Innspillsrapport

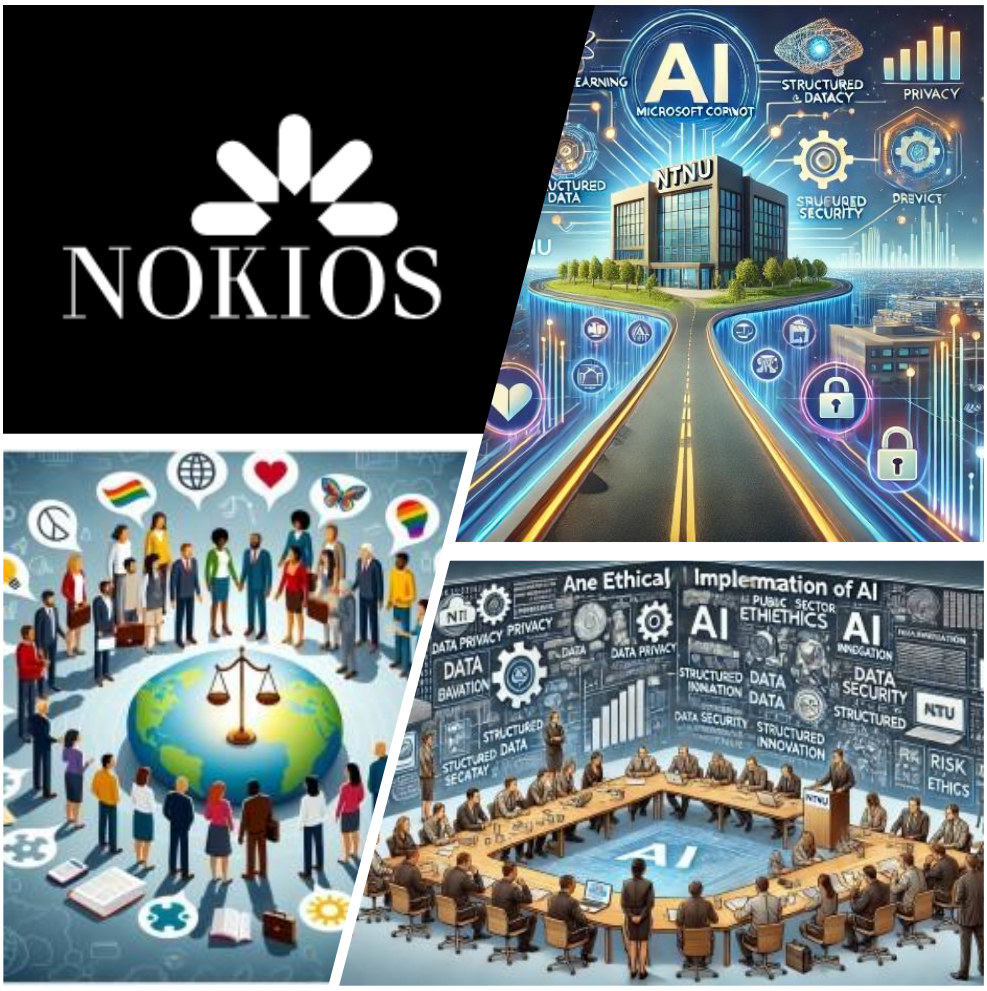
Heine Skipenes, Silje Reiten Blichfeldt,
Hanne Jensen Moe, Marte Nubdal

Et innspill til KI-reisen 2.0?

Innspillsrapport i NTNUs prosjekt i Datatilsynets regulatoriske sandkasse våren 2024

Trondheim, 25. oktober 2024

NTNU
Norges teknisk-naturvitenskapelige universitet
IT-avdelingen / Avdeling for utvikling og virksomhetsstyring



PROSJEKT
Pilotere Copilot for Microsoft 365 i Datatilsynets regulatoriske sandkasse
Innspillsrapport

Innspillsrapport

Et innspill til KI-reisen 2.0?

Datatilsynets regulatoriske sandkasse for personvernvennlig innovasjon og digitalisering våren 2024

VERSJON
Versjonsnummer 1.0

DATO
25. oktober 2024

FORFATTERE
Heine Skipenes, Silje Reiten Blichfeldt, Hanne Jensen Moe, Marte Nubdal

ANTALL SIDER
24 sider

OPPDRAGSGIVER
Datatilsynet

Hvordan ta gode valg når du skal ta i bruk KI-verktøy?

Rapporten bygger på NTNUs erfaringer med å teste Copilot for Microsoft 365 i Datatilsynets regulatoriske sandkasse for personvernvennlig innovasjon. NTNU presenterte sin funnrapport i juni 2024, som inkluderte deres "KI-reise". Datatilsynet og NTNU arrangerte en workshop på NOKIOS-konferansen 22. oktober 2024 hvor målet var å evaluere og forbedre denne KI-reisen, samt å gi innspill og tilbakemeldinger til Datatilsynet om hva det er viktig å fokusere på fremover. Rapporten fremhever flere kritiske temaer for implementering av kunstig intelligens (KI) i offentlig sektor, blant annet:

- **Økonomi og ressursbruk**: For å sikre en vellykket og bærekraftig implementering av KI, må kostnader balanseres mot nytteverdi. Det er viktig å utvikle klare business cases og sikre tilstrekkelige økonomiske ressurser.
- **Personvern og informasjonssikkerhet**: For å opprettholde tilliten til offentlig forvaltning, må KI-systemer håndtere personopplysninger sikkert og unngå diskriminering. Dette krever robuste kontrollsystemer og samsvar med gjeldende personvernlover.
- **Bærekraft**: Implementeringen av KI må være bærekraftig med hensyn til energiforbruk og miljøpåvirkning. Vi bør inkludere bærekraftsvurderinger i egne KI-strategier og ta hensyn til alle kostnader, inkludert eksterne kostnader.
- **Etikk og tillit**: KI må brukes på en måte som ivaretar etiske prinsipper og opprettholder tilliten til offentlige tjenester. Dette innebærer å implementere etiske retningslinjer som fremmer menneskelig verdighet, autonomi og rettferdighet.

Dokumentet er en oppsummering av arbeidet fra NOKIOS-workshopen, og er nesten utelukkende skrevet av kunstig intelligens på bakgrunn av et datamateriale utarbeidet av mennesker med ekte intelligens. Innholdet er kvalitetssikret av et menneske.

For mer informasjon om prosjektet, kan du besøke ntnu.no/copilot.

Dette dokumentet er en innspillsrapport fra en workshop i regi av Datatilsynet og NTNU på NOKIOS-konferansen 22. oktober 2024. Rapporten inngår i prosjektet «Pilotere Copilot for Microsoft 365» i Datatilsynets regulatoriske sandkasse for personvernvennlig innovasjon og digitalisering våren 2024. Rapporten beskriver NTNUs erfaringer med å teste Copilot for Microsoft 365.

Rapporten er ikke en vitenskapelig rapport, men et dokument som oppsummerer tilbakemeldinger fra en 4-timers workshop med ca 40 deltakere. Rapporten er skrevet fra NTNUs perspektiv, og representerer ikke synspunktene til Datatilsynet, Microsoft eller andre samarbeidspartnere. Rapporten er ikke en anbefaling eller en kritikk knyttet til et konkret verktøy, men en beskrivelse av våre observasjoner og refleksjoner i løpet av en begrenset prosjektperiode. Rapporten er ment som et innspill til andre organisasjoner når de gjør sine egne vurderinger. NTNU understreker viktigheten av at alle organisasjoner bør gjøre egne vurderinger, og NTNUs mål har vært å levere rapporter og annet informasjonsmateriell som kan hjelpe til med å komme så raskt som mulig inn i kjernen av muligheter og utfordringer som kunstig intelligente verktøy gir.

Som en del av NTNUs egen læringsprosess er rapporten så langt det er mulig skrevet ved hjelp av verktøyene Copilot with enterprise protection (https://copilot.microsoft.com/) og Copilot for Microsoft 365. Kvalitetssikring er alltid gjort av et menneske. NTNU anbefaler at du leser dokumentet med det i bakhodet, så du kan se hva som er mulig å få til med relativt enkle og tilgjengelige verktøy. Hvis du ønsker mer informasjon om hvordan dette er gjort i praksis, se eksempelbeskrivelse av fremgangsmåte («Steg for steg beskrivelse: Eksempel på bruk av Copilot») i funnrapporten lansert i juni 2024..

Kildehenvisninger og referanser er bakt inn i teksten. Fotnoter er brukt der dette er relevant.

Alle illustrasjoner er generert av kunstig intelligens (Copilot with enterprise protection, ChatGPT eller Copilot for Microsoft 365).

