

# KI og teknologien bak

Tilstandsrapport våren 2024

## Innledning

---

Kopinors KI©K-prosjekt startet opp i desember 2023 og har som mål å gjøre administrasjonen, styret og medlemsorganisasjonene bedre i stand til å utforme politikk og strategi og bedre kunne vurdere risiko og mulig gevinst ved fremtidig forretningsutvikling, spesielt knyttet til å tillate og lisensiere bruk av verk i språkmodeller og maskinlæring.

Prosjektet vil særlig vektlegge juridiske og teknologiske faktorer i utredningen. KI©K-notat nr. 1 tok for seg opphavsrettslige forhold, mens dette notatet mer er en teknologisk tilstandsrapport.

Juridiske og teknologiske forhold griper inn i hverandre, noe som også til en viss grad preger denne rapporten. I tillegg kjennetegnes alle beskrivelser av generativ KI av at feltet er under utvikling i et høyt tempo.

I denne teksten skal vi se på hva man kan med generativ KI, hvordan teknologien er bygd opp, og vi vil kaste noen blikk på fremtiden. Vi skal også se på hvordan rettighetshaverne kan bruke teknologi til å beskytte rettighetene og reservere seg mot tekst- og datautvinning og bruk av åndsverk i KI-modeller.

Rapporten er skrevet av Geir Terje Ruud, med bidrag fra Julie Myhre Barkenæs (avsnittet om teknologi for å beskytte åndsverk).

## Innhold

---

Innledning.....	1
Hvordan virker generativ KI? .....	2
KI-tjenester i bruk.....	4
Beskyttelse av åndsverk .....	6
Veien fremover .....	8
Tillegg.....	10

## Hvordan virker generativ KI?

De fleste som leser dette har selv testet generativ kunstig intelligens i en eller annen form, og kjenner godt til hvordan man kan få hjelp til å finne informasjon og forbedre tekst, kanskje også oversette, transkribere, lage en illustrasjon eller produsere musikk. De som har testet det, har trolig også sett hva som virker godt og når systemene ikke virker slik vi håper: det ekspertene kaller hallusinerer.

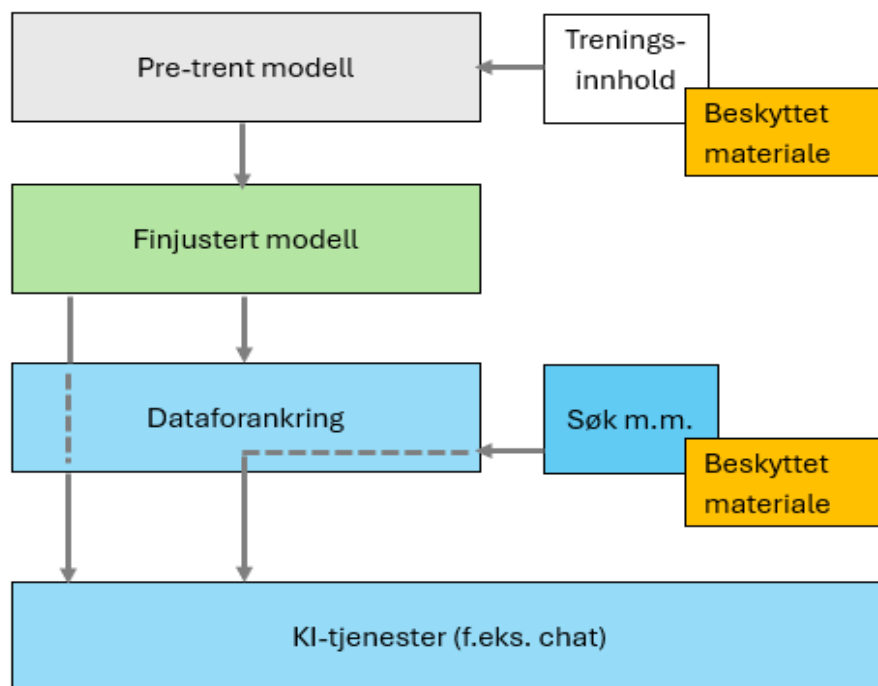
KI-løsningene er ikke sosialt intelligente, og de har ikke følelser eller mulighet til å foreta vurderinger. I en artikkel i *The Economist* om hvordan musikk laget av KI aldri kan treffe oss så godt som musikk laget av mennesker, beskriver en ekspert at det er «fordi en maskin aldri har blitt dumpa».<sup>1</sup>

Rutineoppgaver med behov for presisjon, som oversettelse av en del type tekster, har imidlertid vist seg å treffe relativt godt. Det samme gjelder når generative KI-tjenester blir bedt om å utføre oppgaver innenfor rammene av eksisterende dokumenter.

