

1. notat fra rådgivende ekspertgruppe for KI-satsingen: Nåsituasjon og forventninger

KI (Kunstig intelligens) og digitalisering treffer samfunnet vårt enormt bredt. Digitalisering er avgjørende for å møte sentrale samfunnsutfordringer rundt økonomi, demografiske endringer, kompetansebehov, klimakrisen, global helse, ulikhet, økt polarisering, digitalt utenforskap og tillit. Samtidig vil bruk av de nye teknologiene ha store og til dels uoverskuelige samfunnskonsekvenser.

KI-milliarden alene vil gi begrensede muligheter til å løse alle utfordringene. En egen satsing på KI vil imidlertid kunne sette en retning for forskningsinnsatsen, som igjen kan følges opp med andre forsknings og innovasjonsprosjekter, og nye politiske grep. Samlet kan dette gjøre Norge til en aktiv premissleverandør i den teknologiske utviklingen.

Som beskrevet i mandatet¹ innbefatter satsingen **tre integrerte spor**

1. Konsekvenser av kunstig intelligens og digital teknologi
2. Digitale teknologier som forskningsområde i seg selv
3. Hvordan digitale teknologier kan brukes til innovasjon inkludert innen forskning

‘Digitale teknologier’ er et *enormt* bredt område. Vi har valgt å tolke mandatet til ikke å omfatte alle digitale teknologier, men begrenset dette til KI og digitale teknologier som er relatert til eller muliggjør KI. Dette omtales i det videre samlet som KI. Med dette som utgangspunkt har vi sett på **forskning** innen de tre sporene.

Vi legger til grunn at mer sektororienterte anvendelser og innovasjonsprosjekter primært finansieres innen andre programmer eller fra andre kilder. Store infrastrukturer for generative språkmodeller, data og prosesseringsinfrastruktur for maskinlæring og andre KI-metoder, er også utenfor vårt mandat. Samtidig vil slike prosjekter og infrastrukturer være viktige for at enkelte av effektene av forskningen skal kunne materialisere seg.

Satsingens tre spor henger tett sammen. Sammenhengen mellom sporene understrekes av at en del tema vil være viktige i forskning innen alle sporene i satsingen. Vi vil trekke fram tre eksempler på slike **tverrgående tema**:

- Tillit: I samfunnet, til teknologien, og til at teknologien og tjenester som lages ved hjelp av den er sikre.
- Ansvarlighet: Ansvarlig, inkludert sikker og ikke-diskriminerende teknologi, og ansvarlig og bærekraftig teknologibruk
- Kreativitet: Nytenkende og banebrytende forskning. Nyskapende utvikling og bruk av teknologi. Kreative KI-baserte problemløsningsmetoder for innovasjon og nyskaping.

For å dra nytte av koblingene mellom de tre sporene trengs **tverrfaglig forskning**, som kombinerer teknologifag, samfunnsfag, jus og humaniora. Banebrytende forskning er imidlertid ikke alltid tverrfaglig. Vi trenger også **spisset forskning** fra mange forskjellige fagdisipliner, og virkemidler som sikrer at kunnskap fra disse er tilgjengelig på tvers av fag og spor. Dette vil være relevant innen alle sporene.

¹ Fullstendig mandat i vedlegg

For at satsingen skal ha effekt og få et tydelig fotavtrykk må den ha en tydelig retning. Samtidig må satsingen være tilstrekkelig åpen for muligheter og problemstillinger vi enda ikke kjenner. Vi vil under overordnet beskrive status og muligheter innen de tre sporene.

1 Konsekvenser av kunstig intelligens

1.1 Områder med behov for forskning

Kunstig intelligens tas i bruk innen så godt som alle samfunnsfelt og listen med forskningsbehov er uuttømmelig. Dette gjelder både forskning på KI som er tatt i bruk, følgeforskning når KI tas i bruk, og forskning på mekanismer som er relevante når ny teknologi skal tas i bruk. Vi har derfor identifisert noen overordna tema med eksempler på undertema. Vi understreker at dette ikke er en fullstendig oversikt, og at det vil dukke opp nye områder i 5-års perioden.

Tillit i samfunnet

De nordiske samfunnene er preget av en høy grad av tillit. Dette gjelder både mellommenneskelig tillit, tillit til samfunnsinstitusjonene og forvaltningen, men også tillit til teknologi. Det er viktig å få mer kunnskap om hvordan kunstig intelligens og bruk av kunstig intelligens påvirker tillit (og grunnlaget for tillit). Eksempler på undertema er sikkerhet, personvern og overvåking, informasjon og desinformasjon, ytringsfrihet, demokratiske prosesser, påvirkning av valg og valgoppslutning, KI-støttede beslutningssystemer, etikk, og reguleringer og rammer for teknologibruk.