Alle samarbeidspartnere har dekt sine egne utgifter i forbindelse med prosjektet. Microsoft har gjort tilgjengelig intern utrullingsstøtte for programvareleverandøren Crayon sitt arbeid opp mot NTNU (Copilot readiness workshop). NTNU har godkjent Crayons arbeid hos Microsoft og dokumentert dette i NTNUs arkivsystem (sak 2023/37288)

Prosjektet har mottatt støtte fra EUs DIGITAL EUROPE program (nr 101083966)

This project has received funding from the European Union's DIGITAL EUROPE programme, under Grant Agreement nº 101083966.

Innholdsfortegnelse

Innledning

Kunstig intelligens (KI) har potensial til å revolusjonere offentlig sektor ved å effektivisere tjenester, forbedre beslutningsprosesser og skape nye muligheter for innovasjon. NTNU har utviklet «KI-reisen» som en veiledning for virksomheter som ønsker å ta i bruk KI på en trygg og effektiv måte. KI-reisen dekker mange viktige aspekter ved implementering av KI, men det har blitt identifisert flere mangler og områder som krever ytterligere oppmerksomhet.

I Datatilsynets sandkasse for personvernvennlig innovasjon og digitalisering våren 2024 har det vært tre prosjekt der tema har vært innføring av, eller utvikling og innføring av generative språkmodeller. På NOKIOS-konferansen 22-24. oktober 2024 ble det arrangert en workshop for å sette fokus på hva det innebærer for en organisasjon å ta i bruk slike språkmodeller.

Viser disse prosjektene at man bør ta i bruk slike språkmodeller? Hva må man i så fall være oppmerksom på?

- Hva er forskjellen på egentilpassede løsninger eller språkmodeller integrert i eksisterende plattformer som f.eks. Microsoft Copilot?
- Hva innebærer det å bruke egne strukturerte data på toppen av språkmodellene?
- Hvordan kan man gå fram for å vurdere om språkmodeller er noe som kan være nyttig i egen virksomhet?
- Hva må man være klar over før man starter innføring?
- Har man kompetansen og tilstrekkelig bevissthet om utfordringene?
- Har man tilstrekkelig «orden i eget hus» (styringssystemer, informasjonssikkerhet og dataforvaltning)
- Har man tilstrekkelig behandlingsgrunnlag for de data man ønsker å bruke? (både til trening og til saksbehandling)
- Må man gjennomføre en personvernkonsekvensvurdering (DPIA), og hva skal i tilfelle en slik DPIA fokusere på?

Workshopen ble ledet og gjennomført av Heine Skipenes fra NTNU og Eirik Gulbrandsen fra Datatilsynet, med teknisk og metodisk støtte fra Silje Reiten Blichfeldt (NTNU), Hanne Jensen Moe (NTNU) og Marte Nubdal (NTNU).

På workshopen ble det samlet inn innsikter og synspunkter fra ca. 40 deltakere, hovedsakelig fra offentlig sektor. Deltakerne påpekte flere kritiske temaer og nye aspekter som ikke er tilstrekkelig dekket i KI-reisen. Disse inkluderer økonomi og ressursbruk, personvern og informasjonssikkerhet, bærekraft, juridiske avklaringer, etikk og tillit, opplæring og kompetanse, digitalt utenforskap, kontroll og forvaltning, samt innovasjon og ressursallokering.

NTNU har valgt å skrive denne innspillsrapporten for å adressere manglene og gi en mer utfyllende oversikt over muligheter og utfordringer ved bruk av KI i offentlig sektor. Rapporten er utarbeidet for å gi mer innsikt for Datatilsynet, samt andre offentlige virksomheter som vurderer å implementere KI. Målet er å sikre at implementeringen av KI skjer på en måte som ivaretar personvern, informasjonssikkerhet og bærekraft, samtidig som det skapes reell verdi og grunnlag for effektivisering.

Vi håper at denne rapporten gir en dypere forståelse av kritiske temaer og nye, mer utfyllende perspektiver som deltakerne på workshopen uttrykte. I tillegg håper vi at rapporten inneholder mer konkrete anbefalinger for hvordan disse temaene kan håndteres. Ved å inkludere disse innsiktene, ønsker vi å bidra til en tryggere og mer effektiv implementering av KI i offentlig sektor.

NOKIOS-workshop 22. oktober 2024

På workshopen deltok rundt 40 personer fra omtrent 20 ulike offentlige organisasjoner. Omtrent halvparten av deltakerne rapporterte at deres virksomhet offisielt hadde tatt i bruk språkmodeller, men de aller fleste (78 %) oppga at de selv hadde tatt dem i bruk i jobbsammenheng. De fleste benyttet kjente språkmodeller og rapporterte bruk av både Copilot i Edge, Microsoft 365 Copilot, Microsoft Copilot (med Enterprise Data Protection, tidligere Commercial Data Protection), ChatGPT og andre verktøy.

I og med at dette kun er et øyeblikksbilde fra en 4 timers workshop på et veldig begrenset utvalg av mennesker vil man ikke kunne trekke noen form for konklusjoner - kun indikasjoner. Det er likevel interessant å gjøre enkle undersøkelser da dette kan avdekke både funn og tema som bør undersøkes nærmere. For eksempel: Deltakerne ble bedt om å skåre seg selv på en fempunktsskala fra veldig dårlig/negativ (1) til veldig bra /positiv (5) ut fra følgende tre spørsmål:

Spørsmål (35 respondenter)	1	2	3	4	5
Mitt kunnskapsnivå om kunstig intelligens	*0 %*	*17%*	***40 %***	***40 %***	*3 %*
Min evne til å bruke KI på en god måte i jobben min	*0%*	***54 %***	*20 %*	*26 %*	*0 %*
Min generelle holdning til kunstig intelligens og ny teknologi	*3 %*	*0 %*	*29 %*	***51 %***	*17 %*

Med dette som utgangspunkt: Er det slik at ansatte med jevnt høy egenvurdering av eget kunnskapsnivå og en generell positiv holdning til KI og ny teknologi ikke klarer å ta det i bruk godt nok i jobbsammenheng? Ut fra tallmaterialet kan det se slik ut, og det kan indikere at organisasjonene bør prioritere tiltak for å belyse faktiske nyttige bruksområder, reelt gevinstpotensial og dra nytte av motivasjonen som innovatører («early adopters») kan ha i en organisasjon:

1. *Økt innovasjon*: Early adopters er ofte mer åpne for å eksperimentere med nye teknologier, noe som kan føre til innovative løsninger og forbedringer i arbeidsprosesser[1]
1. *Konkurransefortrinn*: Ved å være tidlig ute med å implementere KI, kan organisasjoner oppnå et konkurransefortrinn ved å forbedre effektiviteten, redusere kostnader og tilby bedre tjenester eller produkter[2].
2. *Bedre beslutningstaking*: KI kan analysere store mengder data raskt og nøyaktig, som igjen kan gi bedre innsikt og grunnlag for beslutninger[3].
3. *Tiltrekke og beholde talenter*: Organisasjoner som er kjent for å være teknologisk avanserte, kan lettere tiltrekke seg og beholde dyktige medarbeidere som ønsker å jobbe med cutting-edge teknologi[4].
4. *Økt produktivitet*: Automatisering av rutineoppgaver gjennom KI kan frigjøre tid for ansatte til å fokusere på mer strategiske oppgaver, noe som øker den samlede produktiviteten[2].
5. *Tilpasning til endringer*: Økt fleksibilitet kan gjøre organisasjoner bedre rustet til å tilpasse seg teknologiske endringer og markedstrender[3].

Men hvordan kan man best gjøre det? På workshopen fikk alle deltakere lov til å tenke høyt og fritt om hvordan man kan tilnærme seg dette temaet fra ulike synsvinkler. Den samlede intelligensen som 40 mennesker på workshop har, resulterte i ekte intelligens som vi har bygd denne rapporten på ved hjelp av KI-verktøy på NTNU. Datainnsamling ble gjort gjennom verktøyet «Mentimeter» (menti.com), og fysisk gjennom post-it på flipover ark. Tekstgjenkjenning ble gjort av ChatGPT og kontrollert av menneske før bearbeiding i KI-verktøy.

[1] (https://www.nho.no/tema/digitalisering/artikler/ny-rapport-kunstig-intelligens-kan-oke-norges-verdiskaping-betydelig/
[2] https://www.regjeringen.no/no/tema/statlig-forvaltning/it-politikk/ny-nasjonal-digitaliseringsstrategi/utnytte-mulighetene-i-kunstig-intelligens/id3054706/
[3] https://teknologiradet.no/rapport-kunstig-intelligens-norge/
[4] https://www.nho.no/tema/digitalisering/Kunstig_intelligens_i_Norge_Rapportsammendrag_SOA2023/

Dilemmaene vi alle står i

Workshopen ble designet på en slik måte at den skulle ende opp med en liste over dilemmaer som er typisk for alle organisasjoner som vurderer bruk av KI-verktøy. Dette er dilemmaer og problemstillinger som kan være nyttige for de som jobber med å utvikle informasjonsmateriell og støttesystemer for særlig offentlig sektor. Nedenfor har vi beskrevet de viktigste dilemmaene som ble identifiser. Temaene er ikke listet opp i prioritert rekkefølge:

Økonomi og ressurser

Deltakerne uttrykte bekymring for om de har tilstrekkelige økonomiske ressurser til å implementere KI-løsninger. De stilte spørsmål ved om investeringene vil gi tilstrekkelig avkastning og om de har nok kunnskap og ressurser til å prioritere og organisere arbeidet internt. Det ble også nevnt at det er vanskelig å lage gode business cases som forsvarer alle nødvendige investeringer.

