan i varierende grad forklares.

Virksomhetene som utvikler KI-systemene, mener at det er lav risiko og få konsekvenser forbundet med at det er gjort færre aktiviteter rettet mot likebehandling, forklarbarhet og åpenhet. De viser til det at KI-systemene brukes som beslutningsstøtte, at en saksbehandler alltid er involvert, og at virksomheten bruker sine egne data og at virksomhetene dermed ivareta disse hensynene godt nok.

Tre av KI-systemene er utviklet av virksomheten selv, mens ett er anskaffet fra leverandør. Undersøkelsen gir et innblikk i hvordan disse to metodene for å ta i bruk KI kan påvirke arbeidet med å sikre ansvarlig bruk av KI. Kravene til ansvarlig bruk av KI er de samme, men undersøkelsen viser at måten de etterleves på, er forskjellig. Undersøkelsen viser at selvutviklerne bruker kunnskap og erfaring fra KI-utviklingen og arbeidet for å ivareta ansvarlig bruk av KI i praksis i sitt prosjekt. Ved anskaffelser av KI fra leverandør stilles det krav til innkjøperen. For å ivareta ansvarlig bruk av KI må man ha kunnskap til å stille de riktige kravene i anskaffelsesprosessen og følger opp at leverandørene oppfyller disse kravene. For sertifiserte produkter kan en del av ansvaret være dekket av sertifiseringen. Sertifiseringens gyldighet kan imidlertid være begrenset til bruksområde og bruksmåte beskrevet i samsvarserklæringen av leverandør. Beskrivelsen dekker nødvendigvis ikke eventuelle tilpasninger av KI-systemet til lokale behov.

Undersøkelsen viser at virksomhetene har ulike modenhetsnivåer i hvordan de følger de etiske prinsippene som utgjør ansvarlig bruk av KI. Det gjelder både systemene for og arbeidet med å sikre ansvarlig utvikling av kunstig intelligens. I utviklingen og bruken av de utvalgte KI-systemene er det særlig hensynet til likebehandling, forklarbarhet og åpenhet som i mindre grad er jobbet med. Disse hensynene skal blant annet hindre diskriminering og usaklig forskjellsbehandling, sikre at beslutninger og resultater av KI-systemet kan forklares, og gjøre vurderingene i saksbehandlingen etterprøvbare. Risikoen ved og konsekvensene av disse svakhetene er lave for de utvalgte KI-systemer som er i bruk i virksomhetene i dag, men de illustrerer utfordringene med å ivareta disse prinsippene, som kan være en risiko i virksomhetenes videre KI-utvikling. Det synliggjør at virksomhetene trenger mer kunnskap og støtte for å kunne ivareta likebehandling, forklarbarhet og åpenhet bedre ved utvikling av KI. Videre viser undersøkelsen at arbeidet med personvern og KI er krevende, noe som kan hemme utviklingen i mange virksomheter.

8.4 Manglende samordning i arbeidet med kunstig intelligens i offentlig sektor

I behandlingen av Meld. St. 30 (2019–2020) viste kommunal- og forvaltningskomiteen til at kunstig intelligens er et eksempel på en teknologi i rask utvikling, som først kan hente ut sitt fulle potensial gjennom et effektivt 5G-nett. Komiteen var svært opptatt av at regjeringen jobber for at Norge får en infrastruktur for kunstig intelligens i verdensklasse, i form av digitaliseringsvennlig regelverk, gode språkressurser, raske og robuste kommunikasjonsnett og tilstrekkelig regnekraft. Disse målene er konkretisert gjennom den nasjonale strategien for kunstig intelligens fra 2020, som inneholder 51

tiltak. Det er ikke gjennomført noen systematisk evaluering av gjennomføringen av strategien, men departementet opplyser at de fleste tiltakene er gjennomført eller påbegynt.

Undersøkelsen viser at Digitaliserings- og forvaltningsdepartementet ikke har systematisk kartlagt bruken av kunstig intelligens i offentlig sektor. Det finnes en ufullstendig oversikt over KI-prosjekter i offentlig sektor som er gjennomført av NORA og Digitaliseringsdirektoratet. Det er imidlertid ikke lagt opp til noen systematisk oppdatering av denne kartleggingen. Det finnes derfor ingen målinger av utviklingen av bruk av kunstig intelligens i offentlig sektor, og den er heller ikke evaluert.

Digitaliserings- og forvaltningsdepartementet opplyser at det gjennom den nasjonale KI-strategien ikke er lagt opp til en satsing på kunstig intelligens. Departementet presiserer imidlertid at det er et politisk mål om å ta i bruk kunstig intelligens til effektivisering i offentlig sektor og for verdiskaping i samfunnet. Departementet kan legge til rette for bruk av kunstig intelligens gjennom generelle virkemidler – særlig gjennom pedagogiske virkemidler som veiledning. Veiledningsarbeidet skjer også gjennom Datatilsynet og Digitaliseringsdirektoratet, som omtalt ovenfor.

Digitaliserings- og forvaltningsdepartementet viser videre til at deres pådriverrolle for kunstig intelligens innebærer å identifisere tverrgående behov og få plass felles løsninger på tvers av sektorene. Det innebærer også at departementet arbeider for å sikre bedre koordinering rundt identifiserte behov. Vi merker oss at Digitaliserings- og forvaltningsdepartementet viderefører den interdepartementale arbeidsgruppen for kunstig intelligens, der alle departementene nå er representert. Formålet med gruppen er å sikre bedre koordinering og informasjonsutveksling på KI-området.

Vi merker oss videre at Digitalisering- og forvaltningsdepartementet nå løfter frem kompetanse og veiledning som en felles problemstilling. Departementet mener at sektorlovgivningen er ulik, men er usikker på hvor langt departementet skal gå inn i den konkrete sektoren, og hvor relevante problemstillingene i én sektor er for andre sektorer. Vi vil påpeke at alle disse spørsmålene har vært sentrale over flere år²⁹⁰, og at det vil være vanskelig å sikre en godt nok tilpasset veiledning når man mangler en systematisk oversikt over bruken av kunstig intelligens i staten.

De enkelte statlige sektorene og etatene må selv vurdere om og eventuelt i hvilket omfang kunstig intelligens bør brukes til for eksempel å forbedre og effektivisere egen oppgaveløsning. Vår vurdering er at sektorovergripende tiltak er sentralt for å kunne sikre en god felles infrastruktur for kunstig intelligens. For å få på plass noen grunnleggende forutsetninger for å lykkes med kunstig intelligens trenger offentlig sektor en mer overordnet og sektorovergripende tilnærming. Det gjelder blant annet å sikre god datakvalitet og deling av data, teknologisk infrastruktur og kompetanse. Videre gjelder det å sikre felles forståelse og tilnærming til juridiske spørsmål om personvern og ansvarlig bruk av kunstig intelligens. Undersøkelsen viser at dette er vedvarende utfordringer for å kunne ta i bruk kunstig intelligens i større grad i offentlig sektor.

Videre vil vi peke på uklarheter rundt hvordan de etiske prinsippene for ansvarlig kunstig intelligens kan ivaretas i praksis. For å kunne kontrollere om et KI-system forskjellsbehandler personer på bakgrunn av for eksempel kjønn eller etnisitet, må slik informasjon foreligge. Det gjør det som regel ikke. KI-forordningen åpner unntaksvis for å innhente slik informasjon for test av skjevheter i KI-prosjekter med høy risiko.²⁹¹ En generell definisjon og analyse av diskriminering i eksempelvis helsevesenet burde dermed først være på plass som grunnlag for å samle inn nødvendig informasjon på sentralt nivå – ikke på bakgrunn av det enkelte KI-prosjekt.²⁹² Relasjon mellom algoritmeskjevhet og diskriminering er i tillegg kompleks, og rettferdighet som et kontekstuellet betinget normativt begrep

²⁹⁰ Se også: Broomfield, H. og Reutter, L. (2019). *Kunstig intelligens/ data science: En kartlegging av status, utfordringer og behov i norsk offentlig sektor. Første resultater.*

²⁹¹ European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act), P9_TA(2024)0138, artikkel 10, 5.

²⁹² Intervju med Akershus universitetssykehus (20. november 2023).

sjelden dekket av matematiske definisjoner som målbare metrikker for skjevhet.²⁹³ Slike grunnleggende avveining mellom personvern og sikring av rettferdighet i KI-systemer er et verdivalg som er relevant for flere sektorer der KI-systemer skal utvikles og tas i bruk, ikke bare i helsesektoren. Etter vår vurdering må den enkelte virksomhet i for stor grad håndtere slike spørsmål på egen hånd.

På bakgrunn av utfordringene ved utvikling av kunstig intelligens i staten vi har påvist, mener vi at Digitaliserings- og forvaltningsdepartementet ikke i tilstrekkelig grad har ivaretatt rollen som samordningsdepartement i oppfølgingen og gjennomføringen av alle tiltakene som skulle sikre at Norge skulle få en infrastruktur for kunstig intelligens i verdensklasse. Men det er også viktig å presisere at alle sektorer og departementer har et felles ansvar for å sikre god måloppnåelse på dette området. I dag fremstår det derfor uklart hvordan Digitaliserings- og forvaltningsdepartementet, sammen med de øvrige departementene, i fellesskap skal gjennomføre politikken som skal sikre forutsetninger som legger til rette for bruk av kunstig intelligens i staten. En ren videreføring av dagens samordningsaktiviteter vil etter vår vurdering neppe sikre at det store potensialet for bruk av kunstig intelligens innenfor ansvarlige rammer blir utnyttet.

²⁹³ Storås, A.M., Prabhu, R., Hammer, H. L. & Strümke, I. (2022), Bias og kvantitativ analyse innen velferd: opphav til skjevheter og relasjon til utfallsrettferdighet. Tidsskrift for velferdsforskning: Vol 25, No 3. <https://doi.org/10.18261/tfv.25.3.3>

9 Referanseliste

Stortingsdokumenter

Innstillinger

- Innst. 84 S (2016–2017) Innstilling fra transport- og kommunikasjonskomiteen om Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet.
- Innst. 16 S (2020–2021) Innstilling fra kommunal- og forvaltningskomiteen om bevilgninger på statsbudsjettet for 2021.
- Innst. 191 S (2020–2021) Innstilling fra kommunal- og forvaltningskomiteen om En innovativ offentlig sektor – Kultur, ledelse og kompetanse.
- Innst 568 S (2021–2022) Innstilling fra næringskomiteen om Data som ressurs – Datadrevet økonomi og innovasjon.

Proposisjoner

- Prop. 1 S (2023–2024) Helse- og omsorgsdepartementet.
- Prop. 1 S (2019–2020) Kommunal- og distriktsdepartementet.
- Prop. 1 S (2021–2022) Kommunal- og distriktsdepartementet.
- Prop. 1 S (2022–2023) Kommunal- og distriktsdepartementet.
- Prop. 1 S (2023–2024) Kommunal- og distriktsdepartementet.
- Prop. 1 S (2023–2024) Kunnskapsdepartementet.

