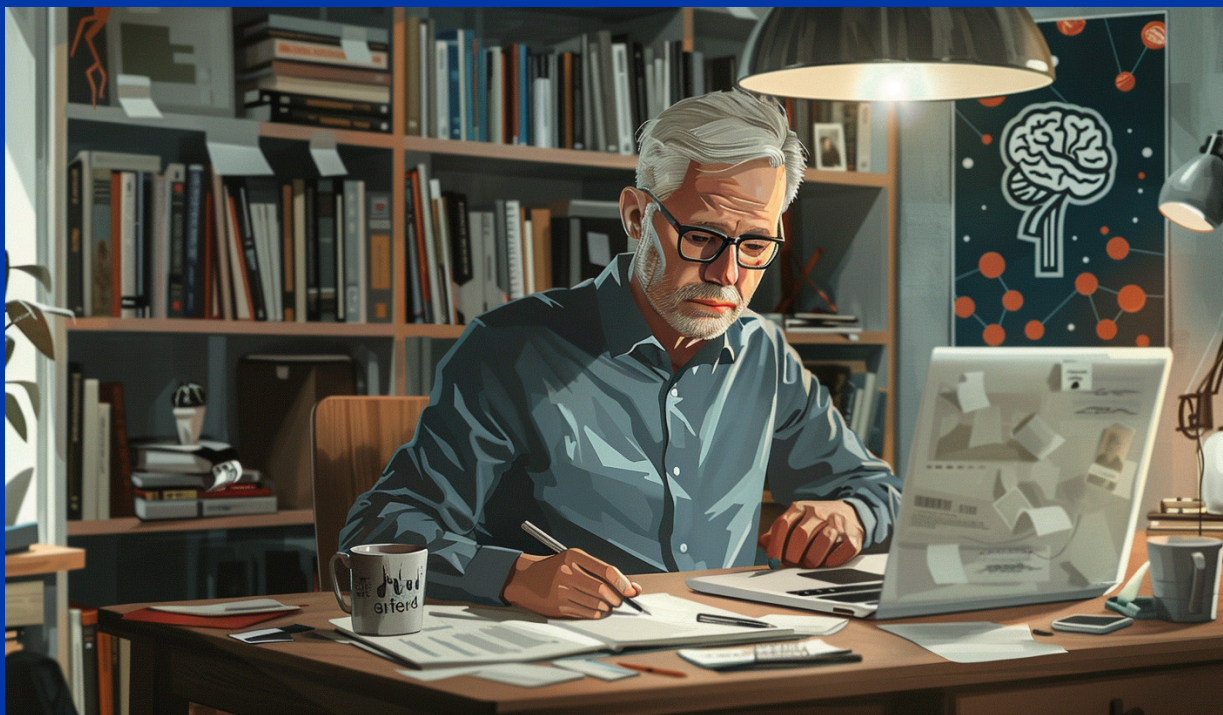


At prompte

En guide af Just Kjærgård Pedersen,
Lektor ved Erhvervsakademi København.



Prompting er et nyt og spændende emne, hvori omdrejningspunktet synes at være at man bør udtrykke sig præcist og tydeligt for derved at få bedre resultater. Generel teori om kommunikation og design kan her hjælpe med at opnå en større grad af præcision og tydelighed. Samtidig bør man være opmærksom på risikoen for bias eller manglende evaluering.

Indholdsfortegnelse

Indledning	3
Hvad er kunstig intelligens	5
Hvad er en prompt	9
Prompt engineering	12
Prompting som problemforståelse	16
Promptens struktur	23
AI som proces	29
Varians og bias	33
Agenter og CustomGPT'er	41
Ophavsret og regulering	47
Prompting som kompetence	52
Videre læsning	54
Figur- og tabelfortegnelse	57
Om forfatteren	58

Indledning

Forestil dig et samfund præcist ligesom vores, men uden biler. Af en eller anden grund er der fine veje, men du og dine medmennesker går, løber eller cykler på dem. Varer transporteres med hestevogn. Byer er indrettet efter, at de fleste bevæger sig til fods. Hvis nogen skal langt væk, tager det tid, og afstande opleves som en reel begrænsning.

Men fra det ene år til det andet sker noget mærkeligt. Pludselig får alle familier en bil. Ikke bare nogle få rige familier, men alle. Bilen står nu i indkørslen. Den kan køre hurtigt, transportere varer og bringe dig steder hen, som tidligere tog timer eller dage at nå.

Problemet er bare, at næsten ingen ved, hvordan denne mærkelige bil egentlig skal bruges.

Nogle kører små ture rundt om blokken for at prøve den af. Andre forsøger at bruge bilen til opgaver, som ret beset ikke giver mening. Nogle bliver frustrerede over, at bilen ikke altid gør præcis det, de forventer. Andre opdager hurtigt, at bilen kan ændre både deres arbejde og deres hverdag. Der bliver talt om kørekort, trafikregler og motorveje, men det er blot idéer, og sandheden er, at trafikken er **et stort kaos**.

Noget lignende er sket med generativ kunstig intelligens.

Siden 2023 er værktøjer som ChatGPT, Copilot og Gemini blevet en del af hverdagen for millioner af mennesker. De kan skrive tekster, forklare komplekse emner, analysere information og hjælpe med en lang række opgaver. Samtidig er der opstået en ny disciplin med et næsten mystisk navn: prompt engineering.

Ideen er enkel. Hvis du vil have gode svar fra en AI, skal du formulere gode prompts. Men hvad betyder det egentlig?

Denne guide handler om netop det spørgsmål. Ikke i betydningen at der findes en magisk formel for den perfekte prompt, men i betydningen at man kan lære at arbejde mere systematisk og reflekteret med generativ AI.

Første udgave af denne guide blev skrevet i **2024**. På det tidspunkt var generativ AI stadig et relativt nyt fænomen for mange organisationer. Fokus var primært på at forstå teknologien og give nogle første praktiske råd om, hvordan man kunne formulere prompts.

Siden da er der sket meget. Generative AI-værktøjer er blevet bedre og mere udbredte. Mange virksomheder er begyndt at integrere AI i deres arbejdsgange. Samtidig er der opstået nye diskussioner om ansvar, bias, ophavsret og regulering, blandt andet i forbindelse med EU's AI-forordning.

Denne udgave er derfor ikke blot en opdatering, men også en udvidelse.

Hvor den første udgave primært handlede om selve prompten, forsøger denne udgave i højere grad at sætte prompting ind i en større sammenhæng. Guiden diskuterer blandt andet, hvordan man formulerer problemer, før man skriver en prompt, hvordan man arbejder iterativt med AI, og hvordan prompts i stigende grad indgår i automatiserede workflows.

Kort sagt: Fokus er flyttet fra enkeltstående prompts til arbejdet med AI som proces.

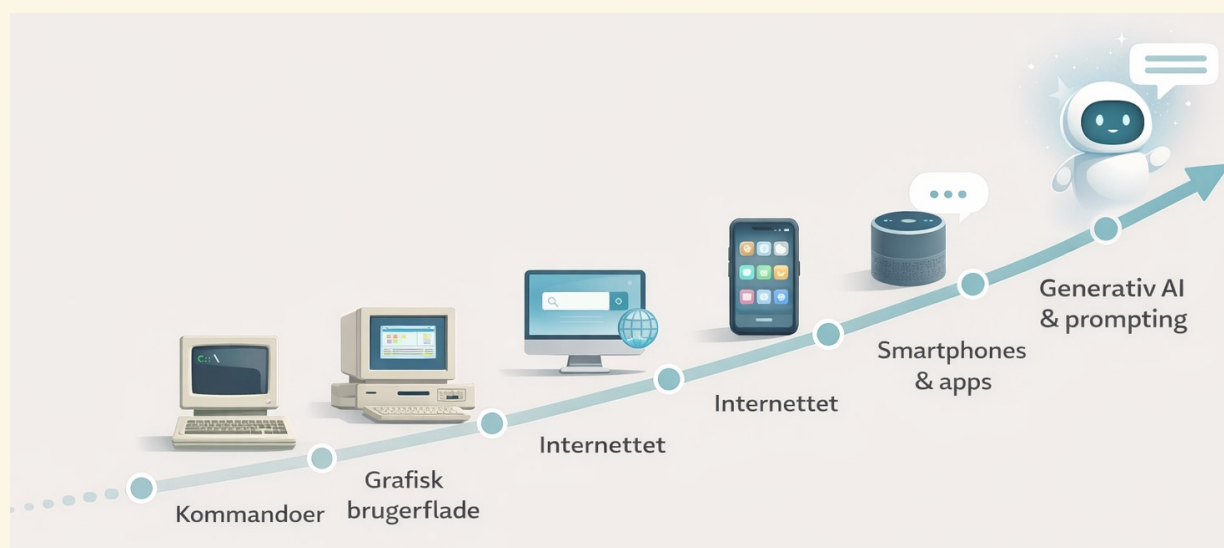
Udviklingen inden for generativ kunstig intelligens går meget hurtigt. Nye modeller lanceres med få måneders mellemrum, og funktioner som for få år siden blev betragtet som avancerede, bliver hurtigt standard.