Tillit og personvern

Bruken av KI kan sette tilliten til offentlig forvaltning på spill. Deltakerne var bekymret for hvordan KI kan påvirke for eksempel personvernet, spesielt med tanke på hallusinasjoner og manglende kompetanse blant brukerne. Det var også bekymringer rundt kontrollen av sensitive data og om verktøyene er for kraftige uten tilstrekkelige kontrollsystemer. I feil hender kan dette gå ordentlig galt.

Juridiske utfordringer

Mange deltakere nevnte at lovverket ikke er oppdatert for å håndtere ny teknologi som KI. Det er uklare og sprikende juridiske vurderinger rundt personvern og andre relaterte områder. Det ble også nevnt at det mangler gode og sikre fellesløsninger for å hente inn og sende data mellom ulike aktører.

Bærekraft

Bærekraft var et annet viktig tema. Deltakerne stilte spørsmål ved om KI kan implementeres på en bærekraftig måte og om bærekraftvurderingene som gjøres tar hensyn til alle kostnadene, inkludert eksterne kostnader. Det var også bekymringer rundt energiforbruket knyttet til KI.

Samfunnsmessige og etiske dilemmaer

Etiske dilemmaer ble også fremhevet. Deltakerne var bekymret for om KI kan brukes på en måte som ivaretar etiske prinsipper og om det er tilstrekkelig vurdering av de potensielle negative konsekvensene av KI. Det ble også diskutert om alle virksomheter skal finne ut av dette selv eller om innsatsen skal samles. Deltakerne uttrykker frykt for at KI kan erstatte menneskelig arbeid og føre til færre jobber. Det er også bekymringer rundt rettferdighet og objektivitet i KI-systemer, samt hvordan teknologien kan påvirke samfunnets verdier og levekår. Noen deltakere er også skeptiske til om KI virkelig vil være spennende og nyttig på lang sikt.

Kompetanse og innovasjon

Det ble påpekt at det er for få personer som forstår både teknologi, juss og gode anvendelsesområder for KI. Dette fører til dårlige eller feilaktige avklaringer og løsninger, samt lite riktig innovasjon. Deltakerne var også bekymret for om de har nok ressurser til å avsette til innovasjon.

Autonomi og kvalitet

Deltakerne diskuterte også dilemmaet mellom graden av autonomi KI-systemer kan ha og kvaliteten på resultatene. Det var bekymringer rundt om KI er vurdert godt nok til å kunne si noe om den faktiske gevinsten eller de nye utfordringene KI bringer med seg.

Teknologiens enorme potensial?

Deltakerne ble utfordret på hvilket potensial de ser i denne teknologien, og samtidig utfordret på hvorvidt de tror på vår (felles) gjennomføringsevne:

Spørsmål (30 respondenter)	Ja	Nei	Vet ikke
Går det an å gjøre det? Har du trua?	*61 %*	*16%*	*24 %*

Selv om deltakerne har identifisert en rekke dilemmaer, er fortsatt 61 % av deltakerne positive til at det er mulig å implementere KI og oppnå suksess. Dette viser at til tross for utfordringene, er optimismen og tilliten til KI sitt potensial fortsatt sterk blant deltakerne:

Effektivisering og automatisering

Deltakerne ser et stort potensial i KI for å effektivisere arbeidsprosesser og redusere manuelt arbeid. KI kan hjelpe med å sortere og navigere i store mengder informasjon, finne relevant informasjon raskere, og redusere tiden brukt på saksbehandling og rutineoppgaver. Dette kan frigjøre ressurser til andre oppgaver og gi mer tid til det som virkelig krever menneskelig innsats.

Beslutningsstøtte og kvalitetsforbedring

KI kan forbedre beslutningsprosesser ved å presentere relevant informasjon mer effektivt og gi støtte til å ta bedre beslutninger. Deltakerne nevner også at KI kan bidra til kvalitetsøkning i tjenestene og redusere menneskelige feil og bias. Dette kan føre til bedre resultater og økt verdiskapning med mindre bruk av menneskelige ressurser.

Demokratisering av informasjon og kunnskap

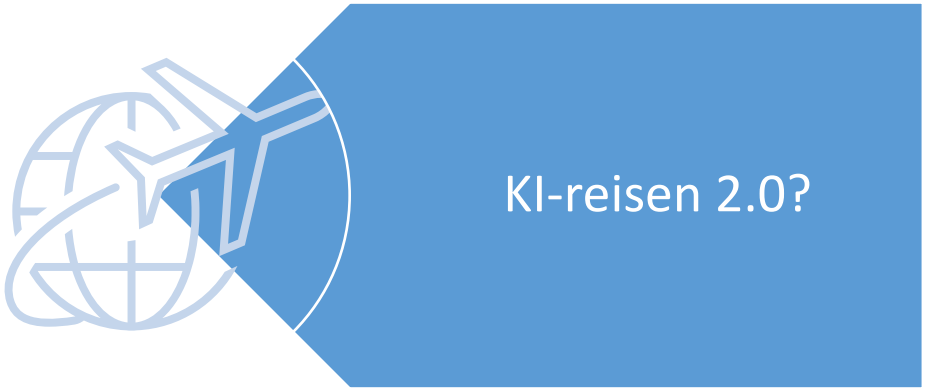
En annen viktig tematikk er hvordan KI kan bidra til demokratisering av informasjon og kunnskap. Deltakerne ser potensialet i at KI kan gjøre informasjon mer tilgjengelig og dele kunnskap på en måte som tidligere ikke var mulig. Dette kan bidra til økt livskvalitet og bedre sosiale evner i fremtidige generasjoner.

Teknologiske fremskritt og innovasjon

KI kan muliggjøre utvikling av autonome agenter og tjenester som kan integreres og utvikles på en skala vi aldri har sett før. Dette kan føre til fremskritt innen forskning og utvikling av nye teknologier. Deltakerne nevner også at KI kan hjelpe med å strukturere og kvalitetssikre data, noe som kan være utfordrende for mennesker.

Fremtidige muligheter og utfordringer

Til slutt ser deltakerne både muligheter og utfordringer i fremtiden med KI. Det er en mulighet for å forbedre psykisk og fysisk helse, øke trygghet og velferd, og redusere behovet for unødvendig rapportering. Samtidig er det viktig å forstå og forme bruken av teknologien for å sikre at den brukes på en måte som er i tråd med samfunnets verdier.


KI-reisen 2.0?


Tilbakemeldinger fra NOKIOS workshop 22 oktober 2024. Hvordan gjør vi KI-reisen bedre?

Hva er det som mangler i NTNUs KI-reise?

NTNU lever etter verdiene kreativ, kritisk, konstruktiv og respektfull i alle sine prosjekter og initiativer. Disse verdiene danner også grunnlaget for vår tilnærming til KI og er avgjørende for å sikre at våre løsninger er innovative, pålitelige og etisk forsvarlige. I tråd med dette ble deltakerne på workshopen oppfordret til å benytte disse verdiene, og spesielt være kritiske i sine tilbakemeldinger på NTNUs KI-reise. NTNU kan på ingen måte garantere at vi har laget en god nok framstilling, og kritisk tenkning er essensielt for å identifisere svakheter og forbedringsområder. Det er gjennom denne typen prosesser vi kan sikre at våre KI-valg møter de høyeste standardene. Ved å oppmuntre til kritisk refleksjon, ønsker vi å skape en kultur der alle bidrar til kontinuerlig forbedring og innovasjon. På denne måten blir vi bedre.

Gjennom workshopen har vi samlet verdifulle innspill som belyser hva som mangler i vår nåværende KI-reise. Det er viktig å påpeke at disse innspillene og tilbakemeldingene ikke kommer i samme rekkefølge som punktene som er satt opp i den opprinnelige KI-reisen. Disse tilbakemeldingene vil være viktige når vi former fremtidige strategier og tiltak.

1. Konseptet «KI-reise» kan være misvisende

Selv om NTNU har valgt å kalle det en "KI-reise", var tilbakemeldingen under workshopen krystallklar på at dette navnet ikke holder mål. Begrepet er misvisende og uegnet for å beskrive den komplekse og dynamiske prosessen som KI-implementering innebærer. Dette poenget ble fremhevet med stor overbevisning, og det er tydelig at navnevalget bør revurderes for å reflektere den virkelige naturen av KI-arbeidet.