Stortingsmeldinger

- Meld. St. 27 (2015–2016) Digital agenda for Norge – IKT for en enklere hverdag og økt produktivitet.
- Meld. St. 7 (2019–2020) Nasjonal helse- og sykehusplan 2020–2023.
- Meld. St. 30 (2019–2020) En innovativ offentlig sektor – Kultur, ledelse og kompetanse.
- Meld. St. 22 (2020–2021) Data som ressurs – Datadrevet økonomi og innovasjon.
- Meld. St. 14 (2022–2023) Utsyn over kompetansebehovet i Norge.

Lover og forskrifter

Lover

- Kongeriket Norges grunnlov (grunnloven).
- Lov om behandlingsmåten i forvaltningsaker (forvaltningsloven) (01.01.1970).
- Lov om behandling av personopplysninger (personopplysningsloven) (20.07.2018).
- Lov om rett til innsyn i dokument i offentlig verksemd (offentleglova) (01.01.2009).
- Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven) (01.01.2018).
- Lov om styrking av menneskerettighetenes stilling i norsk rett (menneskerettsloven) (21.05.1999).
- Lov om virksomheters åpenhet og arbeid med grunnleggende menneskerettigheter og anstendige arbeidsforhold (åpenhetsloven) (01.07.2022).
- Lov om nasjonal sikkerhet (sikkerhetsloven) (01.01.2019).
- Lov om Statens pensjonskasse (01.08.1949).
- Lov om arbeids- og velferdsforvaltningen (arbeids- og velferdsforvaltningsloven) [NAV-loven] (01.07.2006).
- Lov om skatteforvaltning (skatteforvaltningsloven) (01.01.2017).
- Lov om medisinsk utstyr (07.05.2020, 26.05.2021).
- Lov om opphavsrett til åndsverk mv. (åndsverkloven) (01.07.2018).
- Lov om språk (språklova) (01.01.2022).
- Lov om Sametinget og andre samiske rettsforhold (sameloven) (24.02.1989).

- Lov om arkiv (arkivloven) (01.01.1999).

Internasjonale avtaler og konvensjoner

- Den europeiske menneskerettskonvensjon (EMK) (4. november 1950).
- Charter of Fundamental Rights of the European Union, EUR-Lex, Document 12012P. [Hentet 19. april 2024 fra <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>].

Reglement

- Bevilgningsreglementet vedtatt gjennom St.prp. nr. 48 (2004–2005) *Om bevilgningsreglementet*, jf. Innst. S. nr. 187 (2004–2005).
- Reglement for økonomistyring i staten fastsatt 12. desember 2003 med endringer, senest 31. august 2021 (økonomireglementet).
- IVDR: In-vitro diagnostic medical device regulation, Regulation (EU) 2017/746.

Rapporter fra Riksrevisjonen

- Riksrevisjonens Dokument 3:8 (2023–2024). Riksrevisjonens undersøkelse av myndighetenes tilrettelegging for deling og gjenbruk av data i forvaltningen.

Styringsdokumenter

- Digitaliserings- og forvaltningsdepartementet (Kommunal- og distriktsdepartementet). (2019). Tildelingsbrev til Digitaliseringsdirektoratet for 2019.
- Digitaliserings- og forvaltningsdepartementet (Kommunal- og distriktsdepartementet). (2020). Tildelingsbrev til Digitaliseringsdirektoratet for 2020.
- Digitaliserings- og forvaltningsdepartementet (Kommunal- og distriktsdepartementet). (2021). Tildelingsbrev til Digitaliseringsdirektoratet for 2021.
- Digitaliserings- og forvaltningsdepartementet (Kommunal- og distriktsdepartementet). (2022). Tildelingsbrev til Digitaliseringsdirektoratet for 2022 og supplerende tildelingsbrev nr. 1, mars 2022.
- Digitaliserings- og forvaltningsdepartementet (Kommunal- og distriktsdepartementet). (2023). Tildelingsbrev til Digitaliseringsdirektoratet for 2023.
- Finansdepartementet. (2023). Tildelingsbrev til Direktoratet for forvaltning og økonomistyring, Finanstilsynet, skatteetaten, Statistisk sentralbyrå og Tolletaten for 2023.
- Helse- og omsorgsdepartementet. (2024). Oppdragsdokument 2024, Helse Sør-Øst RHF.
- Kommunal- og moderniseringsdepartementet. (2021). Økonomi- og virksomhetsinstruks for Digitaliseringsdirektoratet, 1. januar 2021.
- Kultur- og likestillingsdepartementet. (2024). Tildelingsbrev til Nasjonalbiblioteket for 2021.

Strategier og handlingsplaner

- Kommunal- og moderniseringsdepartementet. (2019). Én digital offentlig sektor, Digitaliseringsstrategi for offentlig sektor 2019–2025.
- Kommunal- og moderniseringsdepartementet. (2020). Nasjonal strategi for kunstig intelligens. <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/>
- Kommunal- og moderniseringsdepartementet. (2015). Handlingsplan for informasjonssikkerhet i statsforvaltningen – 2015–2017.
- Skatteetaten. (2023). Policy for utvikling og bruk av KI i skatteetaten.

Årsrapporter

- Arbeidstilsynet. (2023). Felles årsrapport for etatenes innsats mot arbeidslivskriminalitet for 2022.
- Barne-, ungdoms- og familiedirektoratet. (2023). Årsrapport for 2022.
- Brønnøysundregistrene. (2022). Årsrapport for 2021.

- Digitaliseringsdirektoratet. (2022). Årsrapport for 2021.
- Digitaliseringsdirektoratet. (2023). Årsrapport for 2022.
- Direktoratet for e-helse. (2022). Årsrapport for 2021.
- Direktoratet for høyere utdanning og kompetanse. (2023). Årsrapport 2022.
- Skatteetaten. (2022). Årsrapport for 2021.
- Sivilombudsmannen. (2020). Årsmelding for 2019.
- Sivilombudsmannen. (2021) Årsmelding for 2020.

Rapporter, planer og utredninger

Offentlige utredninger

- NOU 2019: 5 Ny forvaltningslov – Lov om saksbehandlingen i offentlig forvaltning (forvaltningsloven).
- NOU 2022: 1 Ditt personvern – vårt felles ansvar.
- NOU 2023: 4 Tid for handling – Personellet i en bærekraftig helse- og omsorgstjeneste.

Rapporter

- Agenda Kaupang, på oppdrag fra Datatilsynet. (2023). [Evaluering av Datatilsynets sandkasse for kunstig intelligens](#).
- Alan Turing Institute. (2024). *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*.
- Broomfield, H. og Reutter, L. (2019). *Kunstig intelligens / data science: En kartlegging av status, utfordringer og behov i norsk offentlig sektor. Første resultater*.
- Datatilsynet. (2022). NAV – [Sluttrapport fra sandkasseprosjektet med NAV](#). Temaer: rettslig grunnlag, rettferdighet og forklarbarhet.
- Datatilsynet. (2023). *Hjerterom for etisk AI – Sluttrapport fra sandkasseprosjektet til Ahus (EKG AI)* Tema: Algoritmeskjevhet og rettferdige algoritmer.
- Digitaliseringsdirektoratet. (2022). *Viktige tiltak for digitaliseringen i offentlig sektor*.
- Direktoratet for høyere utdanning og kompetanse. (2021). [Behovet for IKT-kompetanse i Norge - En vurdering av kunnskapsgrunnlaget](#).
- Ekspertgruppe nedsatt av kjernegruppen for et bedre kunnskapssystem for håndtering av kriser. (2022). *Juridiske og etiske problemstillinger knyttet til innsamling, tilgjengeliggjøring, deling og bruk av data, samt bruk av randomiserte forsøk, i kriser*.
- Europaparlamentet. (2021). *Artificial intelligence and public services*. [PE 662.936](#).
- Hurtigarbeidende arbeidsgruppen ledet av Justis- og beredskapsdepartementet, nedsatt av regjeringen for å utarbeide en plan for rettidig innlemming og god gjennomføring av EUs forordning om kunstig intelligens, samt vurdere nasjonalt reguleringsbehov på sivil side. (2023). *Regulering av kunstig intelligens*. Arbeidsgrupperapport.
- Menon Economics. (2023). *KI: Betydning for arbeidsstyrken. En analyse av potensialet for kunstig intelligens-drevet effektivisering i norsk næringsliv*.
- National Audit Office. (2024). [Use of Artificial Intelligence in Government](#).
- Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU) og SINTEF. (2023). *Kompetanse for grønn omstilling – En gjennomgang av forskningslitteratur og arbeidslivets kompetansebehov knyttet til miljø- og klimautfordringer*, [Rapport 2023:5](#).
- NSM. (2023). *Konseptvalgutredning for nasjonal skytjeneste*. <https://nsm.no/regelverk-og-hjelp/rapporter/konseptvalgutredning-for-nasjonal-skytjeneste>.
- OECD. (2023). *The State of Implementation of the OECD AI principles four years on*. OECD Artificial Intelligence Papers, No. 3. <https://doi.org/10.1787/835641c9-en>.
- OECD. (2024). *The impact of artificial intelligence on productivity, distribution and growth. Key mechanisms, initial evidence and policy challenges*. [OECD artificial intelligence papers No. 15](#).

- Rambøll Management Consulting og Vestlandsforskning, på oppdrag fra Barne-, ungdoms- og familiedirektoratet (Bufdir). (2022). Bruk av kunstig intelligens i offentlig sektor og risiko for diskriminering. Kunnskapsgrunnlag for arbeidet med å forebygge diskriminerende effekter ved bruk av kunstig intelligens i offentlig virksomhet, [VF-rapport nr. 7-2022](#).
- Rambøll Management Consulting AS. (2023). *IT i praksis 2023*.
- Skate. (2024). *Skates erfaringsrapport 2023*.
- Skate-møte 14. juni 2023 – referat <https://www.digdir.no/skate/skate-mote-14-juni-2023-referat/4690>.
- Teknologirådet. (2021). *Kunstig intelligens i klinikken – seks trender for fremtidens helsetjeneste*. Rapport. Hentet fra ([Teknologirådet \(teknologiradet.no\)](#))
- V. B. Strand, i oppdrag av Likestillings- og diskrimineringsombud. (2024). *Algoritmer, kunstig intelligens og diskriminering. En analyse av likestillings- og diskrimineringslovens muligheter og begrensninger*.

