

Gjelder fra:	Gjelder for:	Godkjent dato:	Godkjent av:	Ansvarlig:
18.03.2024	Trøndelag fylkeskommune	18.03.2024	Fylkesdirektørens ledergruppe	Program for innovasjon og digitalisering

Retningslinjer for bruk av kunstig intelligens

Overordnede retningslinjer og prinsipper for bruk av kunstig intelligens i Trøndelag fylkeskommune

INNHOOLD

1	Formål og omfang	1
2	Begrepsforklaring	2
3	Prinsipper for bruk av kunstig intelligens i TRFK	2
4	Personvern.....	3
5	Vær varsom-prinsipper for bruk av generativ KI.....	3
6	Retningslinjer for bruk av generativ KI.....	4

1 FORMÅL OG OMFANG

Generativ Kunstig Intelligens (KI) representerer en kraftig teknologisk utvikling som har potensiale til å forme måten vi jobber, skaper og samhandler på og kan bidra til at fylkeskommunen kan hente ut gevinster. Formålet med disse retningslinjene er å veilede ansatte i utvikling og bruk av generativ KI som ivaretar våre organisatoriske verdier, lover og forskrifter, personvern, etiske retningslinjer, beskyttelse av virksomhetsdata og brukerbehov. Retningslinjene skal også bidra til å fremme tillit, åpenhet, ansvarlighet og etisk refleksjon rundt KI i fylkeskommunens virksomhet og tjenester.

Retningslinjer for bruk av kunstig intelligens er overordnet og setter minimumskrav til hvordan ansatte bruker generativ kunstig intelligens. Mer detaljerte retningslinjer for de ulike sektorene/tjenesteområdene må etableres og kontinuerlig oppdateres for å følge utviklingen av teknologiene.

Ved andre behov for integrering av KI-funksjonalitet i systemer og løsninger eller egenutvikling av KI må gjeldene retningslinjer i Trøndelag fylkeskommune følges. Mer informasjon finnes her:

<https://tronder.sharepoint.com/sites/intranett/ansatt/>

2 BEGREPSFORKLARING

Kunstig intelligens (KI): Hensikten er å oppnå et gitt mål. Kunstig intelligente systemer utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data. Det omtales tre dimensjoner innen KI:

- **Generativ KI:** Maskinlæringsmodeller som lager nytt innhold basert på informasjonen de har blitt trent på og instruksjonen du gir den. Dette innholdet kan være tekst, bilder, lyd og video.
- **Integrert KI:** KI-teknologier innlemmet inn i eksisterende systemer for å forbedre deres funksjonalitet, for eksempel Microsoft 365 Copilot i Word, Excel, PowerPoint og Teams.
- **Egenutviklet KI:** Fylkeskommunen utvikler selv et KI-system der målet er å løse eller forbedre en oppgave. Dette gjøres ved bruk av data, metoder, teknikker og teknologi.

Dele: I denne konteksten betyr «å dele» å gi informasjon til et KI-system. Dette kan være å skrive inn tekst, lime inn tekst, bilder, dokumenter eller annen informasjon fra andre systemer inn i et KI-system.

Instruks ("Prompt"): Enkle kommandoer man kan gi et kunstig intelligent system for å skape innhold som tekst, bilder, lyd og kode.

Virksomhetsdata: All informasjon som fylkeskommunen samler, lagrer og behandler. Dette kan inkludere data om innbyggere, ansatte, elever, transaksjoner, produkter, tjenester m.m.

Personopplysninger: Alle opplysninger og vurderinger som kan knyttes til en enkeltperson. Dette kan bidra til å identifisere denne enkeltpersonen, enten direkte eller indirekte. Personopplysninger kan være navn, bilde, adresse, telefonnummer, e-post, lydopptak og fødselsnummer. Biometri slik som fingeravtrykk, irismønstre og hodeform (for ansiktsgjenkjenning) er også personopplysninger.

DPIA: Data Protection Impact Assessment, på norsk kalt vurdering av personvernkonsekvenser ([Personvernforordningen Artikkel 35](#)). Dersom det er sannsynlig at en type behandling av personopplysninger vil medføre høy risiko for folks rettigheter og friheter, skal fylkeskommunen vurdere hvilke konsekvenser den planlagte behandlingen vil ha for personvernet. DPIA må gjøres før behandlingen gjennomføres og er av avgjørende betydning for om behandlingen kan skje i det hele tatt. Dette gjelder særlig ved bruk av ny teknologi. Mer informasjon finnes her: <https://tronder.sharepoint.com/sites/intranett/ansatt/Sider/Informasjonssikkerhet.aspx>

Nysgjerrig på kunstig intelligens? Lær det grunnleggende om KI i kurset "Elements of AI": <https://www.elementsofai.com/no/>

3 PRINSIPPER FOR BRUK AV KUNSTIG INTELLIGENS I TRFK

Trøndelag fylkeskommune skal bevare og styrke tilliten vi har som offentlig forvaltningsorgan. For å ivareta denne tilliten og sørge for ansvarlig bruk av KI i tjenesteproduksjonen vår, må vi følge noen grunnleggende prinsipper:

Lovlig – overholde alle gjeldende lover og bestemmelser

Ansvarlig bruk av KI betyr å ivareta rettighetene til ansatte og innbyggere. Grunnloven og menneskerettighetskonvensjonen sier noe om retten til privatliv, ytringsfrihet og likebehandling. Å ta i bruk KI i tjenesteproduksjonen kan utfordre oss på disse områdene.