2. Klare løsninger og oppgaver

Deltakerne på workshopen påpekte behovet for å tydelig definere hvilke løsninger KI skal hjelpe med, samt hvilke oppgaver KI ikke skal brukes på. Dette vil bidra til å sette klare forventninger og rammer for KI-bruken, og bidrar til å unngå forvirring og ineffektivitet, og sikrer at ressursene brukes på en målrettet måte. Ved å ha en klar forståelse av hva KI skal brukes til, kan man lettere måle suksess og identifisere områder for forbedring. Dette innebærer også å sette opp spesifikke mål og indikatorer for å evaluere effekten av KI-løsningene. For eksempel, hvis KI skal brukes til å forbedre kundeservice, bør det være klart definerte mål for responstid, kundetilfredshet og løsningsrate.

3. Mulighet for å avbryte KI-reisen

Det ble fremhevet at det mangler en mekanisme for å avbryte KI-reisen dersom det skulle være nødvendig. Dette kan være viktig for å håndtere uforutsette situasjoner og sikre fleksibilitet i KI-prosessen. Uten en slik mekanisme kan man risikere å fortsette på en feilaktig vei, noe som kan være kostbart og tidkrevende å rette opp senere. Det gir også en trygghet for brukerne om at de kan trekke seg tilbake hvis noe går galt. Dette kan inkludere å ha klare prosedyrer for å stoppe eller pause KI-prosjekter, samt å ha beredskapsplaner for å håndtere eventuelle negative konsekvenser.

4. Definering av målgrupper og behov

For å sikre at KI-løsningene er relevante, er det viktig å definere målgrupper og deres behov. Dette vil bidra til å skreddersy KI-løsningene slik at de møter brukernes forventninger og krav. Uten en klar forståelse av målgruppene kan KI-løsningene bli generelle og mindre effektive. Ved å forstå målgruppenes spesifikke behov, kan man utvikle mer brukervennlige og tilpassede løsninger som gir større verdi. Dette innebærer også å gjennomføre grundige analyser og undersøkelser for å kartlegge målgruppenes preferanser, utfordringer og forventninger.

5. «Prompt-skole»

Etablering av en "prompt-skole" ble foreslått for å lære brukerne hvordan de best kan ta i bruk KI. Dette vil øke brukernes kompetanse og effektivitet i bruken av KI-verktøy. Dette er viktig for å sikre at KI-verktøyene brukes riktig og gir de ønskede resultatene. Uten tilstrekkelig opplæring kan brukerne misforstå eller misbruke teknologien, noe som kan føre til suboptimale resultater. En «prompt-skole» kan også bidra til å bygge en kultur for kontinuerlig læring og forbedring. Dette kan inkludere kurs, workshops og opplæringsprogrammer som dekker alt fra grunnleggende KI-konsepter til avanserte teknikker for å optimalisere bruken av KI-verktøy.

6. Konsekvenser og evaluering

Det er behov for å definere tjenester, evaluere KI-bruken, og vurdere risiko, personvernkonsekvenser og informasjonssikkerhet. Dette vil bidra til å sikre at KI-løsningene er trygge og pålitelige. Uten grundig evaluering kan det oppstå sikkerhetsbrudd eller andre uønskede konsekvenser som kan skade både brukerne og organisasjonen. Evaluering bidrar også til å identifisere forbedringsområder og sikre at KI-løsningene oppfyller de nødvendige kravene. Dette kan inkludere å gjennomføre risikoanalyser, personvernvurderinger og sikkerhetstester, samt å etablere klare retningslinjer og prosedyrer for å håndtere eventuelle problemer som oppstår.

7. Kompetansebehov og opplæring

Deltakerne påpekte behovet for opplæring og kompetanseløft innen ny teknologi. Behovet for opplæring og kompetanseløft innen ny teknologi er avgjørende for å holde tritt med utviklingen. Tydelig informasjon til brukerne om logging og arbeidsprosessene «på baksiden» sikrer transparens og forståelse. Uten tilstrekkelig opplæring kan det være vanskelig å implementere og vedlikeholde KI-løsningene på en effektiv måte. Kompetanseheving bidrar også til å øke brukernes tillit til teknologien og deres evne til å bruke den riktig. Dette kan inkludere å tilby kurs, sertifiseringsprogrammer og kontinuerlig faglig utvikling for ansatte, samt å etablere interne ressurser og støttesystemer for å hjelpe brukerne med å forstå og bruke KI-teknologi.

8. Eksperimentering og pilotering

For å teste og forbedre KI-løsningene, er det behov for eksperimentering og pilotering. Dette bidrar til å identifisere og løse eventuelle utfordringer tidlig i prosessen. Uten denne fasen kan man risikere å implementere løsninger som ikke fungerer optimalt, noe som kan være kostbart å rette opp senere. Pilotering gir også verdifulle tilbakemeldinger fra brukerne som kan brukes til å finjustere løsningene. Dette kan inkludere å gjennomføre småskala piloter, brukertester og eksperimenter for å evaluere effektiviteten og brukervennligheten av KI-løsningene, samt å samle inn data og innsikt som kan brukes til å forbedre løsningene før de rulles ut i stor skala.

9. Usikkerhetsanalyse og bias

Analyse av usikkerhet og bias i språkmodeller trent på globale data, sett i forhold til lokale (norske) kontekster, er nødvendig. Dette vil bidra til å sikre at KI-løsningene er nøyaktige og relevante for den norske konteksten. Uten en slik analyse kan man risikere at løsningene ikke fungerer godt i den lokale konteksten, noe som kan redusere deres effektivitet og pålitelighet. Det er viktig å forstå og håndtere bias for å sikre rettferdige og inkluderende KI-løsninger. Dette kan inkludere å gjennomføre grundige analyser av dataene som brukes til å trene KI-modellene, samt å utvikle metoder og teknikker for å identifisere og redusere bias og usikkerhet i modellene.

10. Identitetsstyring

Det ble påpekt behov for bedre identitetsstyring for å sikre at KI-løsningene fungerer som de skal. Dette vil bidra til å beskytte brukernes identitet og data. Uten god identitetsstyring kan det oppstå sikkerhetsbrudd og misbruk av data, noe som kan ha alvorlige konsekvenser for både brukerne og organisasjonen. Identitetsstyring bidrar også til å bygge tillit hos brukerne og sikre at deres personlige opplysninger håndteres på en sikker måte. Dette kan inkludere å implementere

robuste autentiserings- og autorisasjonsmekanismer, samt å etablere klare retningslinjer og prosedyrer for å beskytte brukernes identitet og data.

11. Gevinstrealisering

Gevinstrealisering er et kritisk steg i enhver KI-reise, da det handler om å sikre at de forventede fordelene og verdiene fra KI-implementeringen faktisk blir oppnådd. Dette innebærer å identifisere, planlegge og følge opp på de konkrete gevinstene som KI-løsningene skal levere. Uten et fokus på gevinstrealisering kan det være vanskelig å måle suksessen av KI-prosjektene og rettferdiggjøre investeringene som er gjort. I NTNUs nåværende KI-reise er steget gevinstrealisering ikke tilstrekkelig adressert. Dette kan føre til at de potensielle fordelene ved KI-implementeringen ikke blir fullt ut realisert, og at det blir utfordrende å demonstrere verdien av KI-løsningene til interessenter og beslutningstakere. Uten klare mål og indikatorer for gevinstrealisering kan det også være vanskelig å identifisere og korrigere eventuelle avvik fra forventet ytelse.

12. Plan for vurdering av risiko og personvernkonsekvenser

En grundig plan for vurdering av risiko og personvernkonsekvenser er avgjørende for å sikre at KI-implementeringen gir de forventede fordelene samtidig som den håndterer potensielle risikoer og beskytter personvernet. Dette innebærer å identifisere, analysere og mitigere risikoer knyttet til personvern og informasjonssikkerhet. Uten en slik plan kan det være vanskelig å oppnå en balansert og ansvarlig KI-implementering. I NTNUs nåværende KI-reise er det ikke tilstrekkelig fokus på å utvikle en omfattende plan for å vurdere risiko og personvernkonsekvenser. Dette kan føre til at risikoer knyttet til personvern og informasjonssikkerhet ikke blir tilstrekkelig håndtert. Uten klare retningslinjer og vurderinger kan det oppstå utfordringer som kan skade både brukerne og organisasjonen. Det er derfor viktig å inkludere en strukturert tilnærming til å identifisere, vurdere og håndtere disse aspektene for å sikre en vellykket og ansvarlig KI-implementering. Deltakerne påpekte et ønske om en standardisert mal for ROS og DPIA, for å sikre en helhetlig og systematisk tilnærming til risikovurdering og personvernkonsekvensanalyser.

13. Avgrensning til egen sektor og relevante lovverk

Dette er et viktig tema å vurdere, da lovverk og reguleringer ofte varierer mellom sektorer. Å avgrense til egen sektor kan være utfordrende, spesielt når det gjelder KI, som ofte krysser sektorer og bransjer. Det er derfor viktig å ha en helhetlig tilnærming som tar hensyn til både sektor-spesifikke og generelle lovverk for å sikre at KI-løsningene er i samsvar med alle relevante reguleringer. Dette vil bidra til å unngå juridiske komplikasjoner og sikre en ansvarlig og etisk implementering av KI.