Rettferdighet, mangfold, språk og kultur

KI reproducerer skjevheter i treningsdatasettene. Generativ KI kan gi dette problemet i enda større skala, ikke bare i form av diskriminering på grunnlag av f.eks. kjønn, hudfarge eller bosted, men også ved at norske data i liten grad er med i treningsdatasettene for store modeller, enten det gjelder språk og bilder, eller medisinske data. Vi trenger forskning som avdekker slik diskriminering, og på hvordan KI påvirker norsk og samisk språk og kultur. Samtidig trenger vi humanistisk og samfunnsvitenskapelig forskning som kan sikre at etiske forutsetninger bygges inn i teknologien og ikke bare blir en ettertanke.

Arbeids- og næringsliv

Arbeidslivet endres med innføring av ny teknologi, og kunstig intelligens kan bidra til store endringer i arbeidsmarkedet og innholdet i ulike yrker. Eksempler på undertema er teknologiens innvirkning på jobbskaping og -tap, endringer av profesjoner, av arbeidsprosesser, ledelse, forretningsmodeller, kompetansebehov i næringsliv, helse og offentlig sektor, og behov for etter- og videreutdanning, samt bærekraftig bruk av KI, og hvilke krav som bør stilles til transparenens.

Kunnskap og kreativitet

Teknologien muliggjør nye måter å lære på og å uttrykke seg på. Kunstig intelligens kan brukes til å innhente og bearbeide informasjon, til store analyser, og til skapende og kreative prosesser. Det er i dag betydelig diskusjon om bruk av digital teknologi i skolen. Kunstig intelligens vil gi økte utfordringer, men også gi nye muligheter. Eksempler på undertema er den enkeltes behov for å mestre teknologien, nye muligheter innen kunst og kultur, kreativitet, samt kunstnerisk utvikling.

1.2 Fag hvor Norge og norske forskningsmiljøer har fortrinn

Norge har siden 1990-tallet hatt betydelige forskningssatsinger på IKT og digitalisering. Det er bygd opp fagmiljøer med sterk kompetanse på teknologiens betydning for samfunn og kultur, for offentligheten, demokratiet, arbeidslivet, og utdanningene. Og vi har betydelig forskning på regulering av teknologi, samt hvordan teknologibruk, brukermedvirkning og kultur bidrar til å forme teknologien. Det er sterke tradisjoner for tverrfaglige samarbeid mellom teknologiske fag og humanistiske og samfunnsvitenskapelige fag i Norge. Mange av disse fagmiljøene er helt i front internasjonalt på sine områder. Disse fagmiljøene orienterer seg nå mot kunstig intelligens, og Norge har her mulighet til å befeste sin posisjon internasjonalt på forskning på konsekvenser av kunstig intelligens.

2 Kunstig intelligens som forskningsområde i seg selv

1.2 Områder med behov for forskning

KI er teknologi i rask utvikling. Mye av teknologien som benyttes i dag er umoden og lite transparent. Den kan mangle robusthet, har utilstrekkelig kvantifisert usikkerhet, klarer ofte ikke å generalisere på tvers av domener, er avhengig av enorme mengder kuraterte treningsdata, og er energikrevende i både trening og bruk. Samtidig ser vi at teknologiens potensiale til å bidra til nødvendig og viktig innovasjon og utvikling er stort (jmf del 3). Det er derfor avgjørende at teknologien videreutvikles. Forskningsinnsatsen bør bidra til en radikal transformasjon av KI for å gjøre den (1) mer nøyaktig, (2) ta høyde for usikkerhet, (3) ansvarlig og forklarbar, og (4) mer bærekraftig.

Mer nøyaktig

Dersom vi skal ha grunn til å ha tillit til kunstig intelligens må den i flere sammenhenger være tilstrekkelig nøyaktig og gi svar vi har grunn til å stole på. Det gjelder om den skal brukes til beslutningsstøtte, analyser eller inn mot andre teknologier. Sentrale forskningsområder for mer pålitelig og nøyaktig kunstig intelligens er: Data-kuratering, inkludert gode datagrunnlag for maskinlæring, metode for å omvandle data til informasjon, for eksempel ved å kombinere 'semantiske' metoder eller annen kunnskap (fysikk og matematikk) med maskinlæringsmetoder og statistiske modeller, samt mulighet for måling av modellkvalitet gjennom transparent bruk av meta-data.