Veiledere, standarder og rammeverk

- Datatilsynet. (2018). *Iverksette styringssystem for informasjonssikkerhet*. <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/informasjonssikkerhet-internkontroll/etablere-internkontroll/iverksette-styringssystem-for-informasjonssikkerhet/>.
- Datatilsynet. (2023). *Vurdering av personvernkonsekvenser (DPIA)*. Hentet fra <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/vurdering-av-personvernkonsekvenser/>.
- Datatilsynet. (u.å.). [Sjekkliste for vurdering av personvernkonsekvenser \(DPIA\)](#).
- Datatilsynet. (2022). *Innebygd personvern og personvern som standard*. [Hentet: 23. november 2023 fra <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/innebygd-personvern-og-personvern-som-standard/>].
- Datatilsynet. (2023). [Personvernprinsippene](#).
- Datatilsynet. (2023). *Databehandleravtale* [Hentet fra [Hvordan lage en databehandleravtale? | Datatilsynet](#)].
- Datatilsynet. (2018). *Informasjonssikkerhet og internkontroll*. [Hentet: 21. november 2022 fra [Virksomhetenes plikter | Informasjonssikkerhet og internkontroll](#)].
- Digitaliseringsdirektoratet. (u.å.). *Innebygd informasjonssikkerheit*. Hentet 23. mai 2024 fra <https://www.digdir.no/informasjonssikkerhet/innebygd-informasjonssikkerheit/2146>.
- Digitaliseringsdirektoratet. (u.å.). *Veileder for orden i eget hus*. [Hentet: 12. februar 2024 fra <https://www.digdir.no/informasjonsforvaltning/veileder-orden-i-eget-hus/2716>].
- Digitaliseringsdirektoratet. (u.å.). *Hvordan vurderer jeg risiko ved bruk av kunstig intelligens?* [Hentet: 29. april 2024 fra <https://www.digdir.no/kunstig-intelligens/hvordan-vurderer-jeg-risiko-ved-bruk-av-kunstig-intelligens/4537>].
- Digitaliseringsdirektoratet. (u.å.). *Råd for ansvarlig utvikling og bruk av kunstig intelligens i offentlig sektor*. [Hentet: 24. mai 2024 fra <https://www.digdir.no/kunstig-intelligens/rad-ansvarlig-utvikling-og-bruk-av-kunstig-intelligens-i-offentlig-sektor/4272>].
- Digitaliseringsdirektoratet. (u.å.). *Prosjektveiviseren*. [Hentet: 23. november 2023 fra [Prosjektveiviseren \(digdir.no\)](#)].
- Digitaliseringsdirektoratet. (u.å.). *Åpenhet og kunstig intelligens*. [Hentet: 11. april 2024 fra <https://www.digdir.no/kunstig-intelligens/apenhet-og-kunstig-intelligens/4581>].
- Digitaliseringsdirektoratet, Digdir.no. (u.å.) *Begrepsliste*. [Hentet: 18. november 2023 fra <https://www.digdir.no/informasjonssikkerhet/begrepsliste/3230#informasjonssikkerhet>].
- Digitaliseringsdirektoratet. (u.å.). ISO/IEC 27002 [Hentet: 17. november 2023 fra <https://www.digdir.no/standarder/isoiec-27002/1696>].
- Direktoratet for medisinske produkter. (Oppdatert 2023). *Samsvarserklæring* [Hentet fra <https://www.dmp.no/medisinsk-utstyr/utvikling-og-produksjon/samsvarsvurdering/samsvarserklæring>].

Brev, notater og nettkilder

Brev

- Brev fra Helse- og omsorgsdepartementet til Riksrevisjonen, 29. februar 2024.
- Brev fra Kultur- og likestillingsdepartementet til Riksrevisjonen, 19. februar 2024.
- Epost fra daværende Kommunal- og Distriktsdepartementet, 09. mars 2023.
- Nasjonalbiblioteket svar på brev til Kultur- og likestillingsdepartementet av 31. august 2023 (elinnsyn).

Nettkilder

- Council of Europe. (u.å.). *CAHAI – Ad hoc Committee on Artificial Intelligence – Artificial Intelligence* (Hentet: 10. juni 2024 fra coe.int).
- Dagens Medisin. (2024). *Vestre Viken har innført benbrudd-KI på alle sykehus – over 11.000 pasienter sjekket*. [Hentet fra dagensmedisin.no].
- Datatilsynet. (u.å.). *Sandkassesiden*. [Hentet: 12. februar 2024 fra [Sandkassesiden | Datatilsynet](https://sandkassesiden.no)].
- McKinsey & Company. (2023). *Unlocking the potential of generative AI: Three key questions for government agencies* [Hentet fra mckinsey.com].
- Nasjonal kommunikasjonsmyndighet (NSM). (2023). 43 prosent bekymret for bruk av kunstig intelligens. Nettartikkel [Hentet fra nkom.no/aktuelt].
- Skatteetaten. (2023). *Slik kan vi bruke kunstig intelligens på en sikker og ansvarlig måte*. Nettartikkel [Hentet fra skatteetaten.no].
- Teknologirådet. (2024). *Gjennombruddet for generativ kunstig intelligens – en tidslinje*. [Nettartikkel](#).

Faglitteratur

- Boe, E. M. (2022). *Læren om myndighetsmisbruk i forvaltningen: krav til saklighet og forsvarlighet*. Jussens Venner, Vol. 57, Utg. 4. doi.org/10.18261/jv.57.4.
- Elgesem, F. (2021). *Tolkning av den europeiske menneskerettighetskonvensjonen*, Jussens Venner, Vol 55, Utg. 6, s. 441-487.
- Gebru T., Morgenstern J. Vecchione B., Vaughan, J. W., Wallach, H. M., Daumé III, H. & Crawford, K. (2018). *Datasheets for Datasets*. [arXiv:1803.09010](https://arxiv.org/abs/1803.09010).
- Lundberg, S. M. & Lee, S. (2017). *A unified approach to interpreting model predictions*. *Advances in Neural Information Processing Systems*.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D. & Gebru, T. (2018). *Model Cards for Model Reporting*. [arXiv:1810.03993](https://arxiv.org/abs/1810.03993).
- Miconi, T. (2017). *The impossibility of "fairness": a generalized impossibility result for decisions*. [arXiv:1707.01195](https://arxiv.org/abs/1707.01195)
- Raz, A., Heinrichs, B., Avnoon, N., Eyal, G., & Inbar, Y. (2024). *Prediction and explainability in AI: Striking a new balance?* *Big Data & Society*, 11(1). <https://doi.org/10.1177/20539517241235871>.
- Ruf, B. & Detyniecki, M. (2020). *Active Fairness Instead of Unawareness*. [arXiv:2009.06251](https://arxiv.org/abs/2009.06251).
- Saleiro, P., Kuester B., Hinkson, L., London J., Stevens A., Anisfeld, A., Rodolfa, K. T. & Ghani, R. (2018). *Aequitas: A Bias and Fairness Audit Toolkit*. [arXiv:1811.05577](https://arxiv.org/abs/1811.05577).
- Sánchez Ávalos, S., González, F. & Ortiz, T. (2021). *Responsible use of AI for public policy: Data science toolkit*. [OECD.AI](https://doi.org/10.18235/0002876). <http://dx.doi.org/10.18235/0002876>.
- Smith, E. & Eckhoff, T. *Forvaltningsrett*. Universitetsforlaget, 11. utgave.
- Storås, A.M., Prabhu, R., Hammer, H. L. & Strümke, I. (2022). *Bias og kvantitativ analyse innen velferd: opphav til skjevheter og relasjon til utfallsrettferdighet*. *Tidsskrift for velferdsforskning*: Vol 25, No 3. <https://doi.org/10.18261/tfv.25.3.3>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. og Polosukhin, I. (2017). *Attention is all you need*. [arXiv:1706.03762](https://arxiv.org/abs/1706.03762).

- Verma, S. & Rubin, J. (2018). *Fairness Definitions Explained*, Proceedings of the IEEE/ACM International Workshop on Software Fairness, <https://doi.org/10.1145/3194770.3194776>.

Annet

- Digitaliseringsdirektoratet. (2023). *Innspill til ny nasjonal digitaliseringsstrategi*. [Hentet fra [Nyheitsarkiv](#)].
- European Court of Human Rights. *Niemietz v. Tyskland*, [J], no. 13710/88, (1992) 16 EHRR 97 og *von Hannover v. Tyskland*, [J] no. 59320/00 ECHR 2004-VI ECHR 294.
- European Parliament. (2024). *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*, Commission draft (2021), [Final draft](#) (2024).
- Muntlig spørretime på Stortinget, 8. mai 2024: [Sak nr. 1 \[10:01:35\] - stortinget.no](#). Hentet 13. juni 2024.

10 Vedlegg

10.1 Vedlegg 1: Brev med spørsmål

Følgende spørrebrev ble sendt til statlige virksomheter og heleide selskaper. Noen mindre justeringer vil forekomme fordi brevene ble sendt til ulike virksomhetstyper.

Kartlegging – Bruk av kunstig intelligens i staten

Riksrevisjonen gjennomfører en forvaltningsrevisjon av bruk av kunstig intelligens (KI) i staten. Revisjonen er per nå rettet mot Kommunal- og distriktsdepartementet. Formålet med revisjonen er å vurdere om statlige virksomheter og heleide selskaper utnytter mulighetene med KI på en ansvarlig måte i samsvar med Stortingets vedtak og forutsetninger.

I denne sammenheng gjennomfører vi en kartlegging av bruk av KI i offentlig sektor, som omfatter alle statlige virksomheter og heleide selskaper. Kartleggingen omfatter KI-systemer som er i bruk og pågående prosjekter for å ta i bruk KI.

Vennligst se definisjon og avgrensning av KI som er relevant for revisjonen nederst i brevet. Dersom bruk av KI ikke er aktuelt for dere, ber vi om at dere kort opplyser om dette gjennom å svare nei på spørsmålene under.

I første omgang skal undersøkelsen gi en oversikt over muligheter, utvikling, bruk og utfordringer med KI. Derfor ber vi om følgende opplysninger:

- (1) Svar på spørsmålene nedenfor.
- (2) Dersom dere aktivt bruker eller utvikler KI, ber vi at dere oppgir en eller flere kontaktperson(er) med faglig ansvar for KI-systemene. Kontaktpersonene vil senere motta en spørreundersøkelse fra Riksrevisjonen med mer detaljerte spørsmål om deres bruk av KI.

Vi ber om svar innen [dato]. På forhånd takk for hjelpen.

Eventuelle spørsmål kan rettes til prosjektleder [navn, kontaktinformasjon].

(1) Spørsmål til alle

Vi ber om at dere svarer på følgende spørsmål, gjerne **kort** og med **stikkord**. Se definisjon av *KI-system* og avgrensning av revisjonen nedenfor.

1. Har dere tatt i bruk, eller har dere pågående prosjekter for å ta i bruk, KI etter definisjonen som er gitt nederst i brevet og med unntak av de avgrensningene som er definert?
 - Hvis ja: Vennligst beskriv kort til hvilke formål (anvendelsesområde og ønsket effekt). Bruk gjerne en liste med et kulepunkt per formål. Vennligst oppgi kontaktpersoner i avsnitt (2).
2. Har dere brukt KI til oppgaver i et tidsavgrenset prosjekt (engangsbruk) som er avsluttet, i tidsrommet 2018 og frem til nå?
 - Hvis ja: Vennligst beskriv kort til hvilke formål (anvendelsesområde og ønsket effekt).
3. Har dere avbrutt eller stoppet en prosess hvor dere vurderte, utviklet eller brukte KI, i tidsrommet 2018 og frem til nå?
 - Hvis ja: Vennligst beskriv kort formål (anvendelsesområde og ønsket effekt) av KI-systemet, og årsakene til at prosjektet ble stanset.