14. Opphavsrett til generert tekst

Et viktig tema som ble tatt opp av en av deltakerne i workshopen, er opphavsrett til generert tekst. Når KI brukes til å generere tekst, oppstår det spørsmål om hvem som eier rettighetene til det skapte innholdet. Dette er spesielt relevant i kontekster der KI-verktøy brukes til å produsere originalt innhold som kan ha kommersiell verdi. Det er avgjørende å avklare opphavsrettslige spørsmål for å sikre at både brukere og utviklere av KI-teknologi har klare retningslinjer å forholde seg til. Dette vil bidra til å unngå juridiske tvister og sikre at rettighetene til generert innhold blir respektert og beskyttet.

Å inkludere opphavsrett til generert tekst i NTNUs KI-reise er essensielt for å sikre en ansvarlig og etisk bruk av KI. Ved å ha klare retningslinjer og prosedyrer på plass, kan NTNU beskytte både sine egne og brukernes interesser, samtidig som de fremmer innovasjon og kreativitet innen KI-feltet. Dette vil også bidra til å bygge tillit blant brukerne og sikre at KI-løsningene brukes på en måte som er i samsvar med gjeldende lovverk og etiske standarder.

15. «Markedspress» og generativ KI

Tilbakemeldinger i workshopen gikk også ut på hvorvidt NTNU opplevde «markedspress» rundt det å ta i bruk generativ KI, gitt NTNUs posisjon i Norge. Dette er et relevant tema, da NTNU som en av de største utdannings- og forskningsinstitusjonene i Norge kan oppleve press fra markedet og samfunnet for å være i forkant av teknologiske innovasjoner. Å ta i bruk generativ KI kan sees på som en nødvendighet for å opprettholde konkurranseevnen og relevansen i et stadig mer digitalisert samfunn. Samtidig er det viktig å balansere dette presset med ansvarlig og etisk implementering av KI-teknolog, for å sikre at de brukes på en måte som gagner både institusjonen og samfunnet som helhet.

16. Grunnlaget – Orden i eget hus

Før man kan implementere og dra nytte av KI, må man først sikre at de grunnleggende systemene og prosessene internt er i orden. Dette inkluderer å ha klare retningslinjer, robuste datasystemer, og en sterk infrastruktur som kan støtte KI-initiativene. Uten et solid fundament kan det være utfordrende å oppnå de ønskede resultatene fra KI-implementeringen, og risikoen for feil og ineffektivitet øker. Å ha "orden i eget hus" er derfor en forutsetning for å kunne utnytte KI-teknologiens fulle potensial på en ansvarlig og effektiv måte.

Øvrige tilbakemeldinger fra deltakerne

I tillegg til punktene over kom deltakerne med en rekke tilbakemeldinger som kan bli nyttige i det videre arbeidet.

Det deltakerne liker

Deltakerne i workshopen uttrykte flere positive aspekter ved KI-reisen. KI-reisen ble beskrevet som oversiktlig og enkel å følge steg for steg, samt at den praktiske tilnærmingen ble beskrevet som god. Deltakerne satte pris på at vurderingen av hva KI faktisk skal brukes til, og at problemstillingen var åpen og relevant for mange. Fokus på opplæring ble også fremhevet som viktig. Deltakerne likte at det var stor vekt på personvern og at KI-reisen består av relevante spørsmål knyttet til dette. Dette er også i tråd med den nye Digitaliseringsstrategien. Det bevisste forholdet til gjeldende regelverk og utgangspunktet i eksisterende rammebetingelser ble også verdsatt. Det at man bør utarbeide en exit-strategi ble også ansett som bra, selv om mange også påpekte at det burde være større fokus på dette punktet. Definerte bruksområder og gevinster ble fremhevet som positive aspekter, og det var klart når KI ikke skulle brukes.

Det deltakerne ikke liker

Deltakerne uttrykte flere bekymringer og kom med en rekke forslag til forbedringer for KI-reisen. De påpekte behovet for å ta høyde for kontinuerlig endring gjennom en sirkulær modell. Deltakerne stilte også spørsmål ved behovet for en egen KI-reise og understreket viktigheten av å vurdere tidligere erfaringer med annen teknologiutvikling.

Flere deltakere mente at det ikke bør kalles en reise, og at vurdering av forvaltning av data, inkludert kvalitet og tilgang, må forbedres. Det er viktig med endringer og oppfølging av driftskrav, samt å at den mangler nok fokus på exit-strategi.

Deltakerne foreslo at det bør være enklere å avgrense hva modellen trenes på enn hva den brukes til, og stilte spørsmål ved om exit-strategi er en reell mulighet. De ønsket også å vite mer om opt-out-alternativer for ansatte og hvordan potensialet ble vurdert, inkludert oppsiden. Forankring av strategien opp mot KI-reisen bør tydeliggjøres.

Det var også bekymringer rundt realismen i bærekraftvurderingen, og deltakerne påpekte at det er store problemer med KI og bærekraft. De stilte spørsmål ved om det egentlig er et reelt valg å ikke bruke KI. Deltakerne foreslo å bytte rekkefølgen på stegene "avdekk relevante lover og

regler for egen sektor" og "hvilke områder skal KI brukes på", og mente at steget som handler om retningslinjer må komme før opplæring.

Generelle innspill satt opp mot dilemmaene vi står i:

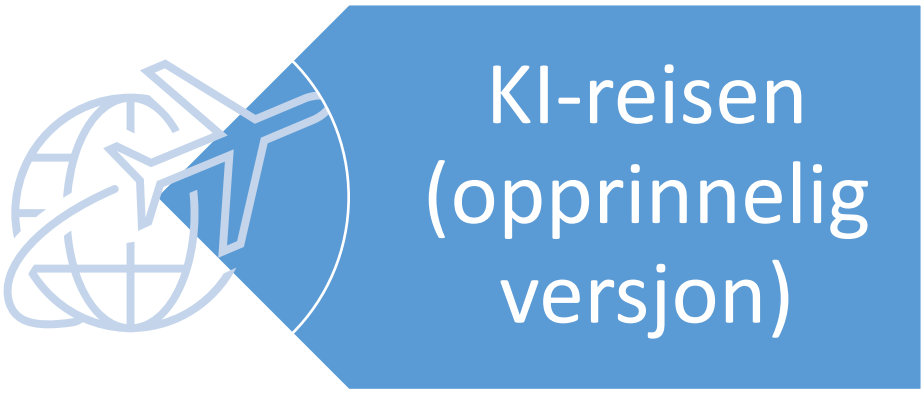
- Økonomi og ressursbruk: Deltakerne i workshopen påpekte at økonomi og ressursbruk er en stor utfordring ved implementering av KI. Det er viktig å balansere kostnader mot nytteverdien av KI, og dette temaet er ikke tilstrekkelig dekket i KI-reisen.
- Personvern og informasjonssikkerhet. Selv om KI-reisen nevner personvern og informasjonssikkerhet, understreker deltakerne at det er utfordrende å ivareta personvern og god bruk av KI. Spesielt er det bekymringer rundt hvordan KI kan prosessere personopplysninger og potensielt diskriminere brukere.
- Bærekraft er et viktig tema som deltakerne mener bør få mer oppmerksomhet. Dette inkluderer vurderinger av energiforbruk, ressursbruk og miljøpåvirkning ved bruk av KI.
- Juridiske avklaringer: Deltakerne påpekte behovet for klare juridiske avklaringer rundt bruk av KI, spesielt med tanke på hvilke data som kan eksponeres for KI og hvordan lovverket må oppdateres for å ta høyde for ny teknologi.
- Etikk og tillit: Etiske vurderinger og tillit til offentlig forvaltning er kritiske temaer som deltakerne mener bør få mer oppmerksomhet. Dette inkluderer hvordan KI kan påvirke tilliten til offentlige tjenester og hvordan etiske retningslinjer kan implementeres.
- Opplæring og kompetanse: Deltakerne understreket viktigheten av kontinuerlig opplæring og kompetanseheving for å sikre riktig bruk av KI. Dette inkluderer både generell opplæring og spesifikk opplæring knyttet til nye oppdateringer og endret bruk.
- Digitalt utenforskap: KI kan føre økt til digitalt utenforskap hvis ikke alle brukere får tilstrekkelig opplæring og støtte. Dette temaet er ikke tilstrekkelig dekket i KI-reisen.
- Kontroll og forvaltning. Deltakerne påpekte at det er utfordrende å forvalte KI-verktøy på en bærekraftig måte, spesielt med tanke på kontinuerlig oppfølging og oppdatering. Dette inkluderer behovet for en klar exit-strategi og oppfølging av utvikling og risiko.
- Innovasjon og ressursallokering. Det ble også diskutert hvordan ressurser kan allokeres til innovasjon og hvordan KI kan brukes til å utvikle nye tjenester og løsninger. Dette temaet er ikke tilstrekkelig dekket i KI-reisen.