Presisjon i KI og det "å stole på" modellen vil variere etter bruksområde. For KI-generert kunst, for eksempel, kan presisjon bety kreativitet og evnen til å følge en bestemt kunstnerisk stil, men også evnen modellene har for å støtte opp om kreativitet og eksperimentering. I prediksjoner, som presisjonsmedisin, vær- eller markedsanalyser, refererer presisjon til nøyaktigheten av forutsigelser sammenlignet med faktiske resultater.

Tar høyde for usikkerhet

Data er oftest ufullstendige og støybefengte (noisy); modeller er oftest ikke nøyaktige nok, og algoritmer bare tilnærmer eksakte løsninger. Derfor er estimater, generaliseringer, prediksjoner og beslutninger produsert av KI i seg selv usikre. For å gi den tilliten som trengs for å informere beslutninger, må usikkerhet om resultater kvantifiseres. Det vil igjen øke påliteligheten til KI-baserte anbefalinger. Det er viktig også at kvantifisering av usikkerhet avdekker uenighet mellom data og kunnskap, så vel som mellom ulike datakilder. Relevante forskningsområder er: usikkerhetskvantifisering, statistikk, robusthetsgarantier, modeller med harde og myke begrensninger, feilmarginer for autonome systemer, brukertolkning, digital humaniora (f.eks. *critical dataset studies*), og teknologifilosofi.

Ansvarlig og forklarbar

Kunstig intelligens som implementeres i samfunnet må i langt større grad enn i dag være transparent og forklarbar. Vi må vite mer om hvordan den faktisk fungerer (datagrunnlag, modeller, algoritmer og epistemologi) hva den kan brukes til, og hva den ikke fungerer godt til. KI må fungere sikkert, samt gjenspeile en norsk kontekst med norske verdier. For å få til dette trengs også høy grad av brukermedvirkning i utviklingen av teknologien. Viktige forskningsområder er for eksempel: Forklarbarhet, sikkerhetsutfordringer, språkteknologi for norsk tekst og tale, alignment av modeller, syntetisk datagenerering og personvern, normer og verdier, diskriminering og bias, og analyser av sammenhenger mellom treningsdata og output.

Mer bærekraftig

Kunstig intelligens krever enorme mengder energi for å være effektiv. Mye av arbeidet innen KI og maskinlæring spesielt bruker brute-force med fokus på stadig større modeller med stadig mer data. Å få omtrent like gode resultater med mindre modeller og databruk er sentralt for at bruk av KI skal være mer bærekraftig. Sentrale forskningsområder for å lykkes med dette er blant annet tinyML, kunnskapsbasert ML, energieffektiv prosessering inkludert nye beregningsmetoder (for eksempel NeuroAI og nevro-morfisk databehandling). Fordi lagring av data også bruker energi, trengs det nye metoder for å redusere og komprimere data, uten større tap av informasjon.

1.3 Fag hvor Norge og norske forskningsmiljøer har fortrinn

Norge har sterke fagmiljøer innen digitale teknologier og kunstig intelligens. Noen områder der Norge har satset over tid er maskinlæring, teoretisk informatikk, statistikk, logikk, anvendt matematikk, energi-effektiv prosessering inkludert nye beregningsmetoder, bildebehandling, teori om læring. Her er det også muligheter for å gjøre gjennombrudd i tverrfaglig samarbeid med humanistiske, samfunnsvitenskapelige og juridiske fag, og domenespesifikke problemer innen biomedisin, fysikk og elektronikk. Et annet område med lang tradisjon er språkteknologi som er viktig for å utvikle og forstå fremtidens språkmodeller. Cybersikkerhetsmiljøet har lang erfaring med å håndtere komplekse data og systemer. Både maskinlæring og mer tradisjonell symbolsk KI fordrer strukturert forståelse og oppfølging av data, der Norge også har sterke miljøer. Av nye områder ser vi en fremvekst av miljøer som jobber med forklarbar KI (XAI), kunnskapsbasert KI (inkl. hybrid-modeller), samt NeuroAI, som springer ut fra det sterke nevrovitenskapsmiljøet i Norge.

3 Hvordan kan kunstig intelligens brukes til innovasjon, inkludert innen forskning

3.1 Områder med behov for forskning

Det er stort potensiale for å bruke kunstig intelligens til innovasjon i privat og offentlig sektor inkludert innen forskning. Vi vil illustrere forskningsbehovet med å vise noen eksempler på muligheter innen disse sektorene. Deretter vil vi gi noen eksempler på tverrgående behov.