4. Vurderer dere å benytte KI, utover det som eventuelt allerede er implementert eller er i utvikling?
 - Hvis ja: Vennligst beskriv kort til hvilke formål (anvendelsesområde og ønsket effekt).
5. Hvis dere har svart ja på ett eller flere spørsmål over, vennligst beskriv utfordringer og/eller hindringer dere har møtt med å ta i bruk KI.

(2) Kontaktperson(er) for spørreundersøkelse om KI-systemer

Dersom dere har svart ja på spørsmål 1 eller 2 over, ber vi at dere oppgir en eller flere kontaktperson(er) med faglig ansvar for KI-systemene. Disse vil motta en spørreundersøkelse med mer detaljerte spørsmål om KI-systemene.

Det er ønskelig med kontaktpersoner med både teknisk forståelse og kunnskap om prosjektets organisasjon, for hvert KI-system. Dette kan bety en kontaktperson med god oversikt over alle/flere KI-systemer i virksomheten (som f.eks. en leder av en utviklingsseksjon, chief data scientist osv.), eller flere ulike kontaktpersoner med ansvar for ett eller noen få KI-system(er).

Kontaktpersonen bør ha et faglig ansvar for implementeringen. Hvis det er flere ulike organisasjonsheter som benytter KI, ber vi om at dere oppgir én kontaktperson for hver av disse avdelingene.

Bruk gjerne en tabell dersom dere har flere kontaktpersoner:

Navn	epost	Stillingstittel

Definisjon KI-system

Med kunstig intelligens mener vi systemer som utfører handlinger basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Kartleggingen følger dermed definisjonen i den nasjonale strategien for kunstig intelligens. Tolking av data kan være fullstendig regelbasert (programvarerobot), basert på en kunnskapsbase (ekspertsystem), eller basert på data gjennom statistisk modellering eller maskinlæring.

Revisjonen gjelder KI-systemer som virksomhetene utvikler eller tilpasser selv, enten in-house eller gjennom oppdrag til eksterne konsulenter.

Vi har avgrenset oss fra følgende KI-systemer i vår revisjon:

- Hylleware som verken er utviklet for eller tilpasset til virksomheten (f.eks. KI innebygd i spam-filter eller programvare for IT-sikkerhet eller regnskap), og som er brukt på samme type data systemet ble utviklet for og derfor ikke krever validering.

Revisjonen skal omfatte hylleware som krever validering pga. usikkert resultat i planlagt anvendelsesområde (inndata), som f.eks. i medisinsk diagnostikk. Helse- og omsorgstjenesten kan bruke data fra den populasjonen et produkt skal virke på, for å validere et KI-produkt.

- Undersøkelsen skal ikke omfatte bruk som faller inn under sikkerhetsloven.
- Undersøkelsen skal ikke omfatte utvikling av KI i forskningsprosjekter. Undersøkelsen vil derimot omfatte KI under utvikling der resultat av utviklingsprosessen skal kunne brukes i produksjon (f.eks. i utøvende helsetjenester). Delprosjekter forankret i forskning skal inkluderes hvis hovedprosjekt er rettet mot utøvende helsetjeneste, helseforvaltning, saksbehandling, andre kjerneoppgaver, virksomhetens administrasjon eller drift.
- Bruk av KI i klinisk sammenheng (f.eks. diagnostikk) skal ikke omfattes av undersøkelsen.

Med *utvikling* mener vi at et prosjekt er påbegynt, men KI-systemet ble ikke enda benyttet for virksomhetenes oppgaver.

Med *i bruk* mener vi KI-systemer som er benyttet i virksomhetens oppgaver, selv om systemet samtidig er utviklet videre, i førstegangsbruk (rollout) og/eller brukt i et tidsavgrenset prosjekt.

Hvis dere er usikre på om et system eller programvare dere bruker faller inn under definisjonen av KI som er relevant for denne revisjonen, vennligst inkluder systemet med en kort beskrivelse i deres svar.

Hvis dere har i bruk eller utvikler ett KI-system som faller under definisjonen over, ber vi om en kontaktperson som kan besvare en spørreundersøkelse om KI-systemer (se avsnitt (2)).

10.2 Vedlegg 2: Spørreskjema

Spørreundersøkelsen ble gjennomført i Questback.

Introduksjonsside til spørreundersøkelsen:

Du ble angitt som ansvarlig for ett eller flere KI-systemer i din virksomhet: Vennligst svar på spørsmålene ut fra dine erfaringer med KI. Spørsmålene handler om hvordan du, og eventuelt kolleger i samme KI-prosjekt, jobber med KI i praksis i enkelte KI-prosjekter.

Besvar spørsmålene med tanke på den organisasjonsenheten med ansvar for det enkelte KI-systemet, for eksempel en avdeling, seksjon eller et prosjektteam. Spørsmålene skal ikke besvares med tanke på hele virksomheten/selskapet.

Vi legger til grunn at det er et stort mangfold i erfaringer, formål og ressurser tilgjengelig til utvikling og bruk av kunstig intelligens hos statlige virksomheter og selskaper. Spørreskjema er utformet for å kunne registrere informasjon om kunstig intelligens med bakgrunn i dette mangfoldet.

Spørreundersøkelsen er inndelt i to hoveddeler; detaljert informasjon om utvalgte KI-systemer, og generelle erfaringer med kunstig intelligens.

Vi bruker følgende begreper i undersøkelsen:

Kunstig intelligens (KI) er informasjonsteknologi som utfører handlinger basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Undersøkelsen følger dermed definisjonen i den nasjonale strategien for kunstig intelligens. Tolking av data kan være fullstendig regelbasert (programvarerobot), basert på en kunnskapsbase (ekspertsystem), eller basert på data gjennom statistisk modellering eller maskinlæring.

KI-system: omfatter hele anvendelsen fra innhenting av data til sluttresultat, ikke bare KI-komponenten som tolker data med for eksempel en maskinlæringsmodell.

KI-prosjekt: Organiseringen av et KI-system, fra planlegging til produksjon eller avviking. Begrepet er uavhengig av en formell struktur med prosjektnummer eller lignende, men kan defineres via målet KI-systemet skal bidra til.

Enhet: Med enhet mener vi organisasjonsenheten med ansvar for KI-systemet, for eksempel en avdeling, seksjon eller et prosjektteam.

Hylleware: Ferdig trente modeller, med mulighet/nødvendighet for tilpassing til egne data

Undersøkelsen tar mellom 15–45 minutter å besvare. Tidsbruken vil avhenge av antall KI-systemer du svarer for.

Spørreundersøkelse om bruk av kunstig intelligens

Oversikt

2) * Hva er din primære rolle i utviklingen av KI-systemer?

- Data scientist, utvikler
- Domeneekspert, ikke-teknisk fagperson
- Produktansvarlig eller prosjektleder
- Leder for data science, KI eller IT
- Annet

3) * Hvor mange KI-systemer kan du svare for?

Dersom det er flere respondenter fra din virksomhet, kryss av for hvor mange KI-systemer du svarer for i dette spørreskjemaet. Hvis du er den eneste respondenten i virksomheten, er dette det samme som det totale antallet KI-systemer.

- 1 2 3 4 5 6 7 8 9 10 Flere enn 10

4) Hvor mange av disse KI-systemene bruker:

Vennligst velg den kategorien som passer best for KI-prosjektene.

- Regelbasert: Robot process automation (RPA), enkel chatbot, enkelt søk
- Statistisk eller enkel maskinlæring: Typisk enkel *ML* er f. eks. lineær eller logistisk regresjon, beslutningstre, naive bayes (hvit boks)
- Mer avansert maskinlæring: *ML* som er mer utfordrende å forklare, f. eks. random forest, boosted trær (XGBoost, Catboost osv.), *SVM*, nevralt nett, dyplæring, generative eller additive modeller

	0	1	2	3	4	5	6	7	8	9	10	Flere enn 10
Regelbasert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statistisk eller enkel maskinlæring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mer avansert maskinlæring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

KI-systemer videre i undersøkelsen

Vennligst oppgi navn for KI-system(ene) du skal svare for, slik at KI-systemene kan identifiseres dersom vi har spørsmål etter spørreundersøkelsen. Bruk gjerne korte, beskrivende navn.

Du svarte at du er satt som respondent for flere enn tre KI-systemer. I resten av spørreskjemaet ønsker vi mer informasjon om **tre** av disse.

Vennligst velg ut **tre** prosjekter basert på følgende kriterier:

- Det teknologisk mest avanserte prosjektet (f. eks. dyplæring over regelrobot)
- Prosjektet som anslagsvis kostet mest å utvikle (enten personalkostnad for egne utvikling, eller kostnad for innkjøp av konsulenter eller programvare)
- Prosjektet med størst konsekvenser for personer. Dersom (de øvrige) KI-systemene ikke har konsekvenser for personer, velg systemet med størst betydning for virksomheten

Prosjekt 1:

Prosjekt 2:

Prosjekt 3:

8) Har du noen kommentarer relatert til denne delen av undersøkelsen?

Informasjon om KI-systemene

9) * Er KI-systemet per i dag i bruk eller under utvikling?

	Under		
	I bruk	utvikling	Annet
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10) * Hvilken type teknologi bruker KI-systemet for tolking av data?

Vennligst velg kategori som passer best for KI-prosjektene:

- Regelbasert: Robot process automation (RPA), enkelt chatbot, enkelt søk
- Statistisk eller enkel maskinlæring: Typisk enkelt *ML* er f. eks. lineær eller logistisk regresjon, beslutningstre, naive bayes (hvit boks)
- Mer avansert maskinlæring: *ML* som er mer utfordrende å forklare, f. eks. random forest, boosted trær (XGBoost, Catboost osv.), *SVM*, nevralt nett, dyplæring, generative eller additive modeller

	Regelbasert	Statistisk eller enkel ML	Mer avansert ML	Vet ikke
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11) Analyserer KI-systemet tekst, tale eller bilder?

Tekst er skrevet språk, ikke f. eks. tidsserier.

Bilder inkluderer alt bildemateriell, f. eks. video og skannede dokumenter.

	Tekst	Tale	Bilder	Ingen av disse
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12) * Benytter KI-systemet sanntidsdata?

Sanntidsdata blir behandlet av KI-systemet umiddelbart, og utfallet sendt umiddelbart til sluttbrukeren.

	Ja	Nei	Vet ikke
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

13) * Hva er resultatet av tolkingen av data i KI-systemet?