Det betyder også, at de værktøjer, som vi bruger i dag, sandsynligvis vil fremstå primitive om et eller to årtier. Mens det er let at overvurdere hvor meget kunstig intelligens vil påvirke samfundet på kort sigt, så kommer påvirkningen på længere sigt til at være stor.

Når du lærer at arbejde med AI, handler det derfor ikke kun om at mestre et bestemt værktøj. Det handler om at udvikle nogle mere grundlæggende kompetencer: at formulere problemer, stille gode spørgsmål og kritisk vurdere de svar, du får.

Det er netop disse kompetencer, denne guide forsøger at udvikle.

Den AI som du bruger i dag, er den dårligste AI som du nogensinde kommer til at bruge.



Figur 1 - Den dårligste AI du nogensinde får - tidslinje

Hvad er kunstig intelligens

Når du begynder at arbejde med generativ kunstig intelligens, opstår der hurtigt en række forestillinger om, hvad systemet egentlig gør. Mange forestiller sig, at AI er en slags digital ekspert, der ved en masse om verden. Andre tænker på AI som en avanceret søgemaskine, der kan finde informationer hurtigere end mennesker.

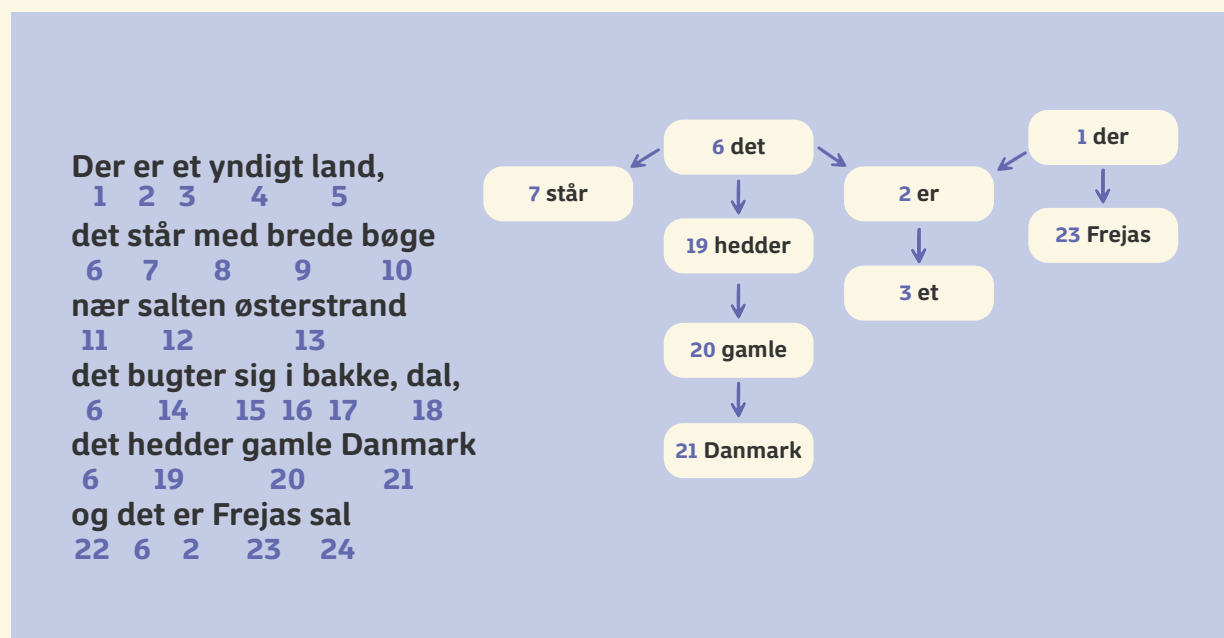
Begge forestillinger rammer noget rigtigt, men de er også misvisende.

For at forstå, hvordan man arbejder med prompts, er det nødvendigt først at forstå, hvad et generativt AI-system faktisk gør.

Den korte forklaring er, at systemet ikke tænker, forstår eller ved noget i menneskelig forstand. Kunstig intelligens **analyserer mønstre** i tekst og beregner statistiske sandsynligheder for, hvilke ord der typisk følger efter hinanden.

Når du skriver en prompt til en AI, starter der derfor en matematisk proces. Systemet analyserer de ord, du har skrevet, sammenligner dem med de mønstre, det har lært under sin træning, og beregner derefter, hvilke ord der sandsynligvis bør komme som næste led i en tekst. Denne proces gentages igen og igen, indtil der er genereret et svar. I nyere modeller kan en tekst blive genereret af et underprogram og rettet til af et andet.

Hvis du beder en AI om at skrive en sætning, vil systemet i praksis omdanne alle ordene til tokens, hvad man kan forstå som et tal. For eksempel vil ChatGPT læse sætningen "Det er en interessant problemstilling" som "3639, 1111, 469, 54257, 4792, 153183"



Figur 2 - Forsimplet eksempel på numerisk tekstanalyse

Hvert token, altså de omskrevne ord, analyseres nu i relation til hinanden og den øvrige kontekst, og ud fra en sandsynlighedsberegning baseret på enorme mængder tekst, som modellen tidligere har analyseret, genereres så et sandsynligt svar i et sprog, som vi mennesker kan genkende.

Resultatet kan ofte virke imponerende, fordi modellerne er trænet på meget store tekstmængder, som blandt andet kan omfatte bøger, artikler, hjemmesider og andre former for digital tekst. De har derfor set millioner af eksempler på, hvordan mennesker formulerer sig. Det betyder imidlertid ikke, at modellen forstår det, den skriver, hvilket kan illustreres med metaforen om det kinesiske rum.

Det kinesiske rum

Den amerikanske filosof John Searle (1980) foreslog tankeeksperiment om det kinesiske rum som en metafor for kunstig intelligens.

Du skal i eksperimentet forestille dig en person, der sidder i et lukket rum og modtager kinesiske tegn gennem en luge. Personen forstår ikke kinesisk, men udvikler stille og roligt en detaljeret regelbog, der beskriver, hvilke tegn der skal sendes retur som svar på bestemte input. Hver gang der sendes noget retur, så får personen at vide om det er godt sat sammen eller ej.

Efter noget tid bliver alle svarene korrekte og for personer udenfor rummet fremstår svarene meningsfulde, selvom der inde i rummet ikke er nogen forståelse af sproget. Pointen hos Searle er, at korrekt output ikke er det samme som forståelse. På samme måde som en digital simulering af et ildsted ikke udvikler varme, så udvikler en digital simulering af ord heller ikke bevidsthed.

Generative AI-systemer fungerer på mange måder på en lignende måde. De producerer svar, der kan virke sammenhængende og intelligente, men processen bag svarene består i at følge statistiske mønstre i tekst.



Figur 3 - Det kinesiske rum

Generativ kunstig intelligens har altså ikke nogen forståelse selv. Når nogen alligevel hævder dette, så er det fordi, det er svært at definere, hvad menneskelig intelligens er. Den amerikanske sociolog James March (2011) har argumenteret for, at mennesker handler i sociale situationer ud fra en vurdering af, hvad der er hensigtsmæssigt, og hvis man tror på det, så må det jo siges at minde om den beregning, som en kunstig intelligens laver.

Det er også vigtigt at huske på kunstig intelligens ikke er kommet ud af ingenting, men er resultatet af en gradvis udvikling af digital teknologi (Chaffey & Ellis-Chadwick, 2019).

Avanceret digital analyse har været udbredt siden 1990'erne, hvor den tidlige brug af kunstig intelligens var baseret på faste **algoritmer**. Det vil sige systemer, hvor mennesker på forhånd havde defineret reglerne for, hvordan data skulle behandles. Et klassisk eksempel er Googles tidlige søgemaskine, hvor rangordningen af søgeresultater byggede på eksplicite kriterier som links mellem websites, sidens struktur og senere også metadata. Googles PageRank-algoritme vurderede en sides relevans ud fra antallet og kvaliteten af links, der pegede på den, og fulgte dermed en relativt klar og regelbaseret logik (Brin & Page, 1998). Hvis du kendte reglerne, kunne du også forsøge at optimere dig frem til et bedre resultat. Meget af den tidlige praksis inden for digital markedsføring handlede netop om at gætte på, hvad algoritmerne ville belønne. I bund og grund var det ikke så forskelligt fra det gamle telefonsbogstrick med at navngive sin virksomhed noget, der begynder med A, for at komme til at stå forrest.