For innovasjon i næringslivet – grønn og digital omstilling

Innovasjon i næringslivet er avgjørende blant annet for at Norge skal få til grønn og digital omstilling. KI vil være nødvendig for omstilling i en lang rekke næringer, fra havbruk og maritime næringer, til prosessindustri, energiproduksjon og distribusjon, bygg og anlegg, batteri/hydrogen, transport, telekom, logistikk, medier og kreative virksomheter. Det er nødvendig med forskning om hvordan KI kan brukes i innovasjonsprosesser for å redusere energibruk, materialbruk, forurensing eller for å skape nye verdikjeder (for eksempel basert på sirkulærøkonomi), tjenester og produkter. Videre kan forskningsbasert innsikt i hvordan KI kan gi innovasjon bidra til viktig omstilling av arbeidsflyt, organisasjon og prosessendringer.

For innovasjon i offentlig sektor – helse og velferd

Offentlig sektor har et svært stort innovasjonsbehov og KI og digitale teknologier kan spille en viktig rolle i dette. Helsepersonellkommissjonen (NOU 2023: 4) illustrerer dette svært godt. Behovet for helse- og omsorgstjenester øker som følge av demografiske endringer. På mange områder, fra diagnostikk til rehabilitering og oppfølging, kan KI gi vesentlige gevinster og bedre tjenester. Behovet er stort, og her trengs både følgeforskning og videreutvikling av KI og andre digitale teknologier som også tar høyde for domenekunnskap og konteksten teknologien skal brukes i. Tilsvarende eksempler kan gis for andre deler av offentlig forvaltning.

For forskning

KI har stort potensiale for gjennombrudd både innen medisin, naturfagene, humaniora og samfunnsvitenskapen. Selv om det er innen naturfaglig forskning vi ser de tydeligste resultatene så langt (for eksempel AlphaFold), er KI på full fart inn i forskning i alle fagområder. Aspekter knyttet til datakuratering og prosesstøtte er sentralt for å øke replikerbarheten av forskningen. Det trengs bedre forståelse av sammenhengen mellom treningsdata og resultat, både for støyete og mangelfulle data, og for ustrukturerte data som tekst og bilder. Det er viktig å utvikle bedre forståelse av hvilke metoder som fungerer på hvilke problemer og typer data. Det er mange eksempler på at det er brukt maskinlæring på feil måter, uten at det er en god tradisjon for å spre denne kunnskapen². Språkmodeller og det at KI er blitt så lett å ta i bruk øker risikoen for slike feil.

Tverrgående behov

For at potensialet med å bruke KI til innovasjon i næringslivet, offentlig sektor og forskning skal kunne realiseres er de tverrgående temaene nevnt innledningsvis viktige: Vi må ha tillit til teknologien og hvordan

² OECD 2023: Artificial Intelligence in Science. <https://doi.org/10.1787/a8d820bd-en>

den utvikles og tas i bruk. Teknologien må være og brukes ansvarlig og ivareta sikkerhet på en god måte. Og vi må bruke teknologien sånn at den fremmer kreativ problemløsning. Videre må teknologien, slik det ble redegjort for i del 2, være nøyaktig, bærekraftig, ansvarlig og forklarbar, og ta høyde for usikkerhet. I tillegg er det avgjørende med transparens rundt prosesser, som hvilken rolle teknologien spiller i beslutnings- eller produksjonsprosesser. KI blir i økende grad del av et sentralt interaksjonssystem der menneske samspiller med maskin i stadig nye kontekster for blant annet beslutningstøtte, kunst, kreativitet, helse og sosiale relasjoner, som gir behov for å utvikle kunnskapen innen systemutvikling og informasjonssystemer for hvordan ta i bruk ny teknologi i organisert virksomhet.

3.2 Fag hvor Norge og norske forskningsmiljøer har fortrinn

Norge har sterke forskningsmiljøer som er godt egnet til å realisere behovene for innovasjon i offentlig sektor og næringsliv nevnt over. I tillegg til det som er nevnt i del 1 og 2 av notatet vil vi trekke fram særlig to forhold: For det første, krever vellykket innovasjonsarbeid brukerinvolvering. Norge har fagmiljøer med sterk kompetanse på å inkludere mennesker i utvikling av teknologi. For det andre er vellykket innovasjon avhengig av å tenke nytt om arbeidsprosesser og forretningsmodeller. Norge har verdensledende miljøer innen systemutvikling og forretningsmessig verdi av IT.