- Informasjonshenting: Sammenstille relevant informasjon fra en databank, f. eks. i søk, chatbots
- Klassifisering: Tildeling til en gruppe, ev. med en sannsynlighetsverdi

- Punktestimat: Estimat av en numerisk verdi, f. eks. en pris, dato, tidsrom, alder osv.
- Generativ: Generering av nytt innhold basert på input

	Informasjonshenting	Klassifisering	Punktestimat	Generativ	Annet
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14) * Hvor ofte kjøres KI-systemet?

	Kontinuerlig (automatisk system)	Periodisk	Ved behov (for eksempel ved brukerforespørsel)	Annet
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15) * Behandler dere persondata i forbindelse med KI-anvendelsen?

GDPR artikkel 4 definerer persondata som enhver opplysning om en identifisert eller identifiserbar fysisk person.

Bruk av data under utvikling regnes som behandling, selv om dataene ikke blir brukt i den endelige versjonen av KI-systemet.

	Vet		
	Ja	Nei	ikke
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

16) * Hvem har utviklet eller utvikler KI-systemet?

Kryss av for alt som passer.

	Egne ansatte	Konsulenter	Leverandør	Andre samarbeidspartnere	Annet
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

17) Hvilke av følgende verktøy bruker eller brukte dere under utvikling av KI-systemet?

Kryss av for alt som passer per prosjekt:

- **Ferdigtrent modell:** en modell utviklet av samarbeidspartnere eller leverandør. Eventuelt kan modellen tilpasses til egne data.
- **KI / ML-plattform:** Programvare som hjelper å trene modeller uten bruk av programmering, ofte brukt til å sammenligne mange ulike ML-modeller med anbefaling av beste modell. Populære tilbydere er f. eks. DataRobot, Azure Machine Learning automated ML, Amazon aws SageMaker, ElasticSearch Machine Learning.
- **Open source kodebiblioteker:** Open source koderammeverk med biblioteker for modellutvikling (R, Python, inkl. Tensorflow, Keras etc.)
- **Lukket kildekode:** Kjøpt programvare, f. eks. SPSS

	Ferdigtrent modell	KI/ML-plattform	Open source kodebiblioteker	Lukket kildekode
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

18) * Har dere benyttet skytjenester i utviklingen eller i produksjonen?

	Ja, i utviklingen/treningen	Ja, KI-systemet kjører i sky	Nei	Vet ikke
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

19) * Hvilke oppgaver brukes eller utvikles KI-systemet for?

Saksbehandling er primært relevant for virksomheter.

Tjenesteleveranser eller andre kjerneoppgaver er vanligvis rettet mot eksterne brukere.

Som **interne prosesser** regnes for eksempel drift, administrasjon, HR, sikkerhet, strategi og styring etc.

Kryss av for **hovedoppgaven** som passer best per prosjekt.

	Saksbehandling	Tjenesteleveranser eller andre kjerneoppgaver	Interne prosesser (f. eks. HR, logistikk, arkiv)	Beslutningsstøtte for strategi og styring av virksomheten/selskapet	Utprøving av ny teknologi/kunnskapsutvikling
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

20) * Hva er hovedformålet med KI-prosjektet og i hvilken grad blir formålet nådd?

Velg **hovedformål** med prosjektet, og hvorvidt formålet ble oppnådd:

	Formål 1						
	Redusere kostnadene eller øke inntjeningen i virksomheten	Raskere saksbehandling eller bearbeidelse av henvendelser	Frigjøre tid til å løse andre eller mer krevende oppgaver	Redusere bemanningen	Forbedre kjerneoppgaver eller lage nye tjenester/tilbud	Uttesting av KI eller intern kompetansebygging	Bedre informasjon som grunnlag for beslutninger
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

21) * Har dere undersøkt nytteeffekter av KI-bruken?

	Ja, kvantitativt	Ja, kvalitativt	Nei	Ikke relevant
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

22) Hvilke andre formål har bruken av KI-systemet?

Kryss av for alt som passer per prosjekt:

Prosjekt 1:

Prosjekt 2:

Prosjekt 3:

	Prosjekt 1	Prosjekt 2	Prosjekt 3
Redusere kostnadene eller øke inntjeningen i virksomheten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Raskere saksbehandling eller bearbeidelse av henvendelser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Prosjekt 1	Prosjekt 2	Prosjekt 3
Frigjøre tid til å løse andre eller mer krevende oppgaver	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Redusere bemanningen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Forbedre kjerneoppgaver eller lage nye tjenester / tilbud	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Uttesting av KI eller intern kompetansebygging	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bedre informasjon som grunnlag for beslutninger	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Øke sikkerheten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bedret etterlevelse av forskrifter og regelverk etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

23) Eventuelle andre formål, eller kommentarer til denne delen av undersøkelsen

Vennligst spesifiser for hvilket prosjekt ev. formål og nytteeffektene gjelder.

24) * Hvem bruker eller skal bruke KI-systemets utfall/resultat?

Kryss av for alt som passer.

	Utviklere av systemet	Andre interne brukere	Eksterne brukere	Annet
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

25) * I hvilken grad kjører KI-systemet autonomt?

For systemer under utvikling, ta planlagt bruk.

	Fullautomatisert	Delvis automatisert, krevende saker håndteres manuelt	Delvis automatisert, alltid manuell verifikasjon	Ikke automatisert	Annet
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

		Delvis automatisert, krevende saker håndteres manuelt	Delvis automatisert, alltid manuell verifikasjon	Ikke automatisert	Annet
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Om bruk av en KI/ML-plattform

Du svarte at prosjektet under bruker en ML/KI-plattform. Vennligst presiser.

26) * Hvor kjører plattformen?

	Lokalt hos oss	I sky	Annet
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

27) Hvilke tjenester brukte eller bruker dere i plattformen under utvikling av KI-systemet?

Kryss av for alt dere bruker eller brukte i KI-systemet:

	Sammenligning av ulike typer modeller	Viktighet av inndata (feature importance)	Analyse av skjevheter i inndata	Analyse av skjevheter i modellens utfall
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

28) Hvilke endringer eller tilpasninger gjennomfører dere til ferdig trent modeller?

Du svarte at et KI-system bruker en ferdig trent modell som ble utviklet av samarbeidspartnere eller leverandør:

Prosjekt 1:

Prosjekt 2:

Prosjekt 3:

	Prosjekt 1	Prosjekt 2	Prosjekt 3
Validering på egne data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Prosjekt 1	Prosjekt 2	Prosjekt 3
Endring av modell basert på egne data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fastsetting av grenseverdier eller lignende uten endring av modell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

29) Eventuelle kommentarer til denne delen av undersøkelsen

Ressursbruk til utvikling av KI-systemet

Vennligst estimer ressursbruk til **utvikling** av KI-systemet, dvs. fra oppstart av teknisk utvikling (for eksempel innhenting av data) til første bruk i produksjon. Planleggingsfasen skal ikke inkluderes her.

Inkluder tilpasning og/eller validering av **hyllevare**, eller av KI-systemer utviklet av samarbeidspartnere.

31) * Kan du gi et anslag på hvor mye det har kostet å utvikle KI-systemet?

Inkluder arbeidstimer, innkjøp av hyllevare og timebruk til eventuelle konsulenter. Gi et grovt estimat i den mest passende priskategorien under.

For KI-systemer under utvikling, ta utgangspunkt i planlagt ferdigstilling av KI-systemet.

	Under 1,0 mill. kroner	1,0-2,5 mill. kroner	2,5-4,9 mill. kroner	5 mill. kroner eller mer	Vet ikke
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

32) * Hvor mange av egne ansatte er eller har vært involvert i utviklingen av KI-systemet?

Ekskluder konsulenter. Inkluder tilpasning av hyllevare.

	1 - 2	3 - 5	6 - 10	Flere enn 10	Ikke relevant
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

33) * Hvor lang tid tok det fra utviklingen startet til første bruk av KI-systemet?

For systemer under utvikling, bruk planlagt dato for første bruk. For modeller eller KI-systemer utviklet av leverandør eller andre partnere, angi tid for tilpasning, validering osv.

	Mindre enn 6 måneder	6 - 12 måneder	13 - 24 måneder	Mer enn 2 år	Ikke relevant
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

34) * Kan enkeltpersoner bli direkte berørt av resultatet fra KI-systemet?

Konsekvens for enkeltpersoner vurderes som **substansiell** dersom KI-systemets utfall kan føre til økonomiske, helsemessige eller rettslige konsekvenser, eller til tiltak som berørte personer kan oppfatte som inngripende.

	Ja, substansiell	Ja, men ikke substansiell	Nei	Usikker:
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

35) Eventuelle kommentarer til denne delen av undersøkelsen

Behandling av personopplysninger

Du svarte at KI-systemet under behandler personopplysninger, eller behandlet persondata under utviklingen.

36) * Behandler dere særlige kategorier av persondata etter GDPR artikkel 9 (sensitive opplysninger som for eksempel religionstilhørighet, politisk oppfatning og helseopplysninger) i forbindelse med KI-systemet?

- | | Ja | Nei |
|----|-----------------------|-----------------------|
| 1. | <input type="radio"/> | <input type="radio"/> |
| 2. | <input type="radio"/> | <input type="radio"/> |
| 3. | <input type="radio"/> | <input type="radio"/> |

37) Bruk av persondata i utvikling og bruk av KI-systemet

Kryss av for alt som passer per prosjekt:

Prosjekt 1:

Prosjekt 2:

Prosjekt 3:

	Prosjekt 1	Prosjekt 2	Prosjekt 3
Det var utfordrende å finne behandlingsgrunnlag for å kunne benytte persondata i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Det var vanskelig å få en avklaring om vi kunne bruke persondata i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi diskuterte internt bruken av persondata i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I organisasjonsenheten har vi tilstrekkelig kompetanse til å vurdere bruken av persondata	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi anonymiserer persondata før bruk i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi benytter pseudonymiserte persondata i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi har evaluert konsekvenser ved bruk av persondata i KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi har evaluert nytteverdi av hver enkelt personopplysning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

38) Eventuelle kommentarer til denne delen av undersøkelsen

Du svarte at effekten av KI-systemet kan berøre privatpersoner:

39) * Har dere testet for skjevheter mellom ulike persongrupper i KI-systemets utfall?

	Vet		
	Ja	Nei	ikke
1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Du svarte at for et eller flere KI-systemer kan effekten berøre privatpersoner, og at dere ikke har testet for skjevheter i KI-systemets utfall:

40) * Var det noe spesielt som hindret at dere testet for skjevheter i KI-systemets utfall?

- Nei
- Vi har ikke tilstrekkelige data
- Vi har ikke lov til å bruke nødvendige personopplysninger for test av skjevheter
- Annet

Hvordan dere jobber med KI

Resten av spørreskjemaet handler mer generelt om måten du og din organisasjonsenhet jobber på i utviklingen og bruk av KI. Kryss "ja" på alt som passer til minst ett KI-system dere har under utvikling, i bruk nå eller har hatt i bruk de tre foregående årene.

41) * Hvilke generelle utsagn treffer for deres KI-systemer?