Enkelt forklart så skaper KI-løsningene resultater etter sannsynlighetsberegning. En tekst fra ChatGPT skapes ved at det neste ordet i en setning velges ut fra en matematisk formel. Hvis et ord som regel følger etter et annet ord i de millioner av setninger som ligger til grunn for ChatGPT, er det også sannsynlig at det er nettopp det ordet som velges i den setningen du får.

### Trinnene i KI-prosessen

Utviklingen av KI-løsninger basert på språkmodeller skjer i flere stadier. Til hjelp har vi oversatt og bearbeidet en figur hentet fra en rapport fra medieorganisasjonen News Media Alliance.<sup>2</sup>



<sup>1</sup> [A new generation of music-making algorithms is here \(economist.com\)](https://www.economist.com/technology-and-innovation/2022/07/27/a-new-generation-of-music-making-algorithms-is-here)

<sup>2</sup> [AI-White-Paper-with-Technical-Analysis.pdf \(newsmediaalliance.org\)](https://www.newsmediaalliance.org/ai-white-paper-with-technical-analysis)

1.

**Språkmodellen** trenes på enorme mengder tekstdata for å identifisere mønstre, sammenhenger og strukturer i språket. Modellen bruker avanserte algoritmer fra et felt kalt maskinlæring, spesielt en underkategori kalt dyp læring, som innebærer nevralt nettverk med mange lag<sup>3</sup>.

2.

I det neste nivået – **finjustering** (*fine tuning*) – trenes språkmodellen videre på mindre, men mer spesifikke datasett. Denne prosessen gjør at modellen bedre kan tilpasse seg og «forstå» den spesifikke oppgaven den skal hjelpe til med. Modellen blir bedre til å gi nøyaktige resultater for spesielle situasjoner.

3.

Det tredje nivået – **dataforankring** (*grounding*) – refererer til prosessen med å koble utdata/svar fra store språkmodeller til faktiske datakilder. Koblingen sikrer at den genererte teksten er konsistent med informasjon fra den virkelige verden, noe som gjør den mer pålitelig og troverdig. Dette kan gjøres på flere måter:

- Justering på datasett hvor faktisk nøyaktighet er avklart, som vitenskapelige artikler eller historiske dokumenter, eventuelt forsterket med menneskelige tilbakemeldinger på faktasjekk og relevans av språkmodellens svar.
- Relevant informasjon hentes fra eksterne kunnskapsbaser og inkorporeres i språkmodellens respons. Verktøy som Google-søk eller spesialiserte kunnskapsbanker kan brukes til dette formålet.

4.

Det nederste nivået i modellen er **KI-tjenestene** – applikasjoner og systemer som bruker språkmodeller for å løse spesifikke oppgaver. Eksempler er ChatGPT, Google Gemini og Anthropic Claude.

Disse tjenestene bygger på underliggende språkmodeller, men tilbyr et brukergrensesnitt og funksjonalitet rettet mot sluttbrukere. Mens språkmodeller kan tilpasses mange formål, er KI-tjenester skreddersydd for bestemte bruksområder som å svare på spørsmål, skrive tekster, oversette, rette skrivefeil, generere kode osv.

Opphavsrettslig beskyttet materiale brukes som treningsdata for språkmodellene, men det er særlig kombinasjonen av KI-tjenester og dataforankring som gjør at en bruker i noen tilfeller kan « snakke » med en lærebok, ghostwrite den neste boken i en romanserie eller lage et nytt automatisk nyhetsnettsted basert på andre mediers nyheter.

## Hallusinerer

En av de store utfordringene med generativ KI er at det kan komme tekster som er fulle av det man i beste fall kan kalle upresis informasjon, såkalt hallusinerer. Det samme skjer ved generering av bilder, der systemet noen ganger tar seg ganske solide friheter. Dette skjer blant annet når språkmodellen prøver å generere tekst over et tema den ikke har treningsgrunnlag for.

Terje Brasethvik i konsultantselskapet Bouvet og professor 2 på NTNU jobber tett med NorwAI-prosjektet som blant annet handler om å bygge en norsk språkmodell. Han sier:

---

<sup>3</sup> For den som ønsker å fordype seg mer i hvordan språkmodeller trenes, anbefales denne teksten fra fysikeren og informatikeren Stephen Wolfram: [What Is ChatGPT Doing ... and Why Does It Work? \(Stephen Wolfram Writings\)](#)

– Hvis datagrunnlaget ikke har tekst over det temaet du spør om, så vil den dessverre ikke svare «vet ikke» men prøver så godt den kan å predikere ord etter ord. Det kan skyldes manglende treningsdata, dårlig treningsdata med mye støy, eller også feil, motsigelser o.l. i treningsdataene. Det kan også være andre statistiske rariteter i modellen, f.eks. at man har forsøkt å trene en for stor modell (for mange parametere) uten tilstrekkelig datagrunnlag. I tillegg til at språk jo er fullt av nyanser som modellen kanskje ikke har fanget helt opp. Det kanskje enkleste botemidlet er å være mer nøye med prompting.

Det pågår et stort forsknings- og utviklingsarbeid med å dempe hallusinerer, og modellene blir stadig bedre.

## KI-tjenester i bruk

---

Det utvikles nå stadig flere typer KI-tjenester som går lenger enn det statiske chat-formatet vi kjenner fra ChatGPT.

### Dialogtjenester

#### Søk

Hvis språkmodellen kobles sammen med materiale som kan søkes opp i sanntid, øker verdien av modellen og den kan gå i retning av en dialogløsning som nærmest umiddelbart kan gi svar på kompliserte spørsmål og oppgaver. Slik kan det oppstå en glidende overgang mellom det som er lagt i KI-modellen, med eller uten avtale, og det som ligger åpent tilgjengelig for søk.

Google har allerede lansert *zero-click-søk* som gir deg svar basert på det underliggende materialet, uten at brukeren nødvendigvis trenger å besøke originalkilden. KI-tjenesten Perplexity<sup>4</sup> tar dette enda lenger, og sammenstiller informasjon fra mange kilder. Brukeren får et oppsummert svar basert på søket, med kildehenvisninger, og kan fortsette dialogen med tjenesten for å få flere detaljer.

#### Beskyttet innhold

Financial Times har lastet sine historiske artikler inn og koblet det til en KI-løsning. Det er et godt eksempel på hvordan man kan la sitt eget, beskyttede materiale skape verdi innenfor rammer som kun er tilgjengelig for et utvalg av deres egne, betalende brukere. Magasinet The Verge har testet løsningen.<sup>5</sup>

#### Egne data

Nyere språkmodeller utvikles med langt større *kontekst-vinduer*, med plass til informasjon som brukeren selv har lagt inn, også innhold man selv ikke har lest. Kontekstvinduet setter rammene for hvordan modellen skal løse oppgaven.<sup>6</sup>

Et større kontekstvindu gir modellen mulighet til å se mer av innholdet samtidig og bidrar til å gjøre oppgavene som skal løses mer spesifikke, med mindre fare for hallusinerer. Dette vil gi mer presise og fullstendige svar, slik Per Kristian Bjørkeng i Aftenposten beskriver det.<sup>7</sup>

---

<sup>4</sup> [Perplexity.ai](https://perplexity.ai)

<sup>5</sup> [Financial Times tests Ask FT, a chatbot trained on decades of its own articles \(The Verge\)](https://www.theverge.com/2023/11/14/financial-times-tests-ask-ft-chatbot)

<sup>6</sup> [What is a context window? \(techtargget.com\)](https://techtargget.com/what-is-a-context-window/)

<sup>7</sup> [Nå kommer KI som kan lære alt om deg og ditt \(aftenposten.no\)](https://aftenposten.no/na-kommer-ki-som-kan-laere-alt-om-deg-og-ditt)

## Visuell generativ KI

KI-tjenestene som leverer bilde- og videogenerering bruker dype nevralt nettverk trent på enorme datasett med bilde-tekst-par. Nettverkene lærer mønstre, stiler og sammenhenger fra disse dataene, slik at de kan generere nye, originale bilder og videoer basert på tekstlige beskrivelser gitt av brukeren.

Prosesen starter med at brukeren gir KI-en en ledetekst (prompt) som beskriver ønsket innhold, stil og komposisjon. KI-en tolker denne teksten og bruker sine lærte mønstre til å generere et passende bilde eller en video, ofte gjennom en teknikk kalt **diffusjon**, der den starter med støy og gradvis jobber fram et bilde. Prosessen er basert på tusenvis av andre bilder som modellen har analysert, plukket fra hverandre og så gradvis erstattet bildet med støy. Dette gjør modellen i stand til å bygge bilder fra støy til noe ferdig.

Mye brukte visuelle KI-tjenester er *Dall-E* (OpenAI), *Stable Diffusion* (Stability AI) og *Midjourney*, mens *Sora* (OpenAI) kan generere video.

Gyldendal har foretatt en gjennomgang av hvordan de forskjellige visuelle tjenestene er bygget opp:

## Bildegeneratorer: hvor ligger landskapet?



	Shutterstock AI	Adobe Firefly	DALL-E 3	Gemini	Midjourney	Stable Diffusion
Modell kun trent på eget innhold/data	✓	✓	✗	?	✗	✗
Tilbyr beskyttelse mot trening av input	✓	✓	✓*	?	✗*	✗*
Tillater kommersiell bruk av output	✓	✓	✓	?	✓	✓
Tilbyr kommersiell <u>beskyttelse</u> på output	✓	✓*	✗	?	✗	✗

\* Uklart eller avhengig av enterprise-modeller eller tredjepartslisens

## Musikk med generativ KI

KI-tjenester for musikkgenerering baseres på modeller som bruker dype nevralt nettverk trent på store musikkdatabaser for å lære mønstre og strukturer i musikk. Brukeren gir inngangsparametere som sjanger og stemning, og KI-tjenesten genererer ny, original musikk basert på dette ved å predikere sannsynlige sekvenser av toner, akkorder og rytmer. Kvaliteten på den genererte musikken avhenger av størrelsen på treningsdataene, kompleksiteten i den underliggende modellen og graden av menneskelig etterbehandling.

Eksempler på musikk-tjenester er de amerikanske *Udio* og *Suno*, som begge er amerikanske, og indiske *Beatoven*.

## Tale til tekst

Transkribering av tale gjennom verktøy som bygger på OpenAIs teknologi *Whisper* har spart mange mennesker for mange titalls timer med tidkrevende arbeid. Etterhvert er de fleste store språk ganske velfungerende i en tale til tekst-løsninger. Systemet skal være trent på nær 700 000 timer med flerspråklige lyddata hentet fra det åpne internettet. Systemet kan

transkribere tale til tekst på flere språk og oversette tale til engelsk, takler aksenter, bakgrunnsstøy og teknisk språk.

Nasjonalbiblioteket har videreutviklet Whisper ved å trene den på transkribert norsk tale fra NRK og Stortinget.<sup>8</sup>

## Tekst til tale

I dag kan man ta opp noen få sekunder med lyd av en person, lage en syntetisk versjon av stemmen og stemmeføringen og få den til å lese opp hva som helst. En tjeneste som *Wondercraft.ai* kan på noen sekunder ta en lang artikkel, gjøre den kortere og få den lest opp av en stemme fra deres system, eller din egen stemme hvis du foretrekker det.

Tekst-til-tale-teknologien (TTS) omdanner skrevet tekst til hørbar tale ved å bruke algoritmer fra kunstig intelligens og maskinlæring for å analysere tekstens lyder, syntaks og kontekst. TTS-systemer bryter ned teksten i fonemer, tolker kontekst og simulerer naturlig uttale og tonefall. Teknologien genererer tale ved å kombinere forhåndsinnspilte lyder/stemmer eller syntetisere nye stemmer fra bunnen av. Resultatet er en til dels svært realistisk og menneskelignende tale. Tale-til-tekst-systemene er en egen teknologi, men ved å koble på/integrere store språkmodeller forbedres systemets evne til å produsere mer dynamisk og naturlig tale, og utvider bruksområdene for begge teknologier.

## Beskyttelse av åndsverk

---

Opphavsrettslig beskyttet materiale som ligger på internett, skrapes til bruk i KI uten at rettighetshavere har gitt sin tilslutning til dette. Etter EUs digitalmarkedsdirektiv artikkel 3 må rettighetshaverne akseptere slik tekst- og datautvinning for forskningsformål. Det er imidlertid en bekymring for at dette materialet brukes i opptreningen av de store språkmodellene utover det som kan anses å være til forskningsformål, uten tillatelse eller at det gis kompensasjon.

Hvilke teknologiske muligheter finnes for å beskytte eget innhold mot utnyttelse, eller trekke egne verk fra et treningsmateriale?

### Avlæring

Uansett om en språkmodell er trent på materiale hvor man har en avtale om bruk, eller om det er «lånt» eller «stjålet», kan det stilles spørsmål om det er teknisk mulig å trekke sitt materiale ut av en ferdig trent språkmodell.

Foreløpig ser det ikke ut til at noen har vist måter dette kan gjøres på, eller vist om det i det hele tatt er mulig. Ken Liu ved Stanford University har nylig skrevet en god tekst som går gjennom en del problemstillinger og gir en oversikt over feltet.<sup>9</sup>

### Teknologi for å beskytte åndsverk

EUs digitalmarkedsdirektivet gir i artikkel 4 anledning til å unnta verk fra tekst- og datautvinning for andre formål enn forskning gjennom formuleringen: «Dette unntaket forutsetter at rettighetshaveren ikke uttrykkelig har nedlagt forbud mot slik bruk.» I departementets forslag til endring av åndsverkloven, som nylig har vært på høring, heter det i den nye § 50 e: «Opphaver kan på en hensiktsmessig måte forbeholde seg retten til slik bruk.»

---

<sup>8</sup> [Ny modell oversetter norsk tale til tekst nesten uten feil \(aftenposten.no\)](https://aftenposten.no/ny-modell-oversetter-norsk-tale-til-tekst-nesten-uten-feil)

<sup>9</sup> [Machine Unlearning in 2024 \(Ken Ziyu Liu - Stanford Computer Science\)](https://stanford.edu/~kenliu/machine-unlearning-in-2024)

Det finnes flere måter å uttrykke et slikt forbehold, som i ulik grad kan sies være hensiktsmessig for opphaver.

### Reservasjonstekst på egne nettsider

Den rent teknisk enkleste måten å reservere seg mot skraping er å ha en reservasjonstekst på egne nettsider som forteller at det ikke lov å skrape eller høste fra nettsiden. Om dette faktisk er en løsning, kan diskuteres. En slik tekst forhindrer hindrer ikke skraping og er ikke standardisert på noen måte som gjør det lesbart for en crawler.

### Robots.txt

Robots.txt-fil brukes av nettstedet for å fortelle søkeroboter eller crawlere hvilke sider eller URL-er som kan gjennomføres og ikke<sup>10</sup>. Robots.txt er filnavnet som brukes for å implementere en egen protokoll som kan leses av en crawler. Det er imidlertid «frivillig» å forholde seg til robots.txt, og det er ikke noe som hindrer en crawler som ignorerer robots.txt tilgang til alle sidene eller URL-ene, også de som ligger i listen over sider en crawler ikke skal ha tilgang til.

Robots.txt er ikke noen offisiell standard, men vedlikeholdes av en uavhengig organisasjon og brukes av svært mange nettsteder for å regulere søketrafikken.

### Ai.txt

Ai.txt er også en tekstfil på nettstedet som inneholder instruksjoner til datautvinner om hvorvidt tekst- og mediefiler på dette nettstedet kan brukes til å trene opp kommersielle KI-modeller<sup>11</sup>. Filen inneholder verdiene «blokker» eller «tillat» for hver enkelt innholdstype (som tekst, bilde, lyd, video og kode).

Ai.txt ble lansert i 2023 av selskapet Spawning.

### International Standard Content Code, ISCC

ISCC er en standard for å lage unike identifikatorer eller fingeravtrykk til digitale objekter<sup>12</sup>. Standarden skal kunne brukes for alle objekttyper (tekst, bilde, lyd og video) og utvikles av The ISCC Foundation.

ISCC har definert en algoritme som utleder ISCC-koden/identifikatoren fra innholdet i det digitale objektet. ISCC er en desentralisert metode og skal være pålitelig for å identifisere innhold og for å sikre mot tap av metadata eller endring av innhold.

Til det digitale objektet kan det knyttes en erklæring om reservasjoner gjennom en egen protokoll (TDM-AI<sup>13</sup>). Erklæringene offentliggjøres i den distribuerte databasen med tilhørende ISCC og er både menneske- og maskinlesbar. Leverandører av KI-applikasjoner *kan* utlede tillatelsene eller restriksjonene fra ISCC, og dermed respektere rettighetshavernes erklæringer.

## Andre tekniske løsninger

Det forskes for tiden på forskjellige tekniske løsninger for å beskytte åndsverk mot skraping. Vi nevner her noen av dem, som i ulik grad kombinerer ren merking av innhold med tekniske beskyttelsessystemer som forhindrer selve innhentingen av treningsdata.

---

<sup>10</sup> [The Web Robots Pages \(robotstxt.org\)](https://robotstxt.org/)

<sup>11</sup> [Spawning ai.txt](https://spawning.ai.txt)

<sup>12</sup> Selve ISO-standarden er under utvikling for publisering, se

[ISO 24138:2024 - Information and documentation — International Standard Content Code \(ISCC\)](https://www.iso.org/standard/79111.html)

<sup>13</sup> [TDM-AI Protocol | TDM-AI \(tdmai.org\)](https://tdmai.org/)

## Kudurru

Tjenesten Kudurru blokker aktivt KI-skrapere fra nettstedet ved hjelp av et nettverk av brukere som identifiserer dem som forsøker å hente innhold<sup>14</sup>. Tjenesten ble lansert av selskapet Spawning i 2023.

## Glaze og Nightshade

En forskergruppe ved University of Chicago<sup>15</sup> har utviklet løsninger for å beskytte visuell kunst fra å bli brukt til KI-trening. For billedkunst er det særlig spørsmålet om beskyttelse mot etterligning av en kunstners stil som er viktig. *Glaze* beskytter mot slike etterligninger ved å beregne et sett med minimale endringer i kunstverket, slik at det ser (nesten) uendret ut for det menneskelige øye, mens det for KI-modeller ser ut som en helt annen stil. *Nightshade* fungerer på lignende måte, men forstyrrer treningsdataene slik at det ødelegger for KI-modellene.

## Digital vannmerking

Det foregår mye forskning på feltet KI og vannmerking. Cornwell University ser på vannmerking for store språkmodeller og på mulighet for digital vannmerking av modellutdata, det vil si å bygge inn signaler i generert tekst som er usynlig for mennesker, men som kan leses av algoritmer.

## Veien fremover

---

Denne rapporten er skrevet som et øyeblikksbilde av generativ KI våren 2024. Teknologien har utviklet seg betydelig de siste 18 månedene.

Det er aldri sett en like rask utbredelse av ny teknologi som etter at ChatGPT ble lansert. På under tre måneder hadde 100 millioner mennesker tatt systemet i bruk. Til sammenligning tok det flere år etter lansering av iPhone før 100 millioner mennesker hadde en smarttelefon, det tok Netflix ti år å få 100 millioner kunder og selv en magnet på unge menneskers liv som TikTok brukte ett år på å oppnå 100 millioner brukere.

Investeringsviljen er stor i utviklingen av KI-løsninger og det er stor nysgjerrighet i hvordan vi best kan ta det i bruk. Som et eksempel fortalte Aftenposten nylig om hvordan det jobbes for at KI-løsningene skal få lov til å forbedre sin egen ytelse.<sup>16</sup> Med å slippe løs såkalte «agenter», øker verdien av kunnskapen i en språkmodell.

Shailesh Prakash, visepresident for nyheter i Google, med topplederbakgrunn fra Washington Post og med oppvekst i Darjeeling, India, forklarte det ganske fint for publikum på den internasjonale journalistfestivalen i Perugia i april 2024:

*– Tenk deg generativ KI som en fotballkamp som normalt varer 90 minutter. Da er vi kanskje fem minutter inne i kampen. Men vi vet at det kan bli forlenget spilletid i begge omganger, og det kan bli ekstraomganger og straffesparkkonkurranse.*

---

<sup>14</sup> [Kudurru.ai](https://kudurru.ai)

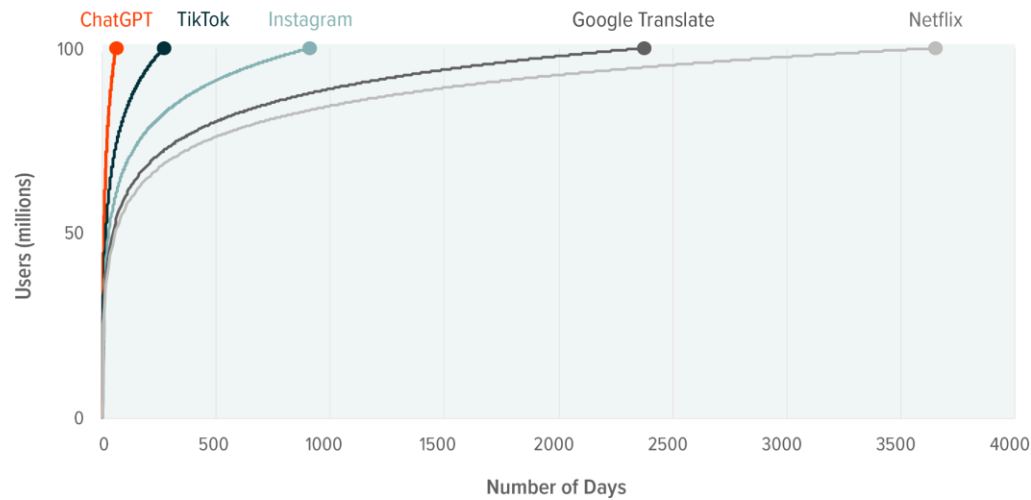
<sup>15</sup> [Glaze - Protecting Artists from Generative AI \(uchicago.edu\)](https://glaze-protecting-artists-from-generative-ai.uchicago.edu)

<sup>16</sup> [I løpet av året får KI-agentene lov til å forbedre seg selv \(aftenposten.no\)](https://aftenposten.no/løpet-av-året-får-ki-agentene-lov-til-å-forbedre-seg-selv)



## TIME IT TOOK COMPANIES TO REACH 100 MILLION USERS

Sources: Global X ETFs with information derived from: BBC News. (2018, January 23). Netflix's history: From DVD rentals to streaming success; Cerullo, M. (2023, February 1). ChatGPT user base is growing faster than TikTok. CBS News.



KI-forskere ved Stanford University leverte i april 2024 en mer enn 500 sider lang rapport om generativ KI.<sup>17</sup> Helt kort konkluderer rapporten med at maskiner nå er dyktigere enn mennesker til de fleste oppgaver, og at det er ekstremt dyrt å lage de aller beste KI-modellene.

Vi avslutter denne teknologiske tilstandsrapporten med en forkortet versjon av forskernes tipunkts sammendrag, som tar utgangspunkt i amerikanske forhold:

1. KI overgår menneskelig ytelse blant annet i bildeklassifisering, visuell resonnering og forståelse av språk, men henger etter i mer komplekse oppgaver.
2. Tek-industrien dominerer forskning og utvikling av KI, heller enn akademia.
3. De mest avanserte modellene blir mye dyrere.
4. USA leder an kappløpet foran Kina, EU og Storbritannia i utviklingen av KI-modeller med høy kvalitet.
5. Robuste regler og standardiserte metoder for å sikre ansvarlighet og transparens i utvikling av språkmodeller mangler.
6. Investeringer i generativ KI skyter i været.
7. Studier viser at KI gjør arbeidstakere mer produktive og fører til arbeid av høyere kvalitet. Likevel advarer andre studier om at bruk av KI uten riktig tilsyn eller oppfølging kan føre til redusert produktivitet.
8. Fremgang i vitenskapelig forskning øker enda mer, takket være KI.
9. Antallet KI-reguleringer i USA øker kraftig.
10. Folk over hele verden er mer bevisste på KIs potensielle innvirkning, men også mer usikre og bekymrede.

<sup>17</sup> [HAI AI-Index-Report-2024.pdf \(stanford.edu\)](#)

# Tillegg

---

## Ordliste

### **Dataforankring**

Dataforankring i språkmodeller (Grounding Data) involverer integrering av eksterne, virkelighetsbaserte informasjonskilder for å forbedre modellens forståelse og nøyaktighet i formidling av fakta, eksempelvis ved å kombinere språkmodellen med Google-søk eller søk i faktadatabaser i sanntid. Dataforankring kan inkludere multimodal læring, hvor tekst kombineres med visuelle eller auditive data, noe som gir en mer helhetlig forståelse.

### **Diffusjonsmodell**

Diffusjonsmodell er en modell som kan generere data som ligner på den de er trent på, særlig brukt ved billedgenerering. Modellen lærer ved å observere mange eksempler på hvordan ekte bilder ser ut, og bruker denne kunnskapen til å gjette seg fram til hvordan den kan gå fra støy til et klart og realistisk bilde. Dette gjør den svært nyttig for å generere nye bilder, forbedre bildekvalitet, eller skape andre typer media som lyd og tekst basert på lært data.

### **KI-crawler**

Språkmodellene bygges opp ved at en KI-crawler beveger seg rundt på ulike nettsteder og samler sammen tekster, både språklige regler og kunnskap fra innholdet. Forskjellen på en KI-crawler og det som tidligere ble brukt for å hente ut informasjon fra internett, er at den ikke har bruk for at innholdet på nettsidene er strukturert eller samlet samme sted.

### **Korpus**

Korpus (tekstkorpus) er en samling av tekster som brukes til språklige eller litterære undersøkelser. Det kan være en elektronisk database med tekster som er samlet for å representere et språk eller en språkvariant. Korpuset kan inneholde ulike typer tekster, som bøker, aviser, tidsskriftartikler, og transkripsjoner fra tale, og det kan være både generelt, dekke mange temaer og sjangre, eller det kan være mer spesialisert og fokusert på et bestemt emne eller felt.

Korpuset kan være *annotert*, «tagget», med lingvistisk informasjon, som ordklasser eller syntaktisk struktur, noe som gjør det mulig å utføre detaljerte analyser av språkbruk og mønstre. Tekstkorpus brukes i utviklingen av språkmodeller og de bidrar til å forbedre teknologier som maskinoversettelse, talegjenkjenning, og automatiserte chatbots/digitale assistenter.

### **Prompt**

Prompt kan oversettes med «ledetekst» og er en instruksjon, spørsmål eller setning som brukes for å initiere en respons eller handling fra en KI-tjeneste. Systemet analyserer prompten og genererer et svar basert på den gitte bestillingen.

### **Retrieval Augmented Generation**

Retrieval Augmented Generation (RAG), er en metode for å utnytte store språkmodeller bedre. Metoden gjør bruk av kvalitetssikret innhold og kan gi mindre risiko for hallusinerer. Det kan også bety at innhold som ikke er i den opprinnelige språkmodellen kan trekkes inn i arbeidet. Det nærmer seg det noen kaller å « snakke med data ».

### **Språkmodell**

En språkmodell er et beregningsverktøy som kan forstå og generere menneskelig språk ved å identifisere mønstre og sammenhenger i store mengder tekstdata. Slike modeller er trent på store datasett med tekstdata for å identifisere mønstre, sammenhenger og strukturer i språket. Språkmodeller spiller en avgjørende rolle i utviklingen av teknologier som stemmeaktiverte assistenter, automatiserte chatbots for kundeservice, og andre applikasjoner som krever forståelse eller generering av naturlig språk.

*Store språkmodeller* (Large Language Models, LLM) som GPT-4 m.fl. har tilegnet seg evnen til språkforståelse og språkgenerering ved å lære statistiske sammenhenger fra tekstdokumenter. Takket være denne «opplæringsprosessen» kan en god LLM brukes til generering av tekster, basert på innholdet i språkdatabasen.

### **Tekst- og datautvinning**

Tekst- og datautvinning (Text and Data Mining, TDM) er kjent fra blant annet språkteknologi gjennom flere tiår. Der det finnes store tekstsamlinger kan man forske på ulike deler av tekstene, som for eksempel hvordan forekomsten av ord og uttrykk endrer seg over tid eller annen type tekstanalyse som er vanskelig å fange opp manuelt. Når det nå bygges store språkmodeller er det et langt mer avansert resultat av tekst- og datautvinning enn det man har sett for seg tidligere.

I digitalmarkedsdirektivet defineres tekst- og datautvinning som «*any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.*»

### **Treningsdata**

Treningsdata er datasettet som brukes til å trene en KI-modell. Den inkluderer eksemplene som modellen lærer å utføre sine oppgaver fra. Kvaliteten, mangfoldet og størrelsen på treningsdataen kan betydelig påvirke språkmodellens ytelse.

## **Produkter og selskaper**

### **OpenAI**

Selskapet ble stiftet av blant andre Elon Musk. Nå er Microsoft største eier. OpenAI lanserte språkmodellen GPT og KI-tjenesten ChatGPT i november 2022 og står også bak generative KI-løsninger som Dall-E og Sora.

### *ChatGPT*

Lanseringen av KI-tjenesten ChatGPT var et gjennombrudd for generativ AI. Det var den første KI-tjenesten som ble lansert med en dialogfunksjon, som ga en opplevelse av reell kommunikasjon med den kunstige intelligensen.. Den første løsningen bygget på GPT3.5 og ble hurtig oppjustert til en betalingsløsning basert på GPT4. I mai 2024 ble en gratisversjon av denne, ChatGPT4o, lansert med begrensninger på antall søk.

### **Gemini**

KI-løsning (tidligere Bard) utviklet av Google som en konkurrent til Open AIs løsninger. Gemini leverer tjenester innen tekst, bilder, lyd og video. Løsningen finnes i tre utgaver, *Ultra*, *Pro* og *Nano*.

### **Mistral**

Fransk KI-tjeneste utviklet av medarbeidere med bakgrunn fra Google og Meta.

### **Claude**

Språkmodeller utviklet av det amerikanske selskapet Anthropic.

### **NorwAI**

Samarbeidsprosjekt for generativ KI-utvikling i Norge, styrt fra NTNU i Trondheim, men med store industribedrifter og medieselskap som partnere.

## **Referanseverk**

For de som vil dykke dypere og få grundigere forklaringer, er det mange muligheter for å lese videre, for det er skrevet mye om KI-løsninger det siste halvannet året. Av bøker på norsk anbefales Per Kristian Bjørkeng: *Knekk ChatGPT-koden* og Inga Strømke: *Maskiner som tenker*, begge utgitt 2023.