KI-reisen er kontinuerlig arbeid

Deltakerne har påpekt behovet for kontinuerlig endring og tilpasning, noe som tyder på at en statisk plan kanskje ikke er tilstrekkelig. En sirkulær modell som tar høyde for kontinuerlig evaluering og justering kan være nøkkelen til å håndtere både teknologiske og organisatoriske utfordringer. Dette innebærer å etablere mekanismer for regelmessig revisjon av strategien, inkludert evaluering av bruken av KI, identifisering av nye kompetansebehov, og tilpasning til endringer i lovverk. Markedspress er også viktig å adressere fortløpende. I tillegg er det viktig at implementeringen skjer på en måte som ivaretar personvern, informasjonssikkerhet og bærekraft.

Videre kan det være verdifullt å fokusere på å bygge en kultur for eksperimentering og læring innen organisasjonen. Dette kan inkludere pilotprosjekter, usikkerhetsanalyser, og en "prompt-skole" for å lære brukerne hvordan de best kan ta i bruk KI. En slik tilnærming kan bidra til å sikre at KI-løsningene er tilpasset brukernes faktiske behov og ferdigheter, samtidig som det gir rom for å identifisere og adressere eventuelle bias i språkmodeller.

Til slutt, å ha en klar exit-strategi og opt-out-alternativer for ansatte kan bidra til å bygge tillit og sikre at KI-reisen oppleves som en reell mulighet, snarere enn en tvang. Dette kan også inkludere en realistisk vurdering av bærekraft og en tydelig beskrivelse av roller og ansvar i KI-prosessene.


KI-reisen
(opprinnelig
versjon)

KI-reisen – Opprinnelig versjon

NTNU har som mange andre i offentlig og privat sektor vurdert mulighetene for å ta i bruk kunstig intelligens (KI) i arbeidshverdagen. KI er en teknologi som kan bidra til å løse mange samfunnsutfordringer, effektivisere offentlige tjenester og skape nye muligheter for innovasjon og verdiskaping. Å ta i bruk KI i riktig form kan gi uante gevinster. Dette er derfor et teknologisk taktskifte man bør ta stilling til om man bør være med på, men det må skje i trygge rammer.

Bruk av KI også fører med seg noen problematiske områder. Prosjektgruppen har gjennom Copilot-prosjektet forsøkt å avdekke hvilke områder som er spesielt viktig å fokusere på før, under og etter anskaffelse av et informasjonssystem som anvender KI. Hva skiller seg fra tidligere informasjonssystemer og er det områder KI bør anvendes og ikke bør anvendes?

KI-reisen er et forsøk på å forberede ulike virksomheter på å kunne ta i bruk KI på en trygg måte. Den bygger på DFØ sin veiledning «Sky-reisen»[5] med momenter som er erfart fra prosjektet. KI-reisen består av flere trinn der det første trinnet omhandler KI-strategi og er det forberedende trinnet til å ta i bruk KI. Deretter kommer vurderinger av tjenester og anskaffelse som gjøres per informasjonssystem, før den tar for seg forvaltningen av systemet i dets levetid.

KI-strategi
Å utarbeide en KI-strategi er et naturlig første steg i prosessen for å klargjøre virksomheten for å ta i bruk KI. Ordet strategi kan for mange virke tungt og vanskelig og det er viktig å tenke på at dette er en del av en modningsprosess. Om det ender opp i et eget strategi-dokument eller dokumenteres på andre måter, for eksempel hektes på eksisterende strategier, retningslinjer og rutiner er opp til hver virksomhet, men å ha vært igjennom prosessen i forkant er viktig.

KI-strategien utarbeidet som forslag fra dette prosjektet består av syv trinn virksomheter bør være igjennom før den tar i bruk KI. Totalt sett utgjør strategien et flytskjema som også illustrerer hvilken rekkefølge disse trinnene bør gjennomføres i. I dette kapittelet vil hvert enkelt av punktene beskrives i større detalj.

Avdekke relevante lover
Første steget i prosessen er å avdekke hvilke lover og regler som er eller vil bli relevante for egen sektor. Listen kan være veldig varierende basert på hva systemet skal brukes til og hvilke informasjon som skal prosesserer, men å ha gjort denne jobben i forkant vil gjøre det lettere å avdekke hvilke områder KI kan anvendes på og ikke kan anvendes på.

Noen lover og regler vil være veldig generelle for alle sektorer. Dette gjelder for eksempel Personvernforordningen, Likestillings- og diskrimineringsloven osv. I tillegg kommer det en rekke nye direktiver og forordninger fra EU som kan være med på å sette føringer for bruk av KI fremover. Både NIS-2 direktivet og KI-forordningen nærmer seg, og det er

[5] https://markedsplassen.anskaffelser.no/veiledning/skyreisen

viktig at virksomhetene setter seg godt inn i hva dette kan bety for egen virksomhet på et tidlig tidspunkt.

For offentlig sektor er det også en rekke ulike regelverk som kan være aktuelle ved bruk av KI. I tillegg skal virksomheter i offentlig sektor ta høyde for det til enhver tid gjeldende digitaliseringsrundskriv. Digitaliseringsrundskrivet er en sammenstilling av pålegg og anbefalinger om digitalisering i offentlig sektor[6]. Rundskrivet refererer til relevant lovverk som er gjeldende, og til eksisterende veiledninger og kan være til hjelp i en slik prosess.

Hvilke områder kan KI brukes på
Hva skal KI hjelpe med, hvilke mål og gevinster ønsker man å oppnå ved å ta i bruk KI, og hvilke områder kan KI brukes eller ikke brukes på? Dette er viktige spørsmål å stille i forberedelsene for å ta i bruk KI. Ved å avdekke hvilke områder virksomheten ønsker hjelp fra KI vil gjøre det lettere å finne relevante KI-verktøy. Våre vurderinger er at KI er best på to ting: Det som ingen mennesker er gode på, altså analysere og sammenstille store datasett, og det som vi kan fra før, men som kan effektiviseres. På den måten kan vi som mennesker kontrollere utfall og forhindre eventuelle feil beslutninger.

Hvilke områder skal KI ikke anvendes?
KI har mange bruksområder og noen av disse områdene vil innebære høyere risiko for virksomheten, eller kan til og med være direkte ulovlig. Dette er helt avhengig av hvordan informasjonssystemet er bygget opp, hvilken informasjon systemet prosesserer, hvor den prosesseres osv. Det kan likevel være fint å avdekke disse områdene for å få gjort en ekstra vurderinger for bruk, om man ikke vurderer helt å droppe slik bruk.

Digdir har kommet med noen anbefalinger for hvor man bør være forsiktig ved bruk av KI[7]. De nevner blant annet at virksomheter bør være forsiktig ved bruk av KI mot innbyggere i tillegg til å være forsiktig å ta i bruk KI som oppslagsverk. I tillegg bør man være forsiktig med å legge inn sensitiv informasjon inn i «prompts» da denne informasjonen kan samles inn og gjenbrukes av andre.

Med bakgrunn i Digdirs anbefalinger og erfaringer og diskusjoner i løpet av prosjektet har vi kommet frem til noen anbefalinger:

- KI skal ikke brukes til å erstatte kompetanse. Vi må kunne avdekke feil som gjøres av verktøyet vi bruker.
- Skal KI brukes til å prosessere personopplysninger må man være ekstra varsom slik at relevant regelverk overholdes. For eksempel; lages det nye sammenstillinger av personopplysninger der man ikke kan kontrollere hvilke opplysninger som er brukt i sammenstilling? Eller kan tilfeldige sammenhenger brukes til å fatte beslutninger som ender med å diskriminere brukere?
- KI skal ikke brukes for å på noen som helst måte overvåke, sammenlikne eller vurdere egne ansatte.

Utarbeid plan for forvaltning

[6] https://www.regjeringen.no/no/dokumenter/digitaliseringsrundskrivet/id3025117/
[7] https://www.digdir.no/kunstig-intelligens/bruk-av-generativ-kunstig-intelligens-i-offentlig-sektor/4670

De aller fleste virksomheter har allerede en forvaltningsplan, men hvordan forvalte informasjonssystemer som brukes allerede og som har endret seg drastisk de siste årene? Fra å være «on prem»-løsninger der det ble kjøpt inn lisenser og ikke gjort noe mer enn oppdateringer før lisenser måtte fornyes, krever sky-løsninger og også nå KI-verktøy at forvaltningen gjøres på en helt ny måte. Den teknologiske utviklingen har skjedd i eksponentiell hastighet de siste årene, og informasjonssystemene som anvendes i dag følger denne teknologiske utviklingen. Å forvalte slike typer informasjonssystemer stiller derfor strengere krav til deler av forvaltningen. Noen områder vil gjøre seg spesielt viktige for å følge opp nye IT-systemer og da spesielt med tanke på IT-systemer som anvender kunstig intelligens:

- Avtaleoppføling
- Oppfølging av utvikling
- Oppdatert produktkunnskap
- Oppdaterte risikovurderinger
- Oppdatert opplæring og veiledning til brukere

Å henge bakpå i slik oppfølging kan føre til en situasjon der man står med et informasjonssystem med uante funksjonaliteter, men også uante risikoer og sårbarheter.

Utarbeid en exit-strategi

Ved å ta i bruk KI-verktøy er det også veldig viktig å ha tenkt over hvordan virksomheten kan avvikle bruken av verktøyet. Det kan være flere årsaker til at man ikke ønsker å bruke verktøyet lenger. Det kan for eksempel være økning i priser, en utvikling som går vekk fra opprinnelig ønske om å bruke verktøyet, en utvikling som gjør at bruk potensielt er ulovlig, økt risiko ved bruk av systemet eller andre årsaker. Exit-strategien bør ha avdekket disse linjene for akseptabel bruk slik at forvaltere kan oppdage og varsle beslutningsmyndighet hvis de overskrides. For virksomhetskritiske systemer bør man også ha vurdert hvilke exit-muligheter man har, er det alternative skyløsninger eller skal systemet tas ut av sky og tilbake til lokale løsninger.

Exit-strategi er også spesielt viktig i forbindelse med KI-verktøy for å kunne ta i bruk verktøy raskere. Da kan man ha lavere terskel for å ta i bruk et verktøy hvis virksomheten også har en plan for oppfølging og avvikling hvis verktøyet ikke gir ønsket gevinst eller hvis verktøyet har uønsket funksjonalitet/gir uønskede effekter.

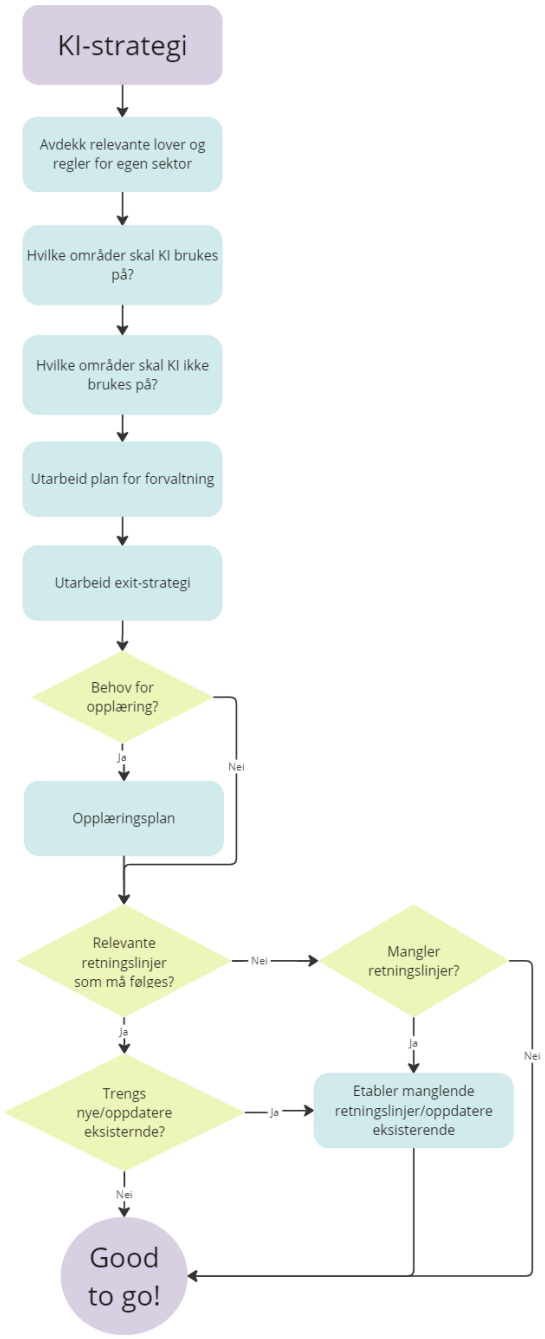
Identifiser behov for opplæring

Bruk av informasjonssystemer som er i kontinuerlig utvikling krever oppdatert informasjon til brukere. KI-verktøy spesielt kan i noen tilfeller kreve en ny måte å arbeide på for størst mulig gevinst, for riktig bruk og ikke minst for å sikre informasjonsverdier tilstrekkelig. Bruk av KI krever også en annen bevissthet i forhold til informasjonskompetansen til hver enkelt bruker. Flere av de åtte funnene i prosjektet hadde anbefaling om opplæring av brukere på ulike områder. Hvordan dette skal realiseres og eventuelt verifiseres eller kontrolleres bør virksomhetene ha en plan på i forkant. Under har vi listet noen momenter som kan vurderes før virksomheten tar i bruk KI:

- Virksomheter bør ha tenkt igjennom hvilken type opplæring som kreves, skal noe være obligatorisk.
- Bør det være en generell opplæring som alle ansatte skal igjennom, og hvordan skal denne gis.

- Skal det være opplæring per system, skal nye oppdateringer eller endret bruk komme frem i dokumentasjon eller opplæring.

Dokumenter strategi
Det siste punktet omhandler å få dokumentert det som har kommet frem i prosessen. Hvor dette gjøres er opp til hver virksomhet. Å hekte det på allerede eksisterende dokumentasjon vil etter prosjektets mening være enklest å gjennomføre og også holde oppdatert i forhold til annen dokumentasjon.



Figur – Flytskjema «steg for steg» i KI-reisen

Vurderinger per tjeneste

Før man tar i bruk et KI-verktøy i en offentlig virksomhet, bør man gjøre en grundig vurdering av de potensielle fordelene og ulempene, samt de juridiske, etiske og praktiske konsekvensene ved å ta i bruk det spesifikke verktøyet. Vi ser på slike vurderinger som spesielt viktig å gjøre i forbindelse med bruk av KI da KI har mange fordeler, men verktøy kan også brukes til andre formål enn opprinnelig tiltenkt. Prosjektet har kommet opp med spørsmål innenfor ulike kategorier som kan være med på å vurdere viktigheten av verktøyet eller systemet man vurderer å anskaffe.

Informasjonssikkerhet
Informasjonssikkerhet handler om å beskytte informasjonsverdiene våre mot uautorisert tilgang, endring, tap eller skade. Hva som har verdi, og hvilken verdi det har, er avhengig av tid, sted og person. Noe kan ha ulik verdi for ulike personer, organisasjoner, samfunn og nasjoner. Oversikt over hvilke verdier som eksisterer i informasjonssystemer der KI-verktøy skal anvendes er derfor grunnleggende for å kunne gjøre vurderinger rundt informasjonssikkerheten til verktøyet og om de er tilstrekkelig. Det henvises ofte til «Orden i eget hus» og er et ansvar hver enkelt ansatt må bidra med.

- Hvilke data prosesseres av KI-verktøyet, og hvor sensitive eller konfidensielle er de?
- Kan vi styre/begrense hvilke data som prosesseres?
- Hvem har tilgang til dataene, og hvordan kontrolleres og loggføres denne tilgangen?
- Hvordan sikres dataene mot uønskede hendelser som hacking, lekkasje, sletting eller manipulering?
- Deles dataene med noen andre aktører, som leverandører, samarbeidspartnere eller tredjeparter, og hvordan sikres dataene i disse tilfellene?

Personvern
Personvern går på mange områder hånd i hånd med informasjonssikkerhet, så flere av vurderingene som gjøres innenfor informasjonssikkerhet vil også være aktuelle for personvern.

Personvern handler om å respektere og beskytte de registrertes rettigheter og interesser når det gjelder deres personopplysninger. Dette blir av mange sett på som regulatoriske hindringer, men til syvende og sist handler det om å ivareta menneskerettigheter. Med det i bakhodet bør virksomheter gjøre noen ekstra vurderinger før KI-verktøy slippes løs på personopplysninger:

- Hvem blir registrert av KI-verktøyet, og hva slags personopplysninger samles inn, lagres og behandles?
- Hvor mye data lagres, og hvor lenge lagres de?
- Lages det nye sammensetninger av opplysninger, som profiler, segmenter eller prediksjoner, basert på dataene?
- Hvilket rettslig grunnlag har vi for å bruke KI-verktøyet, og hvordan informerer vi de registrerte om dette?
- Hvordan ivaretar vi de registrertes rettigheter, som innsyn, retting, sletting, begrensning, dataportabilitet og motstand?

I og med at KI er innovativ bruk eller anvendelse av ny teknologisk eller organisatorisk løsning, og hvis KI-verktøyet skal behandle personopplysninger, bør terskelen være lav for å gjøre en full personvernkonsekvensvurdering (DPIA). Gjennom en DPIA-prosess vil du få svar på spørsmålene over, og du vil få et bedre helhetsbilde over verktøyet og hvordan den registrertes rettigheter og friheter best ivaretas.

Menneskelig perspektiv
Menneskelig perspektiv handler om å ta hensyn til de sosiale, kulturelle og etiske aspektene ved bruk av KI, og å sikre at KI bidrar til å fremme menneskelig verdighet, autonomi, rettferdighet og inkludering.

- Hvordan påvirker KI-verktøyet de ansatte, brukerne og andre berørte parter, både positivt og negativt?
- Hvordan sikrer vi at KI-verktøyet ikke diskriminerer, stigmatiserer eller utestenger noen grupper eller individer på grunnlag av deres personlige kjennetegn, som alder, kjønn, etnisitet, funksjonsevne, religion eller seksuell orientering?
- Hvordan sikrer vi at KI-verktøyet ikke undergraver de ansattes kompetanse, motivasjon, arbeidsmiljø eller arbeidsvilkår?
- Hvordan sikrer vi at KI-verktøyet ikke krenker brukernes integritet, selvbestemmelse, tillit eller forventninger?
- Hvordan sikrer vi at KI-verktøyet er forståelig, forklarlig og kontrollerbart for de ansatte og brukerne, og at de har mulighet til å gi tilbakemelding, klage eller anke på KI-baserte beslutninger?
- Hvordan sikrer vi at KI-verktøyet er i tråd med de etiske verdiene og prinsippene som gjelder for vår virksomhet og sektor?
- Kan bruk av verktøyet føre til digitalt utenforskap?[8]

Bærekraft
Bærekraft handler om å vurdere de økonomiske, miljømessige og samfunnsmessige konsekvensene av å bruke KI, og å sikre at KI bidrar til å oppnå bærekraftsmålene og redusere klima- og miljøavtrykket. Dette innebærer å stille spørsmål som:

- Hvordan påvirker KI-verktøyet vår virksomhets økonomi, effektivitet og innovasjonsevne?
- Hvordan påvirker KI-verktøyet miljøet, både lokalt og globalt, i form av energiforbruk, utslipp, avfall og ressursbruk?
- Hvordan påvirker KI-verktøyet samfunnet, både nasjonalt og internasjonalt, i form av demokrati, rettsstat, menneskerettigheter, likestilling og sosial utjevning?
- Hvordan sikrer vi at KI-verktøyet er i samsvar med de bærekraftsmålene og de klima- og miljømålene som gjelder for vår virksomhet og sektor?
- Hvordan sikrer vi at KI-verktøyet er basert på ansvarlig og transparent datainnsamling, -deling og -bruk, som respekterer menneskers og samfunns interesser og verdier?
- Hvordan sikrer vi at KI-verktøyet er utformet, utviklet og implementert på en måte som fremmer sirkulær økonomi, grønn innovasjon og digital inkludering?

Behov
Å vurdere KI-verktøyet opp mot hvilke behov man har omhandler å vurdere bruksområdet opp mot egen KI-strategi. I tillegg bør det gjøres vurderinger rundt om verktøyet kan brukes til andre formål eller andre områder som er uønsket og potensiell effekt av bruk av verktøyet. Spørsmål man kan vurdere her er:

- Hva er gevinsten ved å ta i bruk verktøyet?
- Har vi mulighet til å kontrollere utfallet av bruken?
- Strider verktøyet med KI-strategien til virksomheten?
- Kan verktøyet misbrukes?
- Hva er mulig utfall hvis verktøyet brukes feil?

Anskaffelse

[8] https://www.regjeringen.no/no/dokumenter/digitaliseringsrundskrivet/id3025117/

Når man skal anskaffe et KI-verktøy i offentlig sektor, må man følge anskaffelsesloven og tilhørende forskrifter, som regulerer hvordan offentlige innkjøp skal gjennomføres. Anskaffelsesloven stiller allerede en del krav til ulike områder og hvordan disse skal vektlegges i en anskaffelsesprosess. Dette er områder som klima og miljøhensyn, menneskerettigheter, miljø og andre menneskehensyn, samt arbeidsforhold, sosiale forhold og arbeidslivskriminalitet.

I tillegg til å følge anskaffelsesloven er det viktig å få kravstilt annet relevant lovverk i forbindelse med den konkrete anskaffelsen. Dette kan være personvernforordningen eller andre lovverk som ble avdekket under arbeidet med en KI-strategi.

Når alle funksjonelle krav er stilt, skal man stille krav til den ikke-funksjonelle delen. Dette omfatter blant annet informasjonssikkerhet. Det er en balansegang mellom mengde A-krav og B-krav. For mange A-krav kan føre til få eller ingen tilbydere. For få A-krav kan føre til at man ender med et system eller verktøy som ikke er i henhold til egen virksomhet sine retningslinjer. A-krav innenfor ikke-funksjonell kategori bør derfor holdes til absolutte krav for virksomheten. Dette kan være sikring av data, hvem som skal ha tilgang til data osv.

Den teknologiske utviklingen gjør det ekstra vanskelig å stille krav til IT-systemer. Hvordan skal man stille krav til et system som er i kontinuerlig utvikling? Dette er viktig å ha med i betraktning i utforming av kravene som stilles slik at den fremtidige versjonen av systemet som anskaffes har samme teknologiske standard som ved anskaffelsestidspunkt. En offentlig anskaffelse kan også ta tid, så kravene som stilles må også ta høyde for utviklingen som skjer fra kravstillelse og til anskaffelse.

Hva som er relevante krav å stille i forbindelse med et KI-verktøy er per nå vanskelig å si. Vi har enda ikke så mye erfaring i anskaffelse av KI-verktøy eller IT-systemer som anvender KI eller på et tidspunkt kan anvende KI. Vi har laget noen forslag til krav som kan stilles under.

Krav	**Type**
KI-teknologi er godt dokumentert og tilgjengelig, inkludert formål og datakilder	A
Er det mulig å kontrollere/spore utfallet av KI	A
Beskriv eksisterende funksjonalitet som anvender KI eller annen automasjon og hva som er gevinsten av slik bruk	B
Beskriv planlagt utvikling av systemet og utvikling av KI spesielt.	B
Beskriv hvordan løsningen møter kravene fra kommende KI-forordningen	B
Beskriv hvordan KI-funksjonalitet er testet og godkjent før det implementeres i systemet	B
Beskriv hvordan bruker kan tilpasse, inkludert slå av/på, KI-funksjonalitet i løsningen	B
Beskriv hvordan løsningen håndterer dataene KI er trent på, spesielt med tanke på informasjonssikkerhet og personvern	B
NB: Legg til spesielle krav for egen organisasjon (Hva er viktig for dere?)	

Forvaltning

Hvordan forvalte IT-systemer på en god og forsvarlig måte er kanskje den største endringen fra tidligere IT-systemer og dagens IT-systemer. Utvikling av systemer og avtaler skjer kontinuerlig og krever derfor kontinuerlig oppfølging. Mangelfull forvaltning som tidligere har vært tilstrekkelig vil nå gi betydelige utfordringer. Dette er ikke nytt for kunstig intelligens, men har gradvis blitt mer gjeldende med sky-løsninger og hyppig teknologisk utvikling. I dette prosjektet har vi valgt å kalle det et forvaltningsmessig gap.

Jo større dette gapet blir, jo mer ressurser vil kreves for å tette det. Nok ressurser fra start er derfor svært viktig når man går til anskaffelse av et slikt system. Derfor må kostnader til forvaltning av systemet synliggjøres sammen med kostnader til lisenser og eventuelt fysisk utstyr. Forvaltning vil innebære kursing og opplæring av forvaltere og driftspersonell, årsverk til oppfølging og drift, opplæring av brukere og liknende.

Det ble i delkapittelet om forvaltningsplan nevnt områder som var spesielt viktig å følge opp i forbindelse med bruk av KI-verktøy. I tillegg til dette er det spesielt viktig å ha avklart ansvars- og beslutningsmyndighet. Dette er avgjørende for å kunne ta løpende vurderinger på risiko, utvikling og eventuelt når en exit-strategi skal iverksettes.

Oppsummering
Kunstig intelligens vil gi stor gevinst hvis det tas i bruk på riktig måte, men konsekvensene av feil bruk kan også være store. Å ha en gjennomarbeidet strategi for bruk er viktig underlag for å kunne gjøre gode vurderinger når et nytt system eller verktøy skal anskaffes. Flere av vurderingene har likhetstrekk med sky-teknologi, men det er likevel noen områder som skiller seg ut eller blir spesielt viktig i forbindelse med kunstig intelligens:

- Avdekk relevant lovverk og krav som stilles
- Avdekk hvor KI kan brukes og IKKE skal brukes
- Avdekk ansvars- og beslutningsmyndighet
- Sett kriterier for exit og gjør det til en realistisk mulighet.
- Følg opp produktet som anskaffes

Anbefaling

- Les hele funnrapporten på https://www.ntnu.no/copilot

This project has received funding from the European Union's DIGITAL EUROPE programme, under Grant Agreement nº 101083966.