	Ikke		
	Ja	Nei	relevant
KI-systemene våre testes for motstandsdyktighet mot manipulasjon og overlagt feilbruk	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KI-systemene våre er testet for svakheter mot cyberangrep eller datainnbrudd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Ja	Nei	Ikke relevant
KI-systemene våre har mekanismer for å håndtere feil eller uventet input	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vi har tilgjengelig dokumentasjon av endringer i KI-systemet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vi tester KI-systemene våre kontinuerlig for å sikre at systemene er stabile over tid	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KI-systemet gir begrunnelser for sine beslutninger eller resultater som er tilpasset brukerens kunnskapsnivå	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informasjon om bruk av KI ligger åpent tilgjengelig for eksterne brukere eller berørte personer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informasjon om hvordan KI-systemet fungerer og treffer beslutning eller gir resultater ligger åpent tilgjengelig for eksterne brukere eller berørte personer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Det er etablert mekanismer for at kunder/brukere kan klage på beslutninger gjort av KI-systemet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

42) * Benyttet din enhet skriftlige retningslinjer eller policyer på følgende områder ved utvikling av KI-systemer?

	Ja	Nei	Ikke relevant
Databehandling (data management)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dokumentasjon av KI-systemer (f. eks. datagrunnlag, modellparameter, valg av modelltype, KPIs, suksessfaktorer)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Risikovurdering av KI-systemer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Plan for innføring, overvåking og vedlikehold av KI-anvendelser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Innebygd personvern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Test for skjevheter i utfall av KI-anvendelser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

43) Hva er konsekvensene av de utfordringene dere har opplevd i utvikling eller bruk av KI-systemer?

Vennligst kryss av for de påstandene som stemmer overens med utfordringer dere har opplevd ved utvikling eller bruk av KI i din enhet.

	Konsekvens					
	Ikke opplevd	Økt kostnad	Forsinket fremdrift	Forringet funksjonalitet	Prosjekt stanset	Annet
Det er manglende forankring internt for utvikling eller bruk av KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Det er motstand i organisasjonen for å endre eksisterende prosesser og/eller arbeidsstrukturer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ledelse eller brukere har urealistiske forventninger til mulighetene med KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi har støtt på utfordringer som burde løses på et høyere nivå i virksomheten/sektor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi er avhengig av kompetanse fra innleide konsulenter eller leverandører for utvikling av KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi mangler relevant ikke-teknisk kompetanse for å utvikle og bruke KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gevinster med KI kunne ikke synliggjøres eller måles godt nok	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi nådde ikke forventet resultat med KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi fikk ikke tilstrekkelig finansiering for KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Det er vanskelig å rekruttere og/eller beholde ansatte med KI-kompetanse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Resultat fra utvikling/proof-of-concept kunne ikke uten videre brukes i produksjon	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vi har ikke nok data tilgjengelig, eller ikke god nok datakvalitet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vår digitale infrastruktur/IT-løsninger er ikke tilpasset KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sentrale forhold rundt KI er uavklarte (for eksempel med sikkerhet, personvern, transparens, forskjellsbehandling, etiske og/eller juridiske vurderinger)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Det er krevende at vi ikke alltid kan forklare hvordan KI-systemet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Konsekvens					
	Ikke opplevd	Økt kostnad	Forsinket fremdrift	Forringet funksjonalitet	Prosjekt stanset	Annet
treffer beslutninger eller generer resultater						
Kontroll av sikkerhetsrisiko knyttet til KI-systemet var krevende	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Forebygging av misbruk av KI-systemet var krevende	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kontroll av <i>skjevheter</i> i KI-systemets utfall var krevende	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Det var krevende å ivareta personvernet i utviklingen eller bruk av KI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Du svarte at din organisasjonsenhet benytter retningslinjer eller policyer for dokumentasjon av KI-systemer:

44) * Hvilke av følgende aspekter er inkludert i deres standard for dokumentasjon av KI-systemer?

	Ja	Nei	Delvis
Datakilde: Populasjon, kontekst og formål ved innsamling, kvalitet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Prosessering og utvalg av data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Valg av modelltype og modellparameter, med begrunnelse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Valg av funksjon- og ytelsenivå med grenseverdier, med begrunnelse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Krav til og implementering av forklaringer av KI-systemets utfall	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Plan for oppdateringer og vedlikehold	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Forebygging av feil bruk, misbruk eller angrep	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ansvar ved feil eller uforutsett konsekvens	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Avslutningsvis:

45) Har du noen kommentarer, tilleggsinformasjon eller lignende til undersøkelsen?

10.3 Vedlegg 3: Rammeverk for vurdering av ansvarlig KI

For vurderinger av hvorvidt prinsippene for ansvarlig KI er ivaretatt i KI-systemene i caseundersøkelser, bruker vi et rammeverk med revisjonsspørsmål basert på [Auditing machine learning algorithms](#)²⁹⁴ og [Audit framework for algorithms](#)²⁹⁵.

Hvert underspørsmål henger sammen med ett av de etiske prinsippene for ansvarlig KI som hovedprinsipp, eller som *styring og ansvar* som relaterer til flere etiske prinsipper.

10.3.1 Underspørsmål vurdert i caseundersøkelser

Styring og ansvar	
Underkategori	Revisjonsspørsmål
Prosesstyring	Dokumentasjonskvalitet: Er alle forhold ved utviklingen, implementeringen/innføringen og bruken av KI-systemet godt nok dokumentert?
Prosesstyring	Har algoritmen et klart definert formål?
Prosesstyring	Har hele prosessen (livssyklusen) for algoritmen blitt dokumentert?
Prosesstyring	Finnes det vedtatte og dokumenterte retningslinjer som stiller krav til kvalitet og resultatmål for algoritmen?
Prosesstyring	Blir algoritmen overvåket jevnlig (i det minste i forbindelse med tilgjengelighet, resultater/kvalitet, sikkerhet, samsvar med lovverk og tjenesteutsetting)?
Ressursstyring	Har organisasjonen god nok tilgang på mye nok og høy nok kompetanse?
Ressursstyring	Har rollene, oppgavene, ansvarsområdene og myndighetene (inkludert eierskap) blitt definert, og har disse blitt satt ut i livet?
Risikostyring	Blir risikoene ved bruk av algoritmen evaluert ved fastsatte, regelmessige intervaller? Hvilke av de følgende forholdene har blitt undersøkt i en risikovurdering, og hva var resultatet? - programvareapplikasjonen - IT-miljøet rundt - tilretteleggingen av data - prosessene som applikasjonen inngår i
Eksterne bidrag	Når deler eller aktiviteter i forbindelse med algoritmen blir tjenesteutsatt, er det lagt til rette for produkt- og kunnskapsoverføring med de aktuelle eksterne partene, og er dette dokumentert?
Eksterne bidrag	Når KI-systemer er kjøpt inn, er det gjort avtale med leverandøren om hyppige oppdateringer og betimelig informasjon fra leverandøren om sikkerhets- eller resultatproblemer, og er denne avtalen dokumentert?
Informasjons- og datarammeverk	Er datakildene, de tekniske/operasjonelle kriteriene for datautvelgelse og det rettslige grunnlaget for innsamling og behandling tilstrekkelig dokumentert?
Utvikling og bruk	Ble de ulike interessentene og sluttbrukerne for algoritmen involvert i utviklingsprosessen?

²⁹⁴ Et rammeverk for revisjon av KI-systemer basert på maskinlæring, utviklet av Riksrevisjoner fra Finland, Tyskland, Nederland, UK og Norge i 2020 (oppdatert 2023), med tilhørende [Excel verktøy](#).

²⁹⁵ Et rammeverk utviklet av den Nederlandske Riksrevisjonen (Netherlands Court and Audit) i forbindelse med revisjonen [Understanding algorithms](#) i 2021, og anvendt i revisjonen [An audit of 9 algorithms used by the Dutch government](#) i 2022²⁹⁵. Dette rammeverket legger ikke maskinlæring til grunn, men også enklere typer KI-systemer.

Personvern	
Innebygd personvern	Er KI-systemet utformet for å oppfylle prinsippet om innebygd personvern?
Datagrunnlag	Er det rettslige grunnlaget for å behandle personopplysninger dokumentert? Skjer databehandlingen på grunnlag av lovfestet plikt?
Datagrunnlag	Er behandlingen av (særskilte) personopplysninger med algoritmen i tråd med det opprinnelige formålet?
Informasjons- og datarammeverk	Er perioden for datalagring definert, og finnes det en policy for sletting?
Informasjons- og datarammeverk	Er det utpekt en behandlingsansvarlig og databehandler for algoritmen og dataene som brukes?
Databehandling	Er det tegn på dataminimering? Er det tatt hensyn til at behandling av personopplysninger skal være rimelig, relevant og begrenset (forholdsmessighet)?
Databehandling	Blir behandlingen av personopplysninger dokumentert i et oppdatert digitalt register?
Databehandling	Er det utført en vurdering av personvernkonsekvenser (DPIA) (hvis relevant)?
Databehandling	Er det lagt til rette for å informere, enten proaktivt eller ved forespørsel, personer som organisasjonen har behandlet eller brukt data om (når det gjelder både data og algoritmen)?
Databehandling	Finnes det en offentlig tilgjengelig personvernerklæring som dekker dataene og algoritmene som er brukt?
Utvikling og bruk	Er det tegn på at det tas automatiserte beslutninger, og er det i så fall i samsvar med lovverket, og har de involverte mulighet til å kreve menneskelig inngripen og klage på beslutningen?
Utvikling og bruk	Er det gjort en vurdering av om det finnes tegn på profilering av mennesker, og om dette er tillatt?
Likebehandling	
Prosesstyring	Hvordan sikres at modellen ikke forskjellsbehandler usaklig eller ulovlig? Er KI-systemet utformet for å oppfylle prinsippet om innebygd rettferdighet?
Datagrunnlag	Er det ingen uønsket skjevhet i dataene?
Datagrunnlag	Er dataene brukt i utviklingen representative for hvordan algoritmen brukes?
Eksterne bidrag	Er datakildene til innkjøpte KI-systemer eller tjenesteutsatt utviklingsarbeid tilstrekkelig dokumenterte for å vurdere deres egnethet for organisasjonens bruksområde?
Utvikling og bruk	Er det iverksatt tiltak for å unngå skjevheter som følge av valgene som er gjort for modellen?
Utvikling og bruk	Forekommer det diskriminering på grunn av dataene og modellen som brukes?
Transparens	
Prosesstyring	Hvordan sikrer modellen transparens? Er KI-systemet utformet for å oppfylle prinsippet om innebygd transparens?
Databehandling	Er logikken i algoritmen, konsekvensene den har, og dataene som brukes, tydelige nok for de registrerte?
Modellstyring	Har algoritmen et delt formål, og er det tydelig og forklarbart for eieren, utvikleren og brukeren?

