

Policy Brief 01.07.2022

# Framtidslaboratorium om kunstig intelligens

Lina Ingeborgrud, Per Koch og Silje Marie Svartefoss

**I dette notatet gir vi en kort oppsummering fra et halvdags fremtidslaboratorium om kunstig intelligens arrangert innenfor rammene av AFINO-senteret for ansvarlig forskning og innovasjon. Øvelsen ble gjennomført den 4. april 2022 i samarbeid mellom NIFU og Handelshøgskolen BI.<sup>1</sup> Målet var å drøfte fremtidige muligheter og utfordringer som kan følge av bruk av kunstig intelligens i Norge og verden for øvrig. AI gir store muligheter for innovasjon på nesten alle samfunnsområder, men gir også utfordringer på områder som personlig frihet, demokrati, rettferdighet og sosial inkludering.**

I dette notatet gir vi en kort oppsummering fra et framtidslaboratorium om kunstig intelligens. Øvelsen ble gjennomført den 4. april 2022, i samarbeid mellom NIFU og BI. Målet var å utforske ulike fortellinger om, og å drøfte fremtidige muligheter og utfordringer for utvikling og bruk av kunstig intelligens, både nasjonalt og internasjonalt.

Algoritmer og kunstig intelligens har allerede blitt en sentral del av livene våre, og brukes for å ta beslutninger og tilby tjenester i arbeidslivet, utdannings-systemet, byer og medisinsk behandling, så vel som i finansmarkeder og sosiale medier. Likevel er kunstig intelligens bygget på lukkede modeller som gjerne er vanskelige å forstå for brukerne. De belager seg på store datasett som kan være mangelfulle, og ikke minst bidra til å forsterke ulike diskrimineringsmønstre. Hvordan kan sosiale, kommersielle og politiske aktører både lokalt og globalt sikre ansvarlighet i algoritmiske beslutningsprosesser? Det er behov for en bred samfunnsdebatt om framtidens bruk av kunstig intelligens.

Framtidslaboratoriet ble gjennomført som en del av forskningsrådsprosjektet 'Algorithmic Accountability: Designing Governance for Responsible Digital Transformations', ledet av Nordisk Senter for Internett og Samfunn ved BI, og som en del av AFINO, senteret for ansvarlig forskning og innovasjon i Norge.

## Kort om metoden

Framtidslaboratoriet ble gjennomført som en gruppeøvelse der deltakerne ble bedt om å utvikle ulike fortellinger om hva fremtiden kan bringe når det gjelder utvikling og bruk av kunstig intelligens. Poenget var ikke å forutsi fremtiden, men å utfordre deltakerne vanetenkning og fordommer.

Øvelsen tok utgangspunkt i en bestemt tilnærming for fremtidensteking utviklet og tatt i bruk av UNESCOs nettverk for *Futures Literacy* eller fremtidskyndighet på norsk. Å være fremtidskyndig betyr å være oppmerksom på at våre valg og handlinger i nåtiden i stor grad er formet av våre antakelser om og forventninger til fremtiden.

I praksis bestod øvelsen av tre hovedsesjoner. De påmeldte ble delt inn i fire grupper med 5-6 deltakere hver. Gruppedeltakerne ble bedt om å skrive ned sine ideer om 2060 på post-it lapper, i form av sitater, avisoverskrifter, beskrivelser eller metaforer. Alle ble deretter bedt om å presentere ideene sine for gruppa, før gruppelederen limte dem opp på veggen, sortert etter tema. Forslagene ble diskutert og vurdert. Arbeidet i gruppene ble ledet av forskere ved NIFU og BI.

Under den første sesjonen ble deltakerne bedt om å beskrive *håpene* de hadde for hvilken rolle kunstig intelligens ville spille i fremtiden. Hensikten var her å kartlegge de bakenforliggende drømmene og verdi-

ene til deltakerne. Deretter brukte deltakerne samme metode på å beskrive de realistiske forventningene de hadde til dette i framtiden. Her er målet å avsløre ulike forforestillinger som kan være med på å begrense folks mulighet til å takle det uforutsette og komme opp med konstruktive ideer som bryter med eksisterende tankemønstre, etablert praksis og institusjonelle strukturer.

Deltakerne ble ikke bedt om å utvikle felles scenarier. Det kan man gjøre når man bruker denne metoden, men her hadde vi for liten tid til en slik øvelse. I stedet bidro deltakerne med ulike visjoner og ideer om hva framtiden kan bringe. Til sammen bidrar dette til et idemangfold som beriker debatten, hjelper deltakerne til å lære av hverandre og som åpner opp for nye muligheter.