Eksempler på nasjonale lover med betydning for bruk av KI:

- Personopplysningsloven med forordning
- Likestillings- og diskrimineringsloven
- Forvaltningsloven

Etisk – overholde etiske prinsipper og verdier

Etiske vurderinger handler om å vurdere positive og negative sider ved ny teknologi, spesielt hvis den benyttes i profilering eller automatiske avgjørelser med betydning for fysiske personer eller samfunnet. Det er viktig at KI ikke forsterker skjevheter vi allerede har i samfunnet, og at vi som er ansvarlige for løsninger med KI vurderer dette grundig.

Personvernregelverket har rettferdighet, gjennomsiktighet og forklarbarhet som viktige etiske prinsipper. Det betyr for eksempel at den det gjelder skal ha mulighet til innsikt i hvordan en automatisert beslutning tas eller at man får informasjon om at det er en maskin man kommuniserer med og ikke et menneske.

Sikker – i et teknisk og samfunnsmessig perspektiv

KI skal være bygget på systemer med teknisk robuste løsninger som forebygger risiko. På samme måte som ved alle IT-systemer, er det nødvendig med en systematisk og helhetlig tilnærming til informasjonssikkerhet før et system med KI settes i drift.

4 PERSONVERN

Det skal ikke behandles personopplysninger i løsninger med KI før det er gjennomført en risikovurdering og en vurdering av hvilke personvernkonsekvenser dette kan få (DPIA). Dette gjelder alle personopplysninger, ikke bare personopplysninger man anser som sensitive. Personvernombudet skal involveres når det gjøres en vurdering av personvernkonsekvenser (DPIA).

Personvernregelverket gjelder for bruk av KI på lik linje med all annen behandling av personopplysninger.

Mulige risikoer ved KI og behandling av personopplysninger:

- Personopplysninger blir behandlet til et ugyldig formål (f.eks. til treningsdata)
- TRFK mangler rettslig grunnlag for å behandle personopplysningene
- Personopplysninger blir behandlet av underleverandør uten gyldig databehandleravtale
- Personopplysningene behandles utenfor EU/EØS uten gyldig overføringsgrunnlag. Det er egne regler for behandling av personopplysninger utenfor EU/EØS.
- De registrerte får ikke oppfylt sine rettigheter (innsyn, retting, sletting, m.m.) Det er fylkeskommunens ansvar å legge til rette for at dette skjer.

5 VÆR VARSOM-PRINSIPPER FOR BRUK AV GENERATIV KI

Husk:

- at KI er en maskin, ikke et menneske – selv om det kan føles slik.
- at KI kan gi opplysninger som ser riktig ut, men som kan være helt feil.
- at KI finner sammenhenger mellom ulike faktorer i en stor mengde data. En sammenheng betyr ikke at det er et årsak-virkning-forhold.
- at KI sine svar ikke nødvendigvis er tilpasset våre verdier.
- at hvis du bruker KI-systemer i jobbsammenheng, vær klar over at informasjon om deg og det du skriver inn blir brukt til å trene opp KI-modellen. Det glemmes ikke.
- at opplæring og forståelse for hvordan KI kan brukes på en forsvarlig og effektiv måte er viktig.
- at hvis du er usikker på om det du ønsker å gjøre er greit, spør et menneske!



6 RETNINGSLINJER FOR BRUK AV GENERATIV KI

Retningslinjene under er basert på Digitaliseringsdirektoratets retningslinjer for generativ KI i offentlig sektor.

Du skal aldri:

- dele personopplysninger om ansatte, elever eller innbyggere.
- dele taushetsbelagt informasjon, gradert informasjon, informasjon unntatt offentligheten.
- dele brukerinformasjon, som påloggingsinformasjon og passord.
- stole på at KI sier sannheten.
- bruke KI til å ta beslutninger som i betydelig grad påvirker ansatte, elever, innbyggere eller bedrifter.

Du skal alltid:

- sjekke om det fins egne retningslinjer i Trøndelag fylkeskommune for det spesifikke KI-systemet du ønsker å bruke.
- vise aktsomhet ved bruk av virksomhetsdata.
- være kildekritisk til det du leser.
- være kritisk til om KI bryter opphavsrett, særlig for tredjepart.
- være nøye med å merke innhold som du har brukt generativ KI til å lage. F.eks. "Dette bilde er laget ved hjelp av KI", i tillegg til referanse.
- være transparent og si ifra når du bruker KI-generert tekst, bilde, kode eller annet.
- melde avvik ved brudd på disse retningslinjene.

Som leder:

- har du ansvar for opplæring og at ansatte har nødvendig kompetanse til å bruke generativ KI forsvarlig og effektivt.
- har du ansvar for at retningslinjene er kjent blant dine ansatte og at brudd på retningslinjene blir meldt som avvik.