Efterhånden som datamængderne voksede, og brugeradfærden blev mere kompleks, blev denne regelbaserede tilgang utilstrækkelig. Det blev ganske enkelt umuligt manuelt at formulere regler, der kunne dække alle variationer i søgeadfærd, klikmønstre og kontekster. Man begyndte nu at tale om **maskinlæring**, hvor du i stedet for at programmere faste regler begyndte at træne modeller til selv at identificere mønstre i store datamængder. Dette betød, at systemer i stigende grad kunne optimere anbefalinger og budskaber baseret på historisk adfærd, uden at mennesker eksplicit havde defineret alle beslutningskriterier (Jordan & Mitchell, 2015). Hvor fokus tidligere havde været på at forstå og arbejde med algoritmens regler, blev det nu vigtigere at forstå data, signaler og feedback. Samtidig blev teknologien mindre gennemsigtig. Når en model lærer ud fra millioner af datapunkter, er det ikke længere muligt entydigt at forklare, hvorfor en bestemt beslutning træffes.

Med **generativ AI** sker der endnu et skift, hvor kunstig intelligens ikke længere kun bruges til at sortere, vælge og optimere, men også til at producere nyt indhold. Sprogmodeller, de såkaldte Large Language Models, kan generere tekst, billeder og andre outputformer på baggrund af sandsynlighedsberegninger baseret på enorme mængder eksisterende data (Bommasani et al., 2021). Ligeledes kan sprogmodeller tolke på og agere ud fra tekst input fra et menneske. Hvor du tidligere skulle lære, hvad knapperne i et program gjorde, altså lære maskinens sprog, så kan sprogmodellerne tolke og handle ud fra dit eget sprog.

Generativ AI er ikke længere kun med til at understøtte beslutninger, men aktivt deltager i at skabe indhold, budskaber og analyser. Det er denne evne til at generere, der gør teknologien så synlig i hverdagen og samtidig så let at forveksle med menneskelig forståelse og kreativitet. Det er vigtigt at understrege, at generativ AI ikke repræsenterer et brud med den tidligere udvikling, men snarere en forlængelse af den. De grundlæggende principper er de samme. Systemerne arbejder stadig med data, mønstre og sandsynligheder. Forskellen er, at outputet nu antager former, som tidligere var forbeholdt mennesker.

Tabel 1 - 3 former for kunstig intelligens

Form for kunstig intelligens	Hvordan den arbejder	Typisk output
Algoritmer	Behandler og tolker data ud fra faste regler og instrukser, som er defineret på forhånd af mennesker. Systemet kan ikke selv ændre sine regler.	Rangordning, sortering og simple beslutninger baseret på klare kriterier.
Maskinlæring	Identificerer mønstre i store datamængder og opstiller selv statistiske regler på baggrund af historiske data. Modellen lærer af feedback.	Forudsigelser, anbefalinger og sandsynlighedsvurderinger, fx klik, køb eller segmenttilhørsforhold.
Generativ kunstig intelligens	Anvender lærte mønstre til at skabe nyt indhold på baggrund af sandsynlighedsberegninger og menneskeligt formuleret input.	Tekst, billeder og andre outputformer, der fremstår menneskelignende, men er statistisk genererede.

Hvad er en prompt

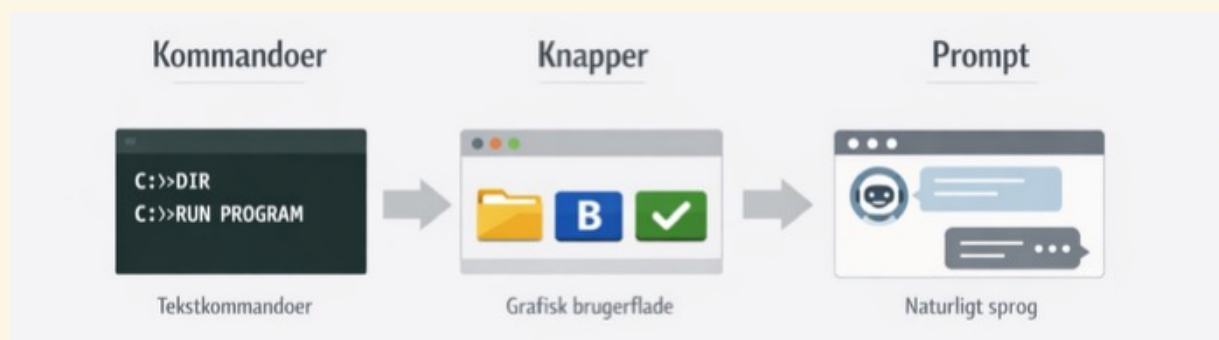
Mens generativ kunstig intelligens bygger videre på fortidens algoritmer og maskinlæring, så er det stadigvæk en ny teknologi. Det kommer tydeligst til udtryk, når vi ser på, hvordan brugerne interagerer med kunstig intelligens. Dette er i sig selv et veletableret forskningsfelt, der enten kaldes UI (user interface) eller HMI (Human Machine Interaction).

I computerens tidlige år foregik denne kommunikation gennem kommandoer. Brugeren skrev instruktioner direkte i et tekstbaseret system, og computeren udførte derefter en handling. Denne form for interaktion krævede, at brugeren kendte computerens syntaks og præcise kommandoer.

Senere blev grafiske brugerflader udbredt. Programmer som Microsoft Word eller Excel gjorde det muligt at arbejde med computere gennem knapper, menuer og ikoner. Brugeren behøvede ikke længere kende de bagvedliggende kommandoer, men skulle i stedet lære, hvad de forskellige funktioner i programmet gjorde.

I begge tilfælde var det imidlertid stadig brugeren, som skulle tilpasse sig maskinen. Man skulle lære programmets logik, funktioner og struktur. Når Microsoft udgiver en ny udgave af Word eller PowerPoint, så skal du lære, hvad de forskellige knapper gør. Til gengæld kan du være sikker på, at når du har lært hvad en knap eller kommando gør, så vil denne knap eller kommando gøre det samme hver gang.

Den amerikanske computerforsker Licklider formulerede allerede i 1960 en vision om et tættere samarbejde mellem mennesker og computere. I stedet for at mennesker skulle lære maskinens sprog, forestillede Licklider (1960) sig et system, hvor maskiner kunne arbejde sammen med mennesker og forstå deres intentioner. Denne vision er endnu ikke fuldt realiseret, men generative sprogmodeller repræsenterer et stort skridt i den retning



Figur 4 – 3 generationer af Human-Machine-Interface

Når du arbejder med en chatbot som ChatGPT, Copilot eller Gemini, kommunikerer du ikke længere gennem knapper eller programmeringssprog. I stedet bruger du dit almindelige sprog. Det betyder, at interaktionen mellem menneske og computer i højere grad ligner en samtale.

Denne ændring kan virke banal, men den har betydelige konsekvenser. Hvor tidligere software krævede, at brugeren lærte programmets struktur, kan generative sprogmodeller i mange tilfælde tolke brugerens hensigt ud fra sproget alene. Det betyder dog ikke, at systemet forstår intentioner på samme måde som et menneske. Som beskrevet tidligere, arbejder sprogmodeller stadig med statistiske sandsynligheder og mønstre i tekst.

Det er også derfor, at den måde brugeren formulerer sig på, får så stor betydning. I traditionelle programmer bestemmes resultatet primært af programmets funktioner. I generative sprogmodeller spiller brugerens formuleringer og den øvrige kontekst en langt større rolle.

Man kan derfor godt sige, at sproget selv er blevet en del af brugerfladen.

Dette er baggrunden for, at begrebet **prompting** er opstået. En prompt er i praksis den tekst, du skriver til systemet for at beskrive en opgave eller stille et spørgsmål. Hvis vi skal oversætte begrebet prompt engineering til dansk, så ville jeg sige, at det er "kunsten at prompte", hvilket kan defineres som "the practice of designing, refining, and implementing prompts or instructions that guide the output of large language models" (Mesko, 2023), eller sagt på jævnt dansk, hvordan du taler, så maskinen forstår dig.

For jo mere præcist du er i stand til at give en koncis instruktion, desto større er sandsynligheden for, at systemet genererer et brugbart svar. Dermed bliver evnen til at formulere instruktioner og i det hele taget rammesætte en opgave en central kompetence i arbejdet med generativ AI. Det er dog også vigtigt at bemærke, at præcise instrukser ikke garanterer perfekte resultater. Hvis du prompter ChatGPT om kun at bruge faktiske kilder, så vil den i højere grad gøre dette, men der vil stadigvæk kunne forekomme opfundne kilder.