Etter den andre sesjonen presenterte imidlertid gruppelederne et *mot-scenario*. Slike motscenarier (*reframing*) skal stimulere til mer radikal nytenking. Dette gjør man gjennom å presentere en virkelighet som på ulike måter bryter med dagens verden. Man kan ta bort institusjonelle strukturer og praksiser som synes selvfølgelig i dag. På den måte tvinges deltakerne til å komme opp med andre ideer, utfordringer og muligheter. Scenariet brukt i denne øvelsen beskrev et 2060 der de store IKT- og medieselskapene hadde fått fritt spillerom globalt, og med enorm makt til å forme verdenssamfunnet etter egne ønsker og behov. I denne beskrivelsen var nasjonalstaten sterkt svekket, og internasjonale organisasjoner viste liten evne til å kontrollere, styre eller begrense bruken av kunstig intelligens på nye områder.

## Sentrale budskap fra gruppediskusjonene

### Håp

Under den første sesjonen skrev deltakerne lapper som reflekterte deres drømmer og håp for framtiden. Formålet var å få fram deres bakenforliggende drømmer for bruk av kunstig intelligens (*artificial intelligence*, AI), og hvordan kunstig intelligens preger tiden i 2060.

Noen uttrykte at håpsesjonen var den enkleste å gjennomføre. Gruppene presenterte en bredde i overordnede tema fra mat, klima, sivilsamfunn, helse, teknologi og energi. Flere/alle gruppene beskrev en verden der kunstig intelligens bidrar til løsninger på store samfunnsutfordringer så vel som i hverdagslivet, og flere trakk fram at kunstig intelligens gjør arbeidslivet enklere: «Monotone og belastende arbeidsoppgaver har blitt automatisert – yrkesskader nær null».

Noen av deltakerne forstilte seg en verden der kunstig intelligens sørger for at vi ikke produserer mer enn vi trenger, og at problemer med overproduksjon, svinn og sløsing dermed er løst. Som en deltaker skrev: «Det er ingen overproduksjon av klær, og all produksjon er behovsstyrt.»

Noen så dessuten for seg en mer rettferdig fordeling av ressurser globalt, blant annet av mat. En deltaker skrev følgende på en av sine lapper: «verdens matvareprogram nedlagt. Vi klarer å brødfø befolkningen med de ressursene vi har», gjerne ved hjelp av økologiske metoder. «Naturen er endelig i balanse – takket være de mulighetene AI gir til å følge med og sette inn riktige tiltak,» skrev en.

Flere så for seg en verden der AI bidrar til utviklingen av et bedre helsevesen, med billige automatiserte helsetjenester og undersøkelser som gjør det mulig å forutsi og forhindre sykdom. Mange påpekte at billigere helsetjenester kan bli mer demokratiske helsetjenester, der marginaliserte i inn- og utland får den hjelpen de trenger. Kreftgåten er løst.

Mange så for seg en verden der klimakrisen er over og der rimelig fornybar energi dominerer. Noen så for seg en verden der AI gir grunnlag for helt nye kunstformer som beriker folks liv.

Andre sentrale håp inkluderte en mer desentralisert utvikling og styring av kunstig intelligens, der tech-gigantene ikke dominerte: «AI bryter tech-gigantenes teknologihegemoni.» En slik utvikling ble ansett for å styrke demokratiet på verdensbasis. I deltakernes håp har dessuten kunstig intelligens ført til bedre helsediagnostisering og behandling av alvorlig sykdom, oppfylning av FNs bærekraftsmål, høyere forventet levealder, og et utdanningssystem som oppfordrer til kreativitet og lek – «kreative fag i skolen er like viktige som de akademiske».

Likevel pekte en av gruppene på at også håp og drømmer som tilsynelatende framstod som kun positive, kunne føre til uventede utfordringer. For eksempel ble det trukket fram hvordan en rasjonell distribusjon av matvarer (som i utgangspunktet skulle løse sultproblemer) kunne føre til en form for overvåking av matinntak. Ikke minst ville dette reise spørsmål om hvem som skulle bestemme hva slags mat som skulle fordeles på hvilken måte.

## Realisme

---

Sesjonen om realisme inneholdt en blanding av pessimistiske og optimistiske forventninger. Det var flere som poengterte at håpsesjonens forslag var realistiske, og noen grupper tok med seg momenter fra denne sesjonen videre. For eksempel forventet noen at bedrifter ble tvunget inn i en tydeligere bærekraftsretning, og en skrev: «Industrier som overproduserer får bot...»

Deltakerne så ut til å være enige om at bruk av kunstig intelligens ville få stor betydning innen medisin og helse, men at flere av dagens utfordringer fortsatt var gjeldende, for eksempel knyttet til rettferdig fordeling av kunnskap og teknologi på helseområdet. Faren for overvåking er reell.

AI vil kunne bidra til en revolusjon i matproduksjon, mente noen, som i landbruksproduksjon, transport, lagring og gjenbruk. En slik utvikling vil forsterkes av nye holdninger til forbruk, blant annet gjennom stor aksept for plantebaserte kjøtterstatninger. Andre var mer pessimistiske og ser for seg en verden der 60 % av befolkningen dør som følge av feilernæring og vannmangel.

Energibehovet er fortsatt ikke løst, mente noen. «80 % av energien brukes til datasentre».

Noen så for seg en verden der AI bidrar til utviklingen av nye kreative læringsformer i dataspill og i metaverse.

Flere gjennomgående problemer ble tatt opp i gruppediskusjonene, som blant annet cyberkrig og autonome våpen. En deltaker var for eksempel opptatt av hvordan krigføringen endres av autonome våpen, siden det ikke lenger var soldater på bakken vil angrepene rettes inn mot sivile mål og infrastruktur.

Noen så for seg en verden der kunstig intelligens ville gi autoritære regimer enda mer kontroll over befolkningen. Her ble det også pekt på en stadig mer konfliktfylt og polarisert mediesituasjon, der det ville være vanskeligere å skille mellom propaganda og ekte nyheter. Andre pekte derimot på at kunstig intelligens kunne styrke demokratiet gjennom bedre tilgang på kunnskap for alle - «stadig vanskeligere å utføre cyberangrep og lage fake news som når ut til alle».

«Personrettingen av medier har nådd punktet der vi ikke lenger har noen felles offentlig samtale,» skrev en, noe som får følger for demokratiet. De store selskaperes algoritmer brukes aktivt til å styre den offentlige debatten.

Flere forventet dessuten noen brudd med dagens utvikling, blant annet i form av flere analoge trender, og at nye generasjoner i førersetet ville ha med seg et

annet tankegods enn dagens ledere. Andre igjen var opptatt av hvordan en fortsettelse av dagens teknologiske utvikling kunne få betydning for utvikling av mellommenneskelige relasjoner og sosiale ferdigheter, spesielt for barn: «alle barn i vestlige land er nærsynt pga. skjermbruk og de har mindre evne til å føre en samtale, smalltalke og snakke med fremmede».

Noen gruppedeltakere forventet at kunstig intelligens ville dominere i enda større grad i 2060, og fryktet utbredelse av digital fattigdom med større forskjeller mellom folk, utfordringer knyttet til personvern og ikke minst eierskap til data. Enkelte pekte på store konsekvenser av identitetstyveri i 2060. Behovet for, og ikke minst utfordringer med å få på plass et globalt regelverk for kunstig intelligens, ble også diskutert. Dette blant annet fordi regioner og nasjonalstater ville legge til grunn ulike verdier for utvikling av kunstig intelligens.

Andre var optimistiske. AI i arbeidslivet vil gi mer fritid. Noen mente borgerlønn vil bli innført og at 50 % omsorgslønn er den nye trenden.

## Mot-scenario

---

Hensikten med de to første sesjonene var å få synliggjort deltakernes verdier og for-forestillinger. Øvelsene ga ikke radikalt nye ideer, rett og slett fordi det er utfordrende å se for seg andre måter å organisere samfunnet på. Likevel kan slike øvelser være nyttige for å hjelpe deltakerne til å bli bevisst sin egen forståelse av dagens verden, samt i å identifisere framtidige utfordringer og muligheter.

I den tredje sesjonen ble deltakerne presentert for et mot-scenario som beskrev en verden i 2060 der de store IKT- og medieselskapene har fritt spillerom globalt. Meta/Facebook, Alphabet/Google, Microsoft, Apple og andre utvikler fortløpende nye tjenester, teknologier og applikasjoner der kunstig intelligens spiller en viktig rolle. Disse selskapene har enorm makt til å forme verdenssamfunnet etter egne ønsker og behov.

I denne verdenen er dessuten nasjonalstaten sterkt svekket, og internasjonale organisasjoner som EU og FN viser liten evne til å kontrollere, styre eller begrense bruken av kunstig intelligens på nye områder.

Deltakerne ble bedt om å reflektere rundt på hvilken måte dette former den verdenen vi ser i 2060, og hva folk gjør for å håndtere utfordringene som følger av dette. Poenget var ikke at mot-scenariet skulle være realistisk eller ønsket – bare at det skulle være annerledes.

Flere deltakere forestilte seg en verden med økt konflikt og økt ulikhet, der statsoverhoder er byttet ut med en «digital» elite – representert ved hvite menn. De som eier aksjer i de store selskapene blir den nye «0,01% adelen», som en uttrykte det. Personvern er et luksusgode forbeholdt de rikeste. Flere var bekymret for hva som skjer med de overnasjonale utfordringene – for selv om de store teknologiselskapene muligens er mer omfattende enn nasjonalstater, vil vi kanskje se de samme problemene som i dag, for eksempel når det gjelder å løse klimakrisen.

Noen mente at økt bruk av cyberspace, herunder metaverse lignende virtuelle verdener, vil føre til at folk tilbringer enda mer tid online, og at det fører til en økt fremmedgjøring fra den «virkelige» verden. Dette kan også føre til mindre interesse for det som skjer i den fysiske verden og økende mistro til alle krav på sannhet.

Samtidig forestilte flere deltakere seg også at en slik verden fører til løsninger på klimaproblemet, fordi finans er en viktig driver mot grønnere løsninger i dag. En deltaker pekte på at samfunnet er prisgitt at de store selskapene har kunnskap om, og setter rammer og lover for kunstig intelligens basert på etikk- og bærekraftshensyn.

Noen deltakere så for seg at et slikt samfunn nødvendigvis måtte innebære en ny form for global styring og politikk – og ikke minst et globalt lovverk. En deltaker skrev: «Cyberforsvaret har flere ansatte enn noen gang, og cybersikkerhet er en del av hverdagen. Passord og brukernavn = pass». På denne måten blir innbyggernes identitet knyttet til hvilket selskap de tilhører, framfor andre identitetsmarkører, som for eksempel nasjonalitet, språk eller etnisitet. Noen pekte på at slike digitale stammer ble viktigere enn familieband. Muligheten for digitale klasser ble også diskutert – at noen har mulighet til å få kjøpe et «pro»-abonnement til et av selskapene, mens andre kanskje måtte nøye seg med å betale ved hjelp av salg av egne data.

Noen grupper tegnet et bilde der teknologiselskapene tar over utdanning så vel som forskning, og tar kontroll over kritisk infrastruktur. Noen fryktet at små- og mellomstore medieaktører dør ut, og at store medieselskap dominerer. I dette bildet fører det også til økt sensur, samt en forringelse av språklig mangfold, der blant annet engelsk er enda mer dominerende.

Flere så for seg motbevegelser, og pekte på at sterke krefter og føringer også ville medføre protester og reaksjoner blant folk. Disse motbevegelsene tar form som analoge enklaver, open source hackerkulturer og

oppblomstring av desentrale «frie» nettverk. Andre pekte på voldelige motbevegelser som drev med massiv sabotasje av teknologiinfrastruktur. Noen motbevegelser inkluderer wifi-frie soner og AI-frie områder, men det er utfordrende å få plass her. En deltaker så for seg oppblomstring av urfolkstradisjoner, og skrev at «Menneskelig utveksling og skaperevne finner nye uttrykksformer eller tar opp urfolkstradisjoner som unndrar seg teknologi (sami, aboriginer).»

En så for seg en egen industri som «gamer» algoritmene til de ulike leverandørene: «De som i 2020 drev med søkemotoroptimalisering, lærer nå folk å optimalisere sjansen for å få jobb, hvordan få billige flybilletter, eller hvordan maksimere pensjonen ved å spille flere leverandører mot hverandre.»

En spenning som ble trukket fram var i hvilken grad teknologigigantene måtte ta hensyn til befolkningens ønsker, behov og personvern for å beholde kundemassen. På den ene siden vil befolkningen bli «låst» inn i de respektive selskapenes tjenester innen helse, jobb, fritid og shopping – vi lever i hver våre «økosystem», som en deltaker skrev. En annen deltaker uttrykte seg på denne måten: «Lojalitet til Meta etc. vil lønne seg, og valg av tech-gigant blir livsvalg.» Illojalitet kan dessuten straffe seg i dette samfunnet, blant annet gjennom manglende tilgang på jobbmarkedet hvis man ikke gir fra seg data.

På den andre siden vil konkurransen mellom selskapene komme innbyggerne til gode. Som en deltaker skrev: «Selskapene må tiltrekke flere kunder, og de etterspør derfor kunnskap og erfaring fra ulike grupper kunder.» Det var også noen deltakere som pekte på at et slikt samfunn tvinger fram ny kompetanse og nye læringsformer, der blant annet samarbeidsevner er særskilt viktig og verdsatt.