De anvendelsesområdene det er særlig viktig for Norge å utvikle kan videre deles i to hovedkategorier. Først har vi de områdene der Norge enten allerede har et næringsliv som bør kunne trekke vekslers på forskningen, eller der hvor det er grunn til å tro at et slikt næringsliv kan utvikles. Helse, ren energi, havbruk og maritime næringer er eksempler på dette. I tillegg til dette finnes det områder der vi ikke nødvendigvis har fortrinn, men der det er avgjørende for Norge at vi utvikler egen kunnskap og kompetanse. Norge har behov for autonomi innen digital sikkerhet. Offentlig sektor og norsk språkstøtte er områder som er viktige uten at vi har fortrinn internasjonalt. Ressursoptimalisering, transport/logistikk og intermodalitet er andre eksempler på slike områder.

4 Mål for satsingen

Satsningen sammen med tilleggsaktiviteter skal resultere i:

- **Forskningstyngde:** Norge plasseres tydelig på kartet som en forskningsnasjon med tung KI-kompetanse. Forskning og nye metoder bidrar til å håndtere utfordringer med avansert KI, i lys av nordiske verdier.
- **Samarbeid på tvers:** Øke antall tverrfaglige prosjekter og samarbeid som krysser sektorer og disipliner for å adressere komplekse KI-relaterte problemstillinger.
- **Kompetanseutvikling:** Resultater fra forskningen og anvendelse tas inn i undervisningen, både av teknologer, men også innen alle andre felt.
- **Kunnskapsmiljøer:** Det utvikles samarbeid mellom forskning og praksisfelt som tilrettelegger for god anvendelse av forskningsresultater.
- **Sikkerhet:** Styrke digital sikkerhet, inkludert motstandsdyktighetstester i KI-systemer.
- **Innovasjon:** Implementering av KI-baserte løsninger i næringslivet, helse og offentlig sektor.
- **Folkeopplysning:** Basis for informative ressurser for allmennheten som øker bevissthet om KI-konsekvenser.
- **Miljø og bærekraft:** Konkrete eksempler på hvordan KI har bidratt til å løse miljøproblemer eller forbedre forvaltningen av naturressurser.

5 Fotavtrykk av satsingen i 2030

Vi beskriver her et ganske ambisiøst fotavtrykk. Dette vil ikke komme av KI-milliarden alene, men ved at KI-milliarden setter en retning og brukes til å opparbeide forskningsbasert kunnskap som kan trekke med seg annen finansiering fra offentlige og private aktører i samme retning. Ambisjonen for bruk av milliarder til kunstig intelligens må være langsiktig, fruktbar, banebrytende og tverrfaglig:

1. Langsiktig: Effekten av milliarden går langt utover de neste 5 årene og selve milliarden, og er i den sammenheng økonomisk bærekraftig.
2. Fruktbar: Effekten av milliarden gir gevinster langt utenfor forskningen. I tillegg til forskningsresultater, bidrar forskningen til konstruktiv og regulert anvendelse av KI spesielt innen helse/velferd, offentlig sektor, og til grønn omstilling i næringslivet.
3. Banebrytende: KI-milliarden produserer viktig, ambisiøs og innflytelsesrik forskning som prioriterer kjerneutfordringene og plasserer Norge på kartet over KI-forskningstunge land. Banebrytende forskning skjer i internasjonalt samarbeid og norske forskere er med på å løse en grand challenge innen vitenskapen med bruk av KI.
4. Tverrfaglig: Resultatene av banebrytende KI-forskning må være til nytte for samfunnet og mennesker generelt, og bidra til en mer rettferdig og bærekraftig verden. Dette kan kun oppnås gjennom tverrfaglig forskning.

Totalt skal satsingen, sammen med tilgrensende prosjekter, plassere Norge som en sentral KI-nasjon. På utvalgte områder skal Norge være anerkjent som en drivkraft på nordisk og europeisk nivå, og gjennom dette på resten av verden. Dette skal gjøre Norge mer attraktivt for internasjonale KI-forskere, både som arbeidsted og samarbeidspartner i internasjonale prosjekter. En større anerkjennelse skal også bidra til at vi har en politisk påvirkning på retningen på KI også internasjonalt.

Avlevert til Norges forskningsråd 15. november 2023

Ekspertgruppens medlemmer:

Tanja Storsul (leder), Arnoldo Frigessi, Cathrine Pihl Lyngstad, Eirik Andreassen, Ieva Martinkenaite, Inga Strümke, Jill W. Rettberg, John Krogstie, Klas Pettersen, Olav Lysne, Petter Bae Brandtzæg, Signe Riemer-Sørensen og Tobias Mahler.

Vedlegg: Mandat for rådgivende ekspertgruppe