Ordet prompt er i sin oprindelige betydning en instruktion eller et stikord, der sætter noget i gang. I forbindelse med generativ kunstig intelligens bruges begrebet om den tekst, som en bruger skriver til et AI-system for at få det til at udføre en opgave.

En prompt kan være et spørgsmål, en instruktion eller en mere omfattende opgavebeskrivelse.

Eksempelvis kan følgende være en prompt:

Forklar hvad inflation er.

Men en prompt kan også være langt mere detaljeret:

Forklar inflation i et kort notat på cirka 200 ord.
Målgruppen er studerende på en erhvervsakademiuddannelse uden økonomisk baggrund.
Brug et konkret eksempel fra hverdagen.

Begge eksempler er prompts, men de vil typisk føre til forskellige svar.

Den første prompt er kort og åben. Den giver systemet relativt stor frihed til selv at vælge perspektiv, niveau og struktur.

Den anden prompt indeholder derimod flere instruktioner om formål, målgruppe og format. Dermed reduceres usikkerheden i systemets fortolkning.

Dette illustrerer et centralt princip i arbejdet med generativ AI: jo mere præcist en opgave beskrives, desto større er sandsynligheden for et relevant svar.

Det betyder dog ikke, at en prompt nødvendigvis skal være meget lang. I mange tilfælde kan korte prompts fungere udmærket. Det afgørende er snarere, om prompten indeholder den information, som modellen har brug for til at forstå opgaven.

En nyttig måde at tænke på en prompt er derfor at se den som en opgavebeskrivelse.

Hvis du skulle forklare en opgave til en kollega, ville du typisk også forsøge at beskrive, hvad der skal laves, hvem det er til, og hvordan resultatet skal se ud.

På samme måde bør en prompt indeholde elementer som:

- **Opgaven**
- **Målgruppen**
- **Konteksten**
- **Formatet**

Disse elementer hjælper modellen med at identificere relevante mønstre i sin træning.

I praksis betyder det, at en prompt ofte fungerer bedst, når den giver systemet tilstrækkelig kontekst til at forstå situationen.

Dette er også grunden til, at prompting i mange tilfælde minder om arbejdet med problemformuleringer i akademiske opgaver. Før man kan formulere en god prompt, må man ofte først overveje, hvad problemet egentlig er.

I første omgang er det imidlertid tilstrækkeligt at forstå, at en prompt er en instruktion til et system, der arbejder med statistiske mønstre i tekst.

Prompt engineering

Der synes generelt at være et behov for at lære, hvordan man bruger ChatGPT og andre generative værktøjer. Her vil det være oplagt at starte med at kigge på den vejledning som OpenAI (2024) selv har udgivet for prompt engineering tilbage i 2024. Vejledningen virkede – som med meget andet materiale om prompt engineering – ustruktureret og mangler kildehenvisninger, men er stadigvæk et logisk sted at starte. Vejledningen opstiller de 6 følgende ”strategier” for ”at få bedre resultater”, som hver især uddybes med 18 underliggende ”taktikker”.

Tabel 2 - OpenAI's råd til prompting

#	Strategi	Taktikker
1	Skriv klare instrukser	<ul style="list-style-type: none"> a. Inkluder detaljer i din forespørgsel b. Bed modellen om at adoptere en persona c. Brug afgrænsningstegn til at adskille dele af inputtet d. Angiv de nødvendige trin for at fuldføre en opgave e. Kom med eksempler f. Angiv den ønskede længde af output
2	Angiv reference tekst	<ul style="list-style-type: none"> a. Instruer modellen til at svare ved hjælp af en referencetekst b. Instruer modellen til at svare med citater fra en referencetekst
3	Opdel komplekse opgaver	<ul style="list-style-type: none"> a. Brug emneopdeling til at strukturere relevante informationer b. For lange samtaler, opsummer tidligere dialog c. Opsummer lange dokumenter stykkevis og konstruer et komplet resumé rekursivt
4	Giv modellen tid til at ”tænke”	<ul style="list-style-type: none"> a. Instruer modellen til at udarbejde sin egen løsning, før den skyn-der sig til en konklusion b. Bed om modellens indre monolog eller en række forespørgsler til at afdække modellens ræsonnement c. Spørg modellen, om den gik glip af noget ved tidligere gennemløb
5	Brug eksterne værktøjer	<ul style="list-style-type: none"> a. Brug indlejningsbaseret søgning til at implementere vidensøgning b. Brug kodeudførelse til at udføre mere nøjagtige beregninger c. Giv modellen adgang til specifikke funktioner
6	Test ændringer systematisk	<ul style="list-style-type: none"> a. Evaluer modeloutput ud fra, hvad du vurderer, som de bedste svar

I 2025 fjernede OpenAI dog denne vejledning uden at give en egentlig begrundelse, men anførte at deres nyere model selv kunne finde ud af at gøre alle disse ting, eller bede om dem, såfremt modellen vurderede, at det var påkrævet. Imidlertid virker det som om at flertallet af taktikkerne stadigvæk har stor relevans. Generelt kan man de nyere modeller fra OpenAI og andre synes at have systemprompter, altså en fast instruks, der opfordrer modellen til at gøre mange af disse ting. Imidlertid kan man jo stadigvæk få mere af det, hvis man eksplicit prompter efter det, hvorfor taktikkerne stadigvæk synes at have deres berettigelse. Lad os kort gennemgå de seks strategier.

Strategi 1: Skriv klare instrukser

Den første strategi er også den mest grundlæggende. Hvis en prompt er uklar, vil modellen ofte producere et generelt eller upræcist svar. Derfor anbefaler OpenAI, at man forsøger at formulere så klare instruktioner som muligt.

En vigtig forklaring på dette er, at generative AI-systemer mangler den sociale kontekst, som normalt hjælper mennesker med at forstå hinanden.

Hvis en studerende under en forelæsning spørger en underviser: "Kan du lige uddybe det?", så vil underviseren typisk forstå spørgsmålet ud fra situationen. Der sidder måske 100 studerende i lokalet, undervisningen er i gang, og spørgsmålet refererer til noget, der netop er blevet sagt. I den situation vil det være oplagt at give en forklaring på måske 20-30 sekunder. Et enkelt ord vil være for kort, mens en ti minutters afstikker vil være for lang.

Denne vurdering sker næsten automatisk, fordi mennesker konstant aflæser sociale signaler. En AI-model har ikke adgang til denne kontekst. Den ved ikke, om du står i et auditorium, skriver en rapport eller blot er nysgerrig. Derfor er det ofte nødvendigt at specificere ting, som mennesker normalt tager for givet. Det gælder specifikt, hvor langt svaret skal være, men også hvad opgaven er, hvem målgruppen er, hvilket format svaret skal have og så videre.

Forklar inflation i et kort notat på cirka 200 ord.
Målgruppen er studerende på en erhvervsakademiuddannelse uden økonomisk baggrund.
Brug et konkret eksempel fra hverdagen.

Strategi 2: Angiv referencetekst

Den anden strategi handler om at give modellen adgang til den tekst eller det materiale, som svaret skal baseres på.

Sprogmodeller er trænet på meget store mængder tekst, men de har ikke adgang til en fuldstændig og opdateret database over verden. Hvis man ønsker et svar baseret på en bestemt artikel, rapport eller tekst, er det derfor ofte en fordel at indsætte denne tekst direkte i prompten.

Hvis man for eksempel skriver:

Opsummer hovedpointerne i følgende artikel.

og derefter indsætter artiklen, vil modellen typisk kunne give et mere præcist svar, end hvis man blot beskriver artiklen i generelle termer.

Denne strategi minder i høj grad om arbejdet med kilder i akademiske opgaver. Hvis analysen skal baseres på en bestemt tekst, er det nødvendigt at arbejde direkte med den. Det bør dog også bemærkes, at RettighedsAlliancen mener, at det er et brud på ophavsretten at vedlægge materiale, som du ikke selv har skabt til en kunstig intelligens.

Strategi 3: Opdel komplekse opgaver

Den tredje strategi bygger på en velkendt indsigt fra både problemløsning og programmering: komplekse opgaver bliver ofte lettere at løse, hvis de opdeles i mindre dele.

Hvis man beder en AI om at løse en meget omfattende opgave i én prompt, kan resultatet blive upræcist eller overfladisk. Det kan derfor være en fordel at opdele opgaven i flere trin.

Et eksempel kan være arbejdet med en længere tekst. I stedet for at bede modellen om at opsummere et dokument på 50 sider i én omgang, kan man først bede den opsummere hvert afsnit og derefter samle disse delresuméer til en samlet konklusion. Dette skyldes noget meget jordnært ved simple sprogmodeller, nemlig begrænsningen i inputtets længde. Beder du en simpel model om at opsummere en tekst på 200 sider, så vil simple modeller typisk kun læse de første 20-30 sider og ud fra det gætte på resten.

På samme måde kan man i længere samtaler med en AI bede modellen om at opsummere den tidligere dialog. Det kan hjælpe med at holde overblik og sikre, at modellen fortsat arbejder ud fra de samme præmisser.

Strategi 4: Giv modellen

Den fjerde strategi handler om at få modellen til at arbejde mere systematisk med en opgave, før den formulerer et svar.

Hvis man blot stiller et spørgsmål, vil modellen ofte forsøge at producere et svar så hurtigt som muligt. I nogle tilfælde kan det være en fordel i stedet at bede modellen gennemgå et ræsonnement trin for trin.

For eksempel kan man skrive:

Forklar først hvilke faktorer der påvirker inflation.
Giv derefter en samlet forklaring.

På den måde nudges modellen til først at strukturere problemet og derefter formulere konklusionen. Denne tilgang bruges ofte i mere analytiske opgaver, hvor svaret afhænger af flere mellemlid.

Strategi 5: Brug eksterne værktøjer

Den femte strategi handler om at kombinere sprogmodeller med andre systemer.

Generative sprogmodeller er gode til at formulere tekst, men de er ikke nødvendigvis gode til præcise beregninger eller til at slå specifik information op i databaser. Det er en kendt sag, at selv de modeller af CoPilot, Claude og ChatGPT, som er udviklet i 2026 stadigvæk ikke kan tælle, medmindre at deres indbyggede matematikmodul bliver aktiveret.

Derfor bør du anvende AI-modeller til tekst sammen med andre mere specialiserede AI-værktøjer til udregning, grafik, lyd og så videre. Det kan også være søgesystemer eller databaser.

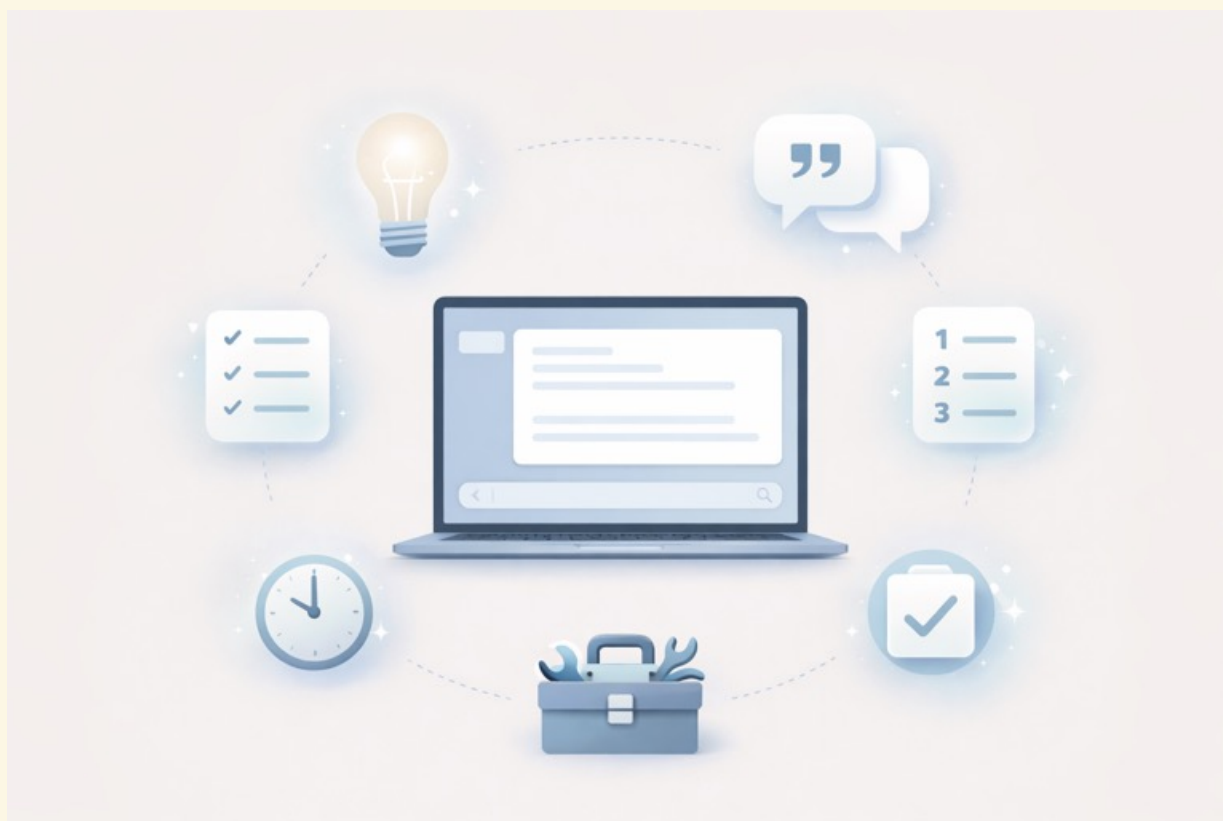
Strategi 6: Test ændringer systematisk

Den sidste og måske mest oversete strategi handler om at evaluere og forbedre sine prompts.

Hvis man arbejder systematisk med generativ AI, vil man ofte opdage, at små ændringer i formuleringen kan føre til forskellige resultater. Det kan derfor være nyttigt at teste forskellige versioner af en prompt og sammenligne svarene – dog altid med det forløb, at en generativ AI har en naturlig variation i sine svar.

Ved denne strategi minder arbejdet med prompts lidt om klassiske eksperimenter eller A/B-test. Man ændrer én ting ad gangen og vurderer, om resultatet bliver bedre eller dårligere.

Denne tilgang er også baggrunden for, at prompting i stigende grad bliver betragtet som en proces frem for en enkelt handling.



Figur 5 - OpenAI's strategier som tankeproces

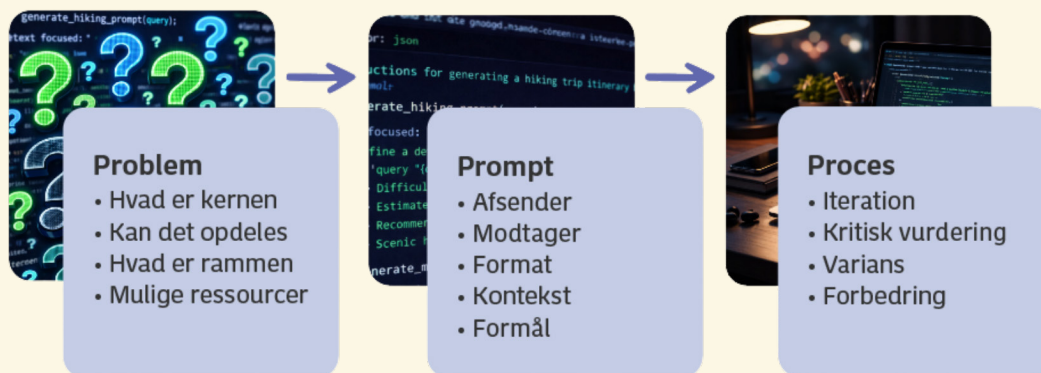
Prompting som problemforståelse

Mange introduktioner til prompt engineering fokuserer på teknikker. Man får lister over taktikker til, hvordan man kan formulere instruktioner mere præcist, give modellen eksempler eller opdele en opgave i mindre dele. Disse råd kan være en god begyndelse, men de har også en stor begrænsning. De siger nemlig relativt lidt om, hvorfor nogle prompts fungerer bedre end andre.

Den engelske forsker Acar (2023) argumenterer imod, at vi anser prompt engineering som en videnskab, men ser det mere som et flygtigt begreb, der hastigt bliver forældet, fordi de generative AI-modeller vil blive udviklet til at stille spørgsmål om de ting, som gør et svar bedre (ud fra brugerens subjektive forståelse) og dermed aftvinge en spørger detaljer om kontekst, persona, reference, længde etc., før den afgiver et svar.

I stedet for at kigge på prompt engineering, så argumenterer Acar (2023) for, at man bør gribe fat i det velkendte akademiske begreb problemforståelse, altså at du bruger tid på at overveje, hvad det er for et problem, som du er i gang med at løse, specifikt at du overvejer kernen, opdelingen, rammen og ressourcerne for det problem, som du er i gang med at løse. Man kan sige, at Acars budskab er, at før du prompter, så skal du tænke over problemet.

Denne proces kan groft opdeles i tre trin: problem, prompt og proces.



Figur 6 - Problem -> Prompt -> Proces

Før du skriver en prompt, må du først forstå, hvad det egentlig er, du forsøger at finde ud af. Hvis problemet er uklart, vil selv en meget detaljeret prompt sjældent føre til et godt resultat. Omvendt vil en relativt simpel prompt ofte være tilstrækkelig, hvis du har en god og veldefineret forståelse af opgaven.

Den første del af modellen handler derfor om selve problemet. Når du altså arbejder med en opgave, kan det være nyttigt at stille nogle grundlæggende spørgsmål:

- **Hvad er kernen i problemet?**
- **Kan problemet opdeles i mindre dele?**
- **Hvad er rammen for opgaven?**
- **Hvilke ressourcer eller informationer har du til rådighed?**

Disse spørgsmål ligner i høj grad de overvejelser, du normalt gør dig i forbindelse med løsningen af komplekse opgaver.

Den næste del af modellen handler om selve prompten.

Når problemet er klart, skal det oversættes til en instruktion, som en sprogmodel kan arbejde med. I praksis betyder det, at du må overveje en række kommunikative elementer:

- **Hvem er afsenderen af teksten?**
- **Hvem er modtageren?**
- **Hvilket format skal svaret have?**
- **Hvilken kontekst skal modellen arbejde ud fra?**
- **Hvad er formålet med svaret?**

Disse elementer hjælper modellen med at aktivere de mest relevante mønstre i sin træning.

Den sidste del af modellen handler om processen.

Arbejdet med generativ AI består sjældent i at skrive én prompt og derefter modtage et færdigt svar. I praksis foregår arbejdet næsten altid gennem en række iterationer, hvor du gradvist forbedrer både prompten og resultatet.

I denne fase spiller flere faktorer en rolle:

- **Iteration**
- **Kritisk vurdering**
- **Varians i modellens svar**
- **Forbedring af prompten**

På den måde minder arbejdet med generativ AI i høj grad om en analytisk proces, hvor man løbende tester og justerer sin tilgang.

Pointen her er, at du ikke kun skal se prompting som en teknik til at formulere stikord til en AI, men som en metode til at strukturere problemløsning. Når du arbejder systematisk med generativ AI, begynder arbejdet derfor ikke med at skrive en prompt, men med at forstå problemet.

Men lad os gå til det første skridt, altså at forstå problemet. Her peger Acar især på rammesætning af problem, den underliggende grund for problemet, opdelingen af problemet og perspektivet på problemet

Tabel 3 - Problemforståelse

Problemaspekt	Beskrivelse	Teknik
Underliggende årsag	Finde frem til hvad kernen af problemet er	RCA - Root Cause Analysis - Spørg hvorfor 5 gange
Opdeling	Finde ud af hvordan problemet kan opdeles	MECE – mutually exclusive commonly exhaustive – opdeling uden overlap, men udtømmende i fælleskab
Perspektiv	Finde ud af hvilken bevidst eller ubevidst rammer som man tænker problemet ind i	Reframing – f.eks. ved en langsom elevator, er problemet så at elevatoren er langsom eller at den opleves langsom.
Rammesætning	Kvantificere problemet så størrelse af det bliver klart	SMART – Specifikt, Målbart, Allokerbart, Realistisk & Tidsrelateret – angiv hvilket resultatet opnås hvornår af hvem

Rammesætning

For at rammesætte et problem, så kan du med fordele bruge klassiske metodetjeklister, såsom SMART-modellen eller Guldlok-testen.

En af de mest udbredte modeller til at formulere mål og problemstillinger er SMART-modellen. Den blev oprindeligt introduceret i i begyndelsen af 1980'erne (Doran, 1981).

SMART står typisk for:

- **Specifik**
- **Målbart**
- **Opnåelig**
- **Relevant**
- **Tidsafgrænset**

blev oprindelig udviklet til at formulere klare årlige mål for afdelinger i større organisationer, men den kan også anvendes i arbejdet med prompts.

Modellen blev oprindelig udviklet til at formulere klare årlige mål for afdelinger i større organisationer, men principperne kan også bruges mere generelt i arbejdet med problemformulering.

Hvis et problem eksempelvis formuleres meget bredt, som for eksempel:
Hvordan kan en virksomhed blive mere succesfuld?

Så er det svært at arbejde analytisk med spørgsmålet. Hvis problemet derimod afgrænses, bliver det mere håndterbart:

Hvordan kan virksomheden øge salget på det danske marked i løbet af de næste 12 måneder?

Her bliver problemet både mere specifikt og mere tidsafgrænset.

Den såkaldte Guldlok-test bygger på en lignende idé. Navnet henviser til eventyret om Guldlok og de tre bjørne, hvor noget hele tiden er enten for stort, for småt eller lige tilpas. I metodeundervisning bruges metaforen til at illustrere, at en problemformulering hverken må være for bred eller for snæver (Saunders, Lewis & Thornhill, 2013).

Hvis problemet er for bredt, bliver analysen ofte overfladisk. Hvis det er for snævert, kan det blive vanskeligt at finde relevant information. En god problemformulering ligger derfor typisk et sted midt imellem.

Underliggende årsager

Når et problem er blevet rammesat, er næste skridt at undersøge de underliggende årsager. I mange situationer viser det sig, at det problem man først identificerer, i virkeligheden blot er et symptom på noget andet.

Et klassisk værktøj til dette er root cause analysis, som anvendes i mange forskellige sammenhænge inden for organisation, kvalitetssikring og ingeniørarbejde (Moaveni & Chou, 2017).

En af de mest kendte teknikker er den såkaldte "five whys"-metode. Her forsøger man gentagne gange at spørge hvorfor et problem opstår, indtil man når frem til en mere grundlæggende forklaring.

Hvis en virksomhed eksempelvis oplever faldende salg, kan analysen udvikle sig i flere trin:

Hvorfor falder salget?

Fordi færre kunder besøger webshoppen.

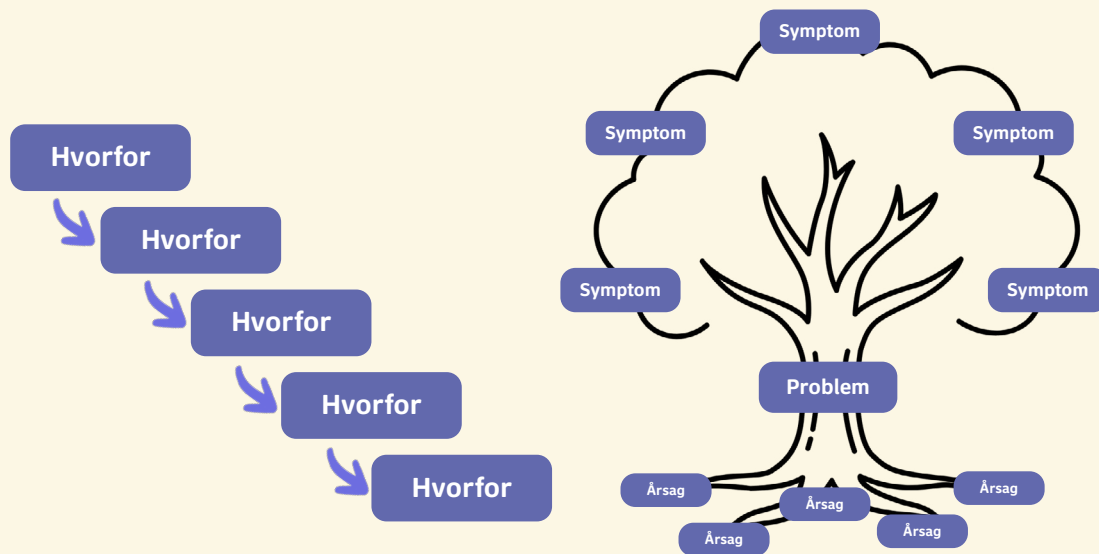
Hvorfor besøger færre kunder webshoppen?

Fordi trafikken fra søgemaskiner er faldet.

Hvorfor er trafikken fra søgemaskiner faldet?

Fordi konkurrenter i stigende grad dominerer søgeresultaterne.

På denne måde bevæger analysen sig gradvist fra symptomer til mere strukturelle årsager.



Figur 7 - Underliggende problemanalyse

Opdeling

Det tredje aspekt handler om at opdele problemet i mindre dele. Mange komplekse problemer bliver lettere at arbejde med, hvis de struktureres i overskuelige kategorier.

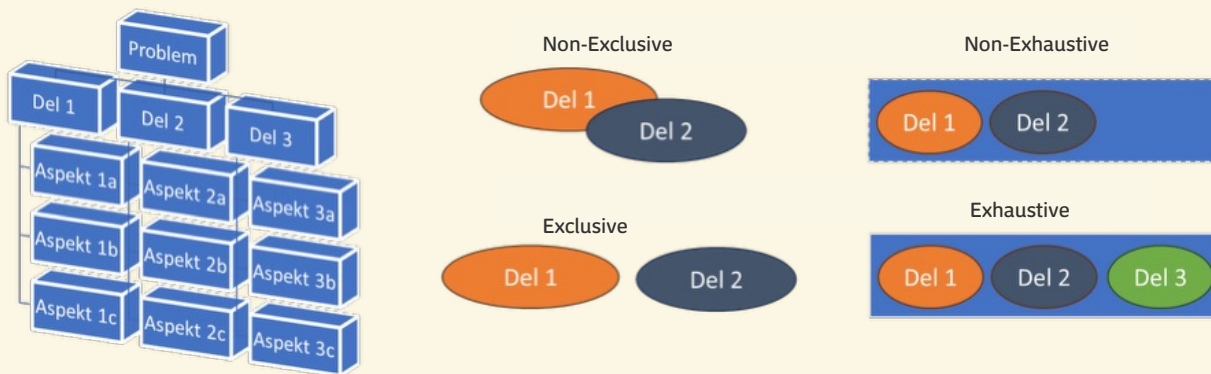
Et klassisk princip til dette er MECE-princippet, som blev populariseret af Barbara Minto i forbindelse med arbejdet i konsulentfirmaet McKinsey (Minto, 1996).

MECE står for Mutually exclusive, Collectively exhaustive, hvad kan oversættes til dansk som ikke-overlappende, men til sammen udtømmende. Ideen er, at et problem bør opdeles i kategorier, der ikke overlapper hinanden, men som tilsammen dækker hele problemfeltet.

Hvis man eksempelvis analyserer årsagerne til faldende salg, kan problemet opdeles i tre overordnede områder:

- **marked**
- **produkt**
- **distribution**

Ved at strukturere problemet på denne måde bliver ens analyse mere systematisk og dermed lettere for andre at overskue.



Figur 8 - Problemopdeling

Perspektiv

Det sidste aspekt handler om perspektivet på problemet. Ofte viser det sig, at et problem kan forstås på flere forskellige måder, afhængigt af hvordan det formuleres. Denne proces kaldes reframing.

Ifølge Wedell-Wedellsborg (2020) er en af de mest almindelige fejl i problemløsning, at man forsøger at løse det forkerte problem. Hvis problemet er defineret forkert fra begyndelsen, vil selv meget grundige analyser ikke nødvendigvis føre til gode løsninger.

Et klassisk eksempel handler om en elevator i en kontorbygning. Brugerne klager over, at elevatoren er for langsom. Hvis man accepterer denne problemformulering, vil løsningen næsten automatisk handle om at gøre elevatoren hurtigere. Man kan installere en ny elevator, opgradere motoren eller forbedre styringsalgoritmen.

Alle disse løsninger kan være både teknisk komplicerede og økonomisk dyre.

Hvis man derimod ændrer perspektivet på problemet, kan situationen se anderledes ud. Måske er det egentlige problem ikke elevatorens hastighed, men at ventetiden føles lang og irriterende for brugerne.

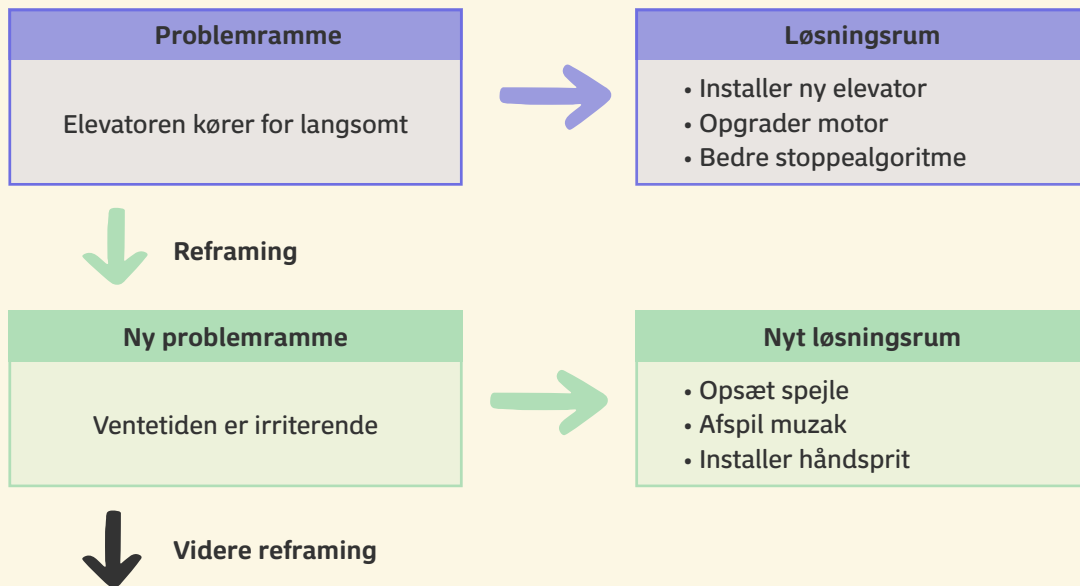
Problemet kan derfor formuleres på en anden måde:

Hvordan kan vi få ventetiden til at føles kortere?

Når problemet formuleres på denne måde, åbner der sig pludselig andre typer løsninger. I mange bygninger har man eksempelvis opsat spejle ved elevatorerne. Det giver folk noget at kigge på, mens de venter. Andre steder har man installeret musik eller små informationskærme.

Elevatoren er ikke blevet hurtigere, men ventetiden opleves alligevel som kortere.

Pointen med eksemplet er, at en ændring i det perspektiv som du har på problemet kan føre til helt andre typer løsninger. Reframing handler derfor ikke om at finde en bedre løsning på det samme problem, men om at undersøge, om problemet i virkeligheden kan eller bør forstås på en anden måde.



Figur 9 - Reframing

Samlet set er pointen med disse fire tilgange, at arbejdet med problemer i høj grad starter med at forstå hvad det problem eller den opgave, som du prøver at løse, egentlig er. Før man forsøger at løse et problem, må man først afgrænse det, forstå dets årsager, opdele det i overskuelige dele og overveje, om problemet er formuleret på den rigtige måde. Hvis du går i den forkerte retning, så er det ligegyldigt, hvor hurtigt du går eller løber – du kommer stadigvæk ikke tættere på dit mål.

Promptens struktur

Når man arbejder med generativ kunstig intelligens, kan det være fristende at se prompting som en teknisk disciplin. Mange beskrivelser af prompt engineering fokuserer på tricks, remser, formuleringer eller bestemte syntaktiske greb. Men kvaliteten af en prompt handler i høj grad om, hvor klart du kan formulere dig. Hvis vi antager, at teknologien nu er kommet dertil, hvor vi har maskiner, der forstår menneskers naturlige sprog, så skaber det jo en opgave for det enkelte menneske om at kunne formulere sig klart og tydeligt.

Det betyder i praksis, at en god prompt minder om god tale. Før man formulerer teksten, må man overveje, hvem der taler, hvad der skal opnås, hvem modtageren er, og hvordan resultatet skal præsenteres. For nogle mennesker kommer det at formulere sig klart intuitivt og naturligt, men der findes teknikker til at blive bedre til det.

Tabel 4 - 6 aspekter for klarhed

#	Aspekt	Spørgsmål	Eksempel
1	Rolle	Hvilken vinkel eller faglig persona?	Rådgiver, coach, redaktør
2	Formål	Hvad skal modtageren få ud af teksten?	Overblik, beslutning, inspiration
3	Format	Outputtype	Notat, liste, opslag, mail, tekstforslag
4	Målgruppe	Hvem er det til – og hvad ved de?	Leder, kollega, kunde, borger
5	Indhold	Hvad skal med?	Fakta, stikord, bilag, budskaber
6	Tone/stil	Hvordan skal det lyde?	Sagligt, entusiastisk, pædagogisk, kortfattet

Rolle

Det første element handler om rollen eller perspektivet. Når du angiver en rolle, beskriver du hvilken faglig eller professionel position, som modellen skal indtage.

Eksempelvis kan modellen instrueres til at agere som rådgiver, redaktør eller coach. Dette påvirker typisk både sproget og indholdet i svaret. Hvis en prompt eksempelvis blot lyder:

Forklar hvordan man reducerer stress på arbejdspladsen.

vil svaret ofte blive generelt. Hvis rollen i stedet præciseres:

Du er arbejdsmiljørådgiver. Forklar tre konkrete tiltag til at reducere stress i en mindre virksomhed.

så ændres både perspektivet og graden af konkretisering.

Formål

Det næste element handler om formålet. I mange tilfælde bliver AI-svar upræcise, fordi formålet med outputtet ikke bliver tydeligt angivet.

En tekst kan have mange forskellige funktioner. Den kan skabe overblik, understøtte en beslutning eller give inspiration.

Et eksempel kan være en organisation, der arbejder med grøn omstilling. En prompt kunne lyde:

Giv et overblik over tre mulige initiativer til at reducere energiforbrug i en kontorbygning.

Her bliver formålet netop at skabe overblik. Hvis formålet i stedet er beslutning, kan prompten formuleres andersledes.

Vurder hvilke tre energitiltag der sandsynligvis giver størst effekt på kort sigt.

Format

Formatet handler om, hvordan svaret skal struktureres. Generative modeller kan producere mange forskellige typer tekst, men uden instruktion vil modellen selv vælge en struktur.

En prompt kan derfor med fordel præcisere outputtypen.

Eksempelvis

Skriv en kort mail til en kunde, der forklarer en forsinkelse i leveringen.

eller

Lav en liste med fem forslag til aktiviteter til en teambuilding-dag.

Ved at angive formatet bliver det lettere for modellen at generere et svar, der passer til den konkrete situation.

Målgruppe

Målgruppen er et klassisk element i kommunikationsteori. Hvem teksten er skrevet til, har stor betydning for både sprog og indhold.

En forklaring til en ekspert vil typisk være anderledes end en forklaring til en nybegynder.

Et eksempel kunne være en offentlig organisation, der arbejder med digital selvbetjening. To prompts kan se næsten ens ud:

Forklar hvordan man bruger MitID.

eller

Forklar hvordan man bruger MitID til ældre borgere uden stor digital erfaring.

I det sidste tilfælde vil svaret typisk blive mere pædagogisk og konkret.

Indhold

Indholdselementet er, hvilke informationer modellen skal arbejde ud fra.

Hvis en prompt ikke indeholder nogen form for baggrund, vil modellen typisk generere et generelt svar. Hvis man derimod angiver stikord, fakta eller kontekst, kan svaret blive mere præcist.

Et eksempel kunne være en organisation, der ønsker hjælp til en jobannonce. En prompt kan indeholde følgende information:

Lav et jobopslag

Virksomhed: mindre IT-konsulenthus

Stilling: projektleder

Erfaring: 3-5 år

Fokus: kundekontakt og koordinering

Når disse oplysninger indgår i prompten, får modellen et mere klart udgangspunkt.

Tone og stil

Det sidste element handler om hvordan teksten skal lyde. Sprog kan være formelt, entusiastisk, pædagogisk eller kortfattet.

Dette påvirker ikke nødvendigvis indholdet, men det påvirker hvordan teksten opleves.

Et eksempel kan være en organisation, der skal skrive et opslag til sociale medier. En prompt kunne lyde:

Skriv et kort LinkedIn-opslag i en professionel men engageret tone om en ny medarbejder.

Hvis man derimod ønsker en mere neutral tekst, kan toneangivelsen ændres.

Den klassiske baggrund

Allerede i antikken beskrev retorikere som Cicero, hvordan en taler må analysere situationen før en tekst eller tale formuleres. I klassisk retorik indgik blandt andet spørgsmål om hvem der taler, hvem der er modtager, hvad formålet er, og hvilken stil der bør anvendes.

Retorikken beskrev også arbejdsfaser i udformningen af en tekst, hvor taleren først analyserer situationen og derefter vælger argumenter, struktur og stil. Et eksempel på en kommunikationsmodel, der er over 2000 år, er 'Ciceros pentagram' (Roer, 2023), der er baseret på Ciceros beskrivelse af en god tale i bogen Om taleren (Cicero, 55 f.v.t.):

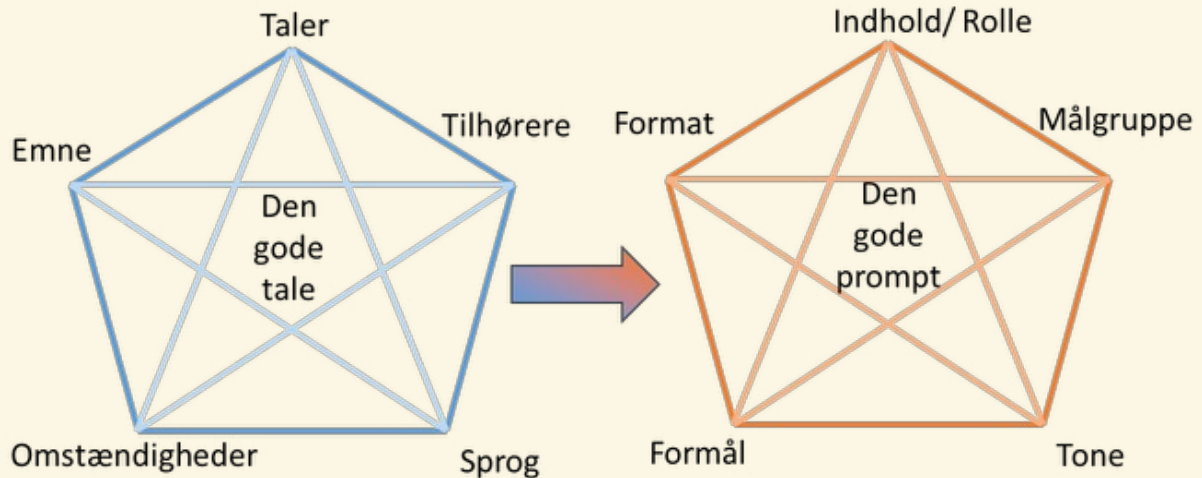
"Siden derfor alle aspekter af spørgsmålet om talens kunstneriske udformning er, om ikke fuldstændigt klarlagt, så dog i det mindste påpeget, så lad os nu se lidt nærmere på begrebet 'det behørig' [aptum], dvs. spørgsmålet om hvad der passer bedst i en tale. Det er jo på forhånd klart, at der ikke findes én og kun én tale-stil [genus-orationis], som passer til enhver sag [causae] eller ethvert publikum [auditori], eller klæder enhver taler [personae] ved enhver lejlighed [tempori]." (som gengivet af Roer, 2023).

Når mennesker kommunikerer med andre mennesker, virker en tydelig struktur ofte bedre, fordi modtageren oplever budskabet som mere målrettet. Hvis en tekst tydeligt adresserer målgruppen og formålet, bliver den lettere at forstå.

Når man kommunikerer med en AI, er mekanismen en anden.

Her handler det ikke om oplevelse eller overbevisning. Sprogmodellen forsøger hele tiden at beregne, hvilke ord der statistisk passer bedst i den aktuelle kontekst. Når en prompt tydeligt angiver rolle, målgruppe, format og formål, hjælper det modellen med at aktivere de sproglige mønstre i træningsdataene, der passer til netop denne situation.

Derfor kan de samme kommunikationsprincipper fungere i begge tilfælde, men af forskellige grunde.



Figur 10 - Fra god tale til god prompting

Afslutningsvis er det vigtigt ikke at læse tabellen ovenfor som en mekanisk opskrift. De seks aspekter er ikke tænkt som en tjekliste, hvor hvert punkt nødvendigvis skal være med i enhver prompt.

I mange situationer vil nogle af elementerne være afgørende, mens andre er mindre relevante. Hvis du eksempelvis beder en AI om at formulere en kort mail, kan rolle og målgruppe være centrale, mens formatet næsten giver sig selv. I andre situationer kan indhold og kontekst være langt vigtigere end tone eller stil.

Tabellen fungerer derfor bedst som en inspirationsliste eller et analytisk værktøj. Den kan hjælpe dig med at overveje, hvilke aspekter af kommunikationen der er relevante i netop den situation, du står i.

Det betyder også, at målet ikke er at lære en bestemt formel udenad.

I de senere år er der opstået en række frameworks for prompting, ofte præsenteret som korte akronymer eller huskeregler. De kan være nyttige som pædagogiske redskaber, fordi de gør det lettere for begyndere at huske nogle centrale principper. På samme måde lærer studerende i mange fag forskellige modeller og forkortelser, der hjælper med at strukturere deres arbejde.

Problemet opstår først, hvis sådanne huskeregler bliver forvