

# Retningslinjer for bruk av språkmodeller i VKM

Alt innhold i VKM publikasjoner skapes av våre medlemmer, men språkmodeller kan benyttes til for eksempel oversettelse, korrektur, dataanalyse, utkast eller sammendrag av innhold. Innhold generert av språkmodeller vurderes og godkjennes alltid av faggruppens medlemmer før publisering.

En **språkmodell** er en type maskinlæringsmodell som er trent til å forstå, lese, skrive og bearbeide tekst på et menneskelig språk. Språkmodeller bruker store mengder tekstdata for å lære seg mønstre og sammenhenger i språket. Slike modeller er ofte basert på dype nevralt nettverk, spesielt transformer-arkitekturer, som gjør dem i stand til å håndtere komplekse språklige oppgaver.

## Formål

VKM ønsker å bruke språkmodeller som et verktøy for å effektivisere og forbedre kvaliteten på sine publikasjoner. All tekst generert av språkmodeller skal vurderes og gjennomgås av et menneske, og uredigert tekst fra generativ kunstig intelligens skal ikke publiseres. Eksempler på mulige anvendelser er korrektur, oversettelse og sammendrag, samt uttrekk av data og annen informasjon. Eksempler på ikke-tillatte anvendelser er tekstskriving som ikke er direkte basert på tekst skrevet av prosjektmedlemmene, samt bruk av åpne spørsmål uten kontekst (noe som gir unøyaktige eller feil svar) for å skrive tekster, og ikke bare for å finne kilder til informasjon.

## Omfang og målsetting

Gjennom disse retningslinjene søker VKM å sikre åpenhet i sine prosesser, slik at andre kan gjenskape arbeidet best mulig. VKM ønsker å opprettholde transparens og bidra til økt tillit til prosessene ved å gjøre metodologiene synlige og følge prinsipper for ansvarlig bruk av kunstig intelligens, slik Digitaliseringsdirektoratet anbefaler (se veiledning: <https://www.digdir.no/kunstig-intelligens/kunstig-intelligens/4132>).

## Kvalitetssikring av innhold

Språkmodellen kan bidra med forslag og utkast, men det er forfatterens ansvar å sikre kvaliteten på det endelige innholdet. All tekst generert av språkmodeller må merkes og synliggjøres tydelig i utarbeidelse av utkast slik at den kan vurderes og kvalitetssikres av fagpersoner. Ved publisering av arbeider som involverer språkmodeller, skal metodikk og verktøy oppgis i methodedelen.

## Datasikkerhet og konfidensialitet

Nettbaserte språkmodeller som ChatGPT, skal ikke brukes til å behandle sensitiv eller konfidensiell informasjon, inkludert personopplysninger eller andre typer sensitiv/klausulert data. Slik data bør kun håndteres av språkmodellimplementeringer som er installert på den lokale PC-en og isolert fra internett.

## Transparens og forklarbarhet

All tekst generert eller bearbeidet av språkmodeller skal omtales og siteres i metodekapittelet i henhold til vitenskapelige standarder for å unngå plagiat. Dette inkluderer genererte oversettelser, sammendrag og annen bearbeiding gjort av språkmodeller.

## Sitering

Tekst generert av en modell skal merkes i førsteutkastet for å sikre at den gjennomgår bearbeiding før eventuell publisering. Etter bearbeiding og integrering av tekst fra en språkmodell, skal det fremgå i metodebeskrivelsen at førsteutkastet av for eksempel et sammendrag er basert på en språkmodell.

Når det genereres sammendrag eller andre hovedpoeng ved bruk av en språkmodell, bør kilden oppgis slik:

- Eksempel: Sammendraget er laget med ChatGPT (2024), og innholdet er kvalitetssikret av VKM.  
OpenAI ChatGPT (25.09.2024) [4o-mini]. Hentet fra <https://chat.openai.com>

### Direkte sitat (klipp og lim)

Om man unntaksvis siterer tekst direkte fra en språkmodell, bør man bruke anførselstegn. En korrekt sitering bør inkludere modellens navn, brukstidspunkt og URL til kilden:

- “Direkte sitert tekst her” OpenAI ChatGPT (25.09.2024) [o1 preview]. Hentet fra <https://chat.openai.com>

### Bruk av “Application programming interface” (API)

API-løsninger muliggjør at datamaskiner kan kommunisere med hverandre via programmering. Ved bruk av språkmodeller via API-er for datahåndtering i VKM-arbeid, skal dette dokumenteres og beskrives i metodekapittelet. Data og script skal lagres i henhold til VKMs rutiner for datahåndtering.

### Gråsoner

Språkmodeller brukes i flere vanlige verktøy, eksempelvis Word eller Grammarly, eller i vanlige oversettelses program som Google-oversettelse. Mange verktøy anses som støttefunksjoner og regnes som en naturlig del av skriveprosessen og krever ikke eksplisitt dokumentasjon i metodekapittelet eller referanser, med mindre de brukes til mer omfattende og vesentlig endringer av innholdet.