## Avsluttende betraktninger

Diskusjonene i gruppene vitner ikke om noen ensidig teknologioptimisme når det gjelder bruk av AI. Deltakerne setter også bruken av AI inn i en større sosiokulturell kontekst, noe som skaper mange rike og komplekse fortellinger om hvordan kunstig intelligens kan være med på å forandre den verdenen vi lever i. Den til dels sterke skepsisen reflekterer nok et generelt skifte i holdninger til fremtiden, blant annet som følge av et økende antall kriser, fra pandemi via klimakrise til krig i Ukraina.

Selv om mange viser til en rekke måter AI kan brukes til å gjøre verden bedre, er vi langt fra den formen for håpefull optimisme vi så for bare få år siden.

Denne teknologipessimismen fører til en form for motløshet hos noen, i den forstand at det etterlatte

inntrykket er at utviklingen er ut av kontroll og at vi står overfor krefter som få eller ingen kan håndtere. Det kan se ut som en økt fremtidskompetanse eller fremtidsbevissthet fører til handlingslammelse snarere enn handlekraft.

Men denne formen for handlingslammelse kan like gjerne være et tegn på manglende fremtidskompetanse, ettersom man alt for lett projiserer dagens virkelighet over på fremtiden. De krisene som er nevnt ovenfor viser oss at vi svært ofte tar feil om fremtiden – at vi «koloniserer» fremtiden med våre begrensede forestillinger om hvordan verden bør være og kommer til og bli – og uten at vi klarer å komme oss ut av de begrensningene eksisterende institusjonelle og kulturelle strukturer setter for endret adferd.

Flere av deltakerne mener da også at menneskenes kreativitet, intelligens og sosiale bevissthet kan bidra til at man motvirker mange av de negative konsekvensene av AI og at man kan gjøre bruk av de positive effektene på en måte som ikke ødelegger miljøet eller menneskenes frihet.

Alle de ulike fremtidsbildene som denne øvelsen fikk frem vitner om et mangfold av muligheter som man ikke alltid kan planlegge for, men som allikevel gjør det lettere for folk å «tenke utenfor boksen» og være åpne for det uforutsette når det skjer. Det kan bidra til økt handlekraft, ikke mindre, rett og slett fordi man nå er i stand til å forestille seg andre scenarier med andre muligheter.

AFINO er et nettverks- og læringscenter for ansvarlig innovasjon og bedrifters samfunnsansvar i Norge. Senteret ble etablert i 2019 for en fem-års periode, og finansieres gjennom Forskningsrådets program «SAMANSVAR – Ansvarlig innovasjon og bedrifters samfunnsansvar».

AFINO består av et nettverk av forskningsinstitusjoner og prosjekter som utforsker hvordan forskning og innovasjon kan utføres på en samfunnsansvarlig måte. Målet er å utvikle kompetanse og nye metoder som sikrer at forskning og innovasjon er ansvarlig, bærekraftig og rettferdig.

En viktig del av senterets arbeid er å etablere nye læringsarenaer på tvers av fagfelt og sektorer. Gjennom å organisere aktiviteter og arrangementer for forskere, innovatører og beslutningstakere ønsker AFINO å skape en felles forståelse av hva ansvarlig innovasjon er.

AFINO ledes av NTNU, BI, UiB og OsloMet, og NTNU er koordinator (Program for anvendt etikk ved Institutt for filosofi og religionsvitenskap). GenØk, UiS, NIFU og HVL er med som samarbeidspartnere.

## Note

1. Laboratoriet ble ledet av Per Koch, spesialrådgiver ved NIFU. Lina Ingeborgrud, seniorforsker ved NIFU, Silje Marie Svartefoss, forskningsassistent ved NIFU, og Celine Marie Hasle, seniorrådgiver for organisasjonsutvikling ved BI, bidro som gruppeledere. Synniva Larsen, leder for studentopplevelser ved BI, har vært sentral i planleggingsarbeidet med laboratoriet. Algorithmic Accountability: Designing Governance for Responsible Digital Transformations ledes av professor Christian Fieseler.

## NIFU

Nordisk institutt for studier av innovasjon, forskning og utdanning

Nordic Institute for Studies in Innovation, Research and Education

NIFU er et uavhengig samfunnsvitenskapelig forskningsinstitutt som tilbyr handlings- og beslutningsorientert forskning til offentlig og privat sektor. Forskningen omfatter hele det kunnskapspolitiske området – fra grunnopplæring, via høyere utdanning til forskning, innovasjon og kompetanseutvikling i arbeidslivet.

### NIFU

PB 2815 Tøyen, NO-0608 Oslo  
www.nifu.no | post@nifu.no