Utvikling og bruk	Har interessenter tilgang til informasjon om bruken av KI, en beskrivelse av de underliggende dataene (så langt det er mulig) og metoden for å analysere dataene? Har bruken av modellen eller algoritmen, inkludert begrensningene (altså hva den kan og ikke kan gjøre), blitt formidlet til alle interessenter?
Utvikling og bruk	Er algoritmen forklarbar, og er det gjort en innsats for å avveie modellens forklarbarhet mot resultatet?
Utvikling og bruk	Finnes det dokumentasjon som beskriver utformingen og implementeringen/innføringen av algoritmen?
Utvikling og bruk	Er begrunnelsene for valgene som er tatt i utformingen og implementeringen/innføringen av algoritmen, dokumentert?
Utvikling og bruk	Har kvaliteten til modellen blir dokumentert, inkludert resultatmål og andre indikatorer?
Robusthet og sikkerhet	
Prosesstyring	Har målene blir oppfylt, og har bruken oppnådd den ønskede hensikten?
Risikostyring	Hvordan påvirkes KI-systemet av feilaktige eller manipulerede datasett?
Generelle IT-kontroller – sikkerhetskopi	Hva slags løsninger for sikkerhetskopier finnes for KI-systemet (det vil si sikkerhetskopiering, hyppighet, lagringsplass og testrutiner)?
Generelle IT-kontroller – loggføring	Hvilke rutiner for loggføring følges for algoritmen og dens programkjøringsmiljø (inkludert databaser), og hvordan blir loggene overvåket?
Generelle IT-kontroller – tilgangsstyring	Hvordan blir brukere med spesialrettigheter administrert i driftsmiljøet i KI-systemet? (Beskriv minst opprettelse, endring, deaktivering, regelmessig gjennomgang, oversikt/register over autorisasjoner med personen ansvarlig for godkjenning, har administratorer normale brukerkontoer til ikke-administratorbruk, i hvilken grad brukes upersonlige kontoer?)
Generelle IT-kontroller – tilgangsstyring	Hva slags brukere har tilgang til KI-systemet (person/bruk, rettighetsnivå), og hvilken autentiseringsmetode brukes? Hvis flerfaktoraутentisering brukes, hva slags og i hvilken grad med tanke på systemer, intern/ekstern tilgang og brukergrupper? Hva er kravene til passord, hvordan er disse kravene implementert, og er det ulike krav til ulike brukergrupper? Hvis det brukes krypteringsnøkler, er de knyttet til identifiserbare eiere, og hvor er nøklene lagret?
Generelle IT-kontroller – endringshåndtering	Hva slags endringshåndtering brukes for KI-systemet, inkludert evaluering av konsekvenser og autorisasjon? Hvilke krav er det til endring av kode eller algoritme/modell?
Interne IT-kontroller – eksterne bidrag	For anskaffede KI-systemer: Har avtaler blitt inngått med leverandørene for tilstrekkelig loggføring, sikkerhetskopiering, endringshåndtering og tilgangskontroll, i den grad dette skal håndteres av leverandøren? Er informasjon som systemet sender mellom kunden og leverandøren, sikkert nok?
Eksterne bidrag	For anskaffede KI-systemer: Er det avtale med leverandøren om kontinuerlig støtte av systemet med rimelig forutsatt levetid, eller er det utforsket alternativer i tilfelle støtten fra leverandøren opphører?

Styring	Har algoritmen et formål, og har det blitt oversatt til praktiske funksjoner med tanke på modellen og dataene som brukes? Hvilke konkrete oppgaver eller sider av operasjonsledelsen er algoritmen ment å støtte?
Styring	Er det lagt til rette for vedlikehold og håndtering av algoritmen? Blir modellen regelmessig oppdatert, inkludert for å samsvare med gjeldende lovgivning?
Utvikling og bruk	Hvilke trenings-/validerings-/test-datasett ble brukt? Hvordan ble disse laget eller valgt ut, og hvordan oppdateres disse gjennom systemets levetid? Er det iverksatt tiltak rettet mot kvaliteten til valgene som tas om trenings- og testdata? Blir trenings-, test- og valideringsdata behandlet separat? Er mengden data tilstrekkelig til å separere dataene på denne måten?
Utvikling og bruk	Var valget av hyperparametere støttet av argumenter og dokumentasjon?
Utvikling og bruk	Ble modellens ytelse sammenlignet med alternative metoder/modeller? Utdyp.
Utvikling og bruk	Blir modellens resultater overvåket?
Eksterne bidrag	For anskaffede KI-systemer: Er det tilgjengelig dokumentasjon fra leverandøren om kvaliteten til modellen, inkludert ytelsesmål eller andre resultatindikatorer? Hvilke ytelsesmål ble gitt av leverandøren?

Noen av de mer omfattende revisjonsspørsmålene er operasjonalisert i neste avsnitt.

10.3.2 Operasjonalisering av rammeverket

Hvert revisjonsspørsmål i rammeverket blir vurdert som et av tre nivåer, lignende systemet for prosessers evne- eller modenhetsnivå i COBIT 2019²⁹⁶, men forenklet til bare tre modenhetsnivåer: lav, middels og høy. Siden ikke alle kravene i rammeverket er like relevante for alle KI-systemer, kan en vurdering på lav eller middels være tilstrekkelig for KI-systemet. Derfor bruker vi begrepet aktivitetsnivå i beskrivelse av resultater.

På grunnlag av disse enkeltvurderinger beregnes et oppsummert aktivitetsnivå per etisk prinsipp som prosent av kravene som er oppfylt i kategorien, med vektning 0, 0,5 eller 1 for henholdsvis lavt, middels eller høyt nivå i hvert krav.

10.3.3 Styringsprinsipper vurdert i rammeverket

Noen styringsprinsipper som er vurdert i rammeverket, er omfattende og blir derfor delt opp i underpunkter her.

KI-livssyklusstyring (AI lifecycle management)

KI-livssyklusstyring er et ledelsessystem for styring av hele prosessen for KI-utvikling, fra planlegging til drift eller avvikling. En plan for slik prosessstyring inneholder tids- og ressursplanlegging av arbeidet med KI-systemets formål og krav, design og utvikling, kvalitetssikring og testing, produksjonssetting og vedlikehold. For leverandørsystemer er design og utvikling erstattet av anskaffelse og eventuell tilpasning. Sammenlignet med tradisjonell programvare er risikoen med KI er større i test- og vedlikeholdsfasen, siden for eksempel maskinlæringsmodeller er kjent å være mottakelige for uønsket endring av prestasjon over tid. Underpunktene er derfor

²⁹⁶ Capability Maturity Model Integration® (CMMI)-based process-capability scheme, fra 0 til 5, COBIT 2019 Framework: Governance and management objectives, ISACA

1. tids- og ressursplan i utviklingsfasen, inkludert definert mål
2. tids- og ressursplan for overvåking av KI-systemets resultater
3. tids- og ressursplan for vedlikehold og oppdatering, inkluderer re-trening av modeller

I sammenheng med overvåking og vedlikehold betyr ressursplanen at ansvaret er tydelig definert. Merk at det som skal overvåkes i denne omgangen av vurderingen bare er KI-systemets ytelse. Ideelt sett burde KI-livssyklusstyring kobles til innebygd personvern, likebehandling og transparens, og tilsvarende oppførsel av KI-systemet bli inkludert i utvikling, overvåking og vedlikehold. For å kunne vurdere disse aspektene separat har vi valgt å redusere livssyklusstyring til ytelse og vurderer innebygde etiske prinsipper separat (se neste avsnitt).

Innebygde etiske prinsipper

OECD anbefaler²⁹⁷ å iverksette mekanismer og sikkerhetstiltak for å ivareta etiske prinsipper gjennom hele KI-systemets livssyklus. Prinsippet innebygd ansvarlig KI, eller ansvarlig KI «by design», betyr å ta hensyn til personvern, rettferdighet, transparens/åpenhet og sikkerhet i alle KI-systemets livsfaser.

Datatilsynet definerer for eksempel innebygd personvern som «prinsipp om at et teknisk system eller en løsning blir utviklet slik at personvernet blir ivaretatt», og peker på at en forutsetning er at å ha tilstrekkelig kompetanse på de underliggende personvernprinsippene²⁹⁸. På lignende måte beskriver Digdir innebygd informasjonssikkerhet som oppnås blant annet med «innarbeidd i prosessane og prosjekta i verksemda frå starten av» og «tatt omsyn til i heile livssyklusen til IKT-løysingar».²⁹⁹

I rammeverket for KI-revisjon ser vi etter indikasjoner på innebygd personvern, rettferdighet og transparens som styringsprinsipper. Det inkluderer minst en evaluering av relevante prinsipper og tilknyttet risiko samt planlegging og implementering av eventuelle tiltak.

Vi ser ikke på innebygd sikkerhet, siden sikkerhetsaspekter som er uavhengig av KI, er håndtert av virksomhetens generelle sikkerhetsstyring, og ingen av de KI-systemene i caseundersøkelsen er utsatt for nye, KI-spesifikke sikkerhetsrisiko som *poisoning* eller *adversarial attacks*. I tillegg faller ingen av de tre selvutviklede KI-systemene innenfor klassifiseringen av høyrisiko³⁰⁰ KI-systemer, i henhold til risikoklassifiseringen i forslag til KI-forordningen (*AI Act*), og det er ingen fare for naturlige personers liv og helse hvis det oppstår feil på disse KI-systemene. Et unntak er KI-systemet i helsesektor, som faller under klassifisering av høyrisiko fordi det brukes i medisinsk diagnostikk, men her er KI-spesifikk innebygd sikkerhet tett knyttet til overvåking av systemets ytelse, som dekkes under KI-livssyklusstyring.

Innebygd personvern

For å oppfylle kravet om innebygd personvern etter personvernforordningen artikkel 25 kan virksomhetene orientere seg på personvernprinsippene i [Datatilsynets veiledning på innebygd personvern og personvern som standard](#), som lister opp nøkkelelementer for personvernprinsippene med eksempler for implementering. Personvernprinsippene som virksomheter bør vurdere i kontekst av KI-systemet, er lovlighet, rettferdighet, åpenhet, formålsbegrensning, dataminimering, riktighet, lagringsbegrensning, integritet og konfidensialitet, og ansvarlighet.

Personvernprinsippene rettferdighet og åpenhet overlapper dermed med styringsprinsippene om innebygd rettferdighet og transparens. For å minimere overlappet i vår vurdering er krav for å oppfylle rettferdighet eller åpenhet som en av ni personvernprinsipper lavere enn det vi vurderer som innebygd prinsipp. Ikke minst er de innebygde prinsippene ikke begrenset på personvernforordningen som

²⁹⁷ OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449

²⁹⁸ Datatilsynet. (2023, 27. juli). *Innebygd personvern og personvern som standard*. <https://www.datatilsynet.no/rettigheter-og-plikter/virksomhetenes-plikter/innebygd-personvern-og-personvern-som-standard/>

²⁹⁹ Digdir. (u.å.). Innebygd informasjonssikkerheit. Hentet 23.05.2024 fra <https://www.digdir.no/informasjonssikkerhet/innebygd-informasjonssikkerheit/2146>

³⁰⁰ Høyrisiko KI etter kategorisering av EUs forslag til KI-forordningen, artikkel 6 og Annex III.

kontekst, men blir også vurdert i henhold til de lovfestede og ulovfestede prinsippene for saksbehandling i offentlig forvaltning.

Innebygd rettferdighet

Rettferdighet har mange definisjoner, og forståelsen avhenger av kontekst. En mulig praktisk metode for å undersøke rettferdighet er å sammenligne om like saker behandles likt. For individuell rettferdighet er det ofte vanskelig å bestemme «likt» for både utgangssituasjon, behandling og resultat. Derfor brukes ofte gruppebasert rettferdighet, der man sammenligner behandlingen eller resultat av ulike grupper mennesker som er like hverandre og derfor burde behandles likt.

Likestillings- og diskrimineringsloven definerer menneskelige egenskaper, såkalte diskrimineringsgrunnlag, som ikke er lov å bruke som grunnlag for ulik behandling. I offentlig forvaltning skal vedtak ikke innebære usaklig forskjellsbehandling, uten at *usaklig* er begrenset til diskrimineringsgrunnlag etter likestillings- og diskrimineringsloven.

Grunnen til at KI-systemer kan fungere ulikt for ulike persongrupper kan for eksempel være historiske forskjeller som modeller lærer av dataene. En vanlig utfordring ved utvikling av maskinlæringsystem er videre at noen grupper er underrepresentert i treningsdataene, noe som fører til at modellen gir dårligere prediksjoner for disse gruppene.³⁰¹ Denne formen for skjevhet omtales som representasjonsskjevhet.³⁰²

Standard definisjoner av rettferdighet

Metodikken for innebygd diskrimineringsvern i Likestillings- og diskrimineringsombudets (LDO) veileder³⁰³ bygger på **gruppebasert rettferdighet**. Denne måten å definere rettferdighet på er lett å relatere til lovverket, og samtidig er det lett å beregne rettferdighetsmål³⁰⁴. Derfor legger vi også denne definisjonen til grunn. Typiske rettferdighetsmål i en klassifisering baserer seg på antall feilklassifiseringer i ulike grupper, for eksempel hvor mange kvinner blir feilaktig kontrollert sammenlignet med menn.

Et alternativ som allerede er nevnt, er **individuell rettferdighet**, som betyr at like saker behandles likt. Definisjonen er mindre brukt, fordi det er utfordrende å måle likhet av inndata og likhet av resultat.

Kontrafaktisk rettferdighet (*counterfactual fairness*) er noe mer brukt. Det måler hvorvidt en egenskap eller variabel er årsaken til resultatet, ved å gjøre samme beregning også med omvendt egenskap, for eksempel teste utfallet av en maskinlæringsmodell for samme person når bare egenskapen kjønn er endret i inndata. Denne metoden egner seg godt til å utforske mekanismer bak ulik behandling, men det er ofte vanskelig å behandle korrelerte egenskaper riktig (for eksempel er en variabel som yrkesfag relatert til kjønn). Begrepet **kontrafaktisk forklaring** betegner en relatert metodikk som går ut på å forklare grunnlaget til maskinlæringsmodellens utfall med den minste endringen i inndata som fører til endret utfall. Denne metoden kan også egne seg for å avdekke om usaklige grunn påvirker en prediksjon.

Disse tre tilnærmingene til å vurdere rettferdighet kan anses som standard, med de tekniske utfordringene ved individuell og kontrafaktisk rettferdighet som beskrevet. En fjerde tilnærming som brukes iblant, kalles **rettferdighet ved uvitenhet** (*fairness through unawareness*), og betyr at et KI-system vurderes som rettferdig hvis ingen av egenskapene som inngår, kan være diskrimineringsgrunnlag. Denne definisjonen anser vi og forskningsmiljøet at **ikke** er tilstrekkelig, siden

³⁰¹ Likestillings- og diskrimineringsombudet. (2023). *Innebygd diskrimineringsvern: En veileder for å avdekke og forebygge diskriminering i utvikling og bruk av kunstig intelligens*. https://ldo.no/globalassets/ldo_2019/bilder-til-nye-nettsider/ki/ldo-innebygd-diskrimineringsvern.pdf

³⁰² Storås, A. M., Prabhu, R., Hammer, H. L. & Strømke, I. (2022). Bias og kvantitativ analyse innen velferd: opphav til skjevheter og relasjon til utfallsrettferdighet. *Tidsskrift for velferdsforskning*, 25(3). <https://www.idunn.no/doi/10.18261/tfv.25.3.3>

³⁰³ Likestillings- og diskrimineringsombudet. (2023). *Innebygd diskrimineringsvern: En veileder for å avdekke og forebygge diskriminering i utvikling og bruk av kunstig intelligens*. https://ldo.no/globalassets/ldo_2019/bilder-til-nye-nettsider/ki/ldo-innebygd-diskrimineringsvern.pdf

³⁰⁴ En utfordring er fortsatt å begrunne hvilket rettferdighetsmål som er mest relevant å oppfylle, siden det ofte er umulig å oppfylle alle, jf. *impossibility theorem of fairness*.

det i mange tilfeller ble vist at korrelasjoner med variabler som inngår, kan føre til diskriminering. I forbindelse med testen av KI-systemet viser LDOs veileder for innebygd diskrimineringsvern derfor til at «selv om personopplysninger fjernes, kan systemet finne mønstre i dataene slik at opplysninger som sammenfaller med diskrimineringsgrunnlagene vektlegges ulovlig.»

Likebehandling i tilsyn med virksomheter

Rettferdighetsdefinisjoner baserer seg oftest på rettferdighet i behandling av naturlige personer. Kravene til saklighet, forsvarlighet og proporsjonalitet i offentlig forvaltning er imidlertid ikke begrenset til naturlige personer, men gjelder hele saksbehandling³⁰⁵. Lignende definisjoner og metoder kan derfor brukes til å vurdere likebehandling av virksomheter. Hovedforskjellen er hvilke egenskaper ved virksomheter som kan føre til usaklig forskjellsbehandling.

Den mulige risikoen for usaklig forskjellsbehandling av virksomheter må vurderes for hvert KI-system.

Innebygd transparens

Transparens i KI består av åpenhet og forklarbarhet. Åpenhet handler om å gi relevant informasjon til interessenter (*stakeholders*), mens KI-systemers forklarbarhet handler om å forstå hvordan utgangssituasjonen (inndata) fører til et visst resultat (utfall). Forklarbarhet kan dermed være en forutsetning for åpenhet der en slik forståelse er relevant. For eksempel skal enkeltvedtak grunngis etter forvaltningsloven § 24 (åpenhet), og § 25 presiserer at begrunnelsen skal inneholde både de relevante reglene og de faktiske forholdene vedtaket bygger på (forklaring).

I vurderingen av innebygd transparens ser vi på planlegging for og implementering av åpenhet og forklarbarhet. I tillegg ser vi på om det finnes en indikasjon på at transparens er et mål.

Ved KI-utvikling må det blant annet vurderes hvordan hensynet til nøyaktige og høyt ytende modeller skal balanseres med muligheten til å forstå og forklare beslutningene tatt av KI-systemet.³⁰⁶

Robusthet og sikkerhet

Sikkerhet i forbindelse med kunstig intelligens omfatter blant annet informasjonssikkerhet, menneskers sikkerhet og trygg bruk av KI. Informasjonssikkerhet handler om hvordan informasjonen blir beskyttet mot uønsket innsyn (konfidensialitet), er tilgjengelig ved behov (tilgjengelighet) og er beskyttet mot manipulering (integritet). For at KI-systemer skal være forebygget mot skade, må de være teknisk sikre og robuste ved å blant annet sikre mot manipulasjon og misbruk. Kunstig intelligens skal være bygget på systemer med teknisk robuste løsninger som forebygger risiko og bidrar til at systemene fungerer slik de er tiltenkt.³⁰⁷

³⁰⁵ Jf. ulovfestede prinsipper i norsk forvaltningsrett, [NOU 2019: 5 \(2019\). Justis- og beredskapsdepartementet](#). Kapittel 11.7.

³⁰⁶ Raz, A., Heinrichs, B., Avnoon, N., Eyal, G. & Inbar, E. (2024). Prediction and explainability in AI: Striking a new balance? *Big Data & Society*, 11(1) <https://doi.org/10.1177/20539517241235871>.

10.4 Vedlegg 4: Bruksområder for KI i statlige virksomheter og heleide selskaper

Liste over ulike bruksområder for KI i statlige virksomheter og heleide selskaper, ikke uttømmende: Eksempler fra kartleggingsundersøkelsen.

Hovedbruksområder KI	Beskrivelse	Eksempel
Kartlegging av tilstand og vedlikeholdsbehov	Bruk av sensordata og tilstandskartlegging for planlegging av drift og vedlikehold av utstyr og infrastruktur	Analysere dronebilder for å oppdage feil på komponenter
Prediktive modeller	Forutsi produksjon, bruk av tjenester, ressursbehov, sykdomsforløp og sykefravær	Predikere antall kundesenter for å justere bemanningsbehov på et kundesenter
Prisoptimalisering	Optimalisere prissetting, beregne grenseverdier og kostnader for ulike produkt- eller kundegrupper	Beregne støtte gjennom å estimere pris på alternative produkter. Det alternative produktet avviker fra opprinnelig produkt som følge av én bestemt egenskap.
Kundestøtte og innsikt	Anbefalingssystemer for kunder og produkter, veiledning og chatbots for kunder. Automatisert refusjon. Prediksjon og forebygging av kundeavgang.	Chatbot gir 24/7støtte for standardhenvendelser og avlaster kundesenteret
Risikomodellering og overvåkning	Vurdering av sannsynlighet for regelverksbrudd, skade eller risikofylt adferd til planlegging av tilsyn og kontroll, beregning av fare for ulykker eller naturkatastrofer (skred, flom osv.).	Finne statistiske avvik i overvåkingsdata som er forbundet med økt sannsynlighet for ulykker
Ressurs- og områdekartlegging	KI støttet analyse av kart og andre data for å estimere forekomster av naturressurser og kartlegge områder av interesse	Undersøke tilstand og utvikling av naturområder med analyse av satellittbilder
Bearbeiding og analyse av ulike datakilder	Automatisk transkribering, tekstanalyse og skriftgjenkjenning. Systematisk analyse av kvalitative kilder med bruk av språkmodeller	Automatisere teksting av bilder eller lydopptak.
Automatisering av arbeidsprosesser	Bruk av KI til å automatisere regelbaserte rutineoppgaver og saksbehandling, sammenstille eller kvalitetssikre data, eller kategorisere henvendelser.	Automatisere arkivering av dokumenter med robotisert prosessautomatisering (RPA)
Autonome roboter og systemer	Autonome systemer i landbruk, transport og logistikk	Selvkjørende robot automatiserer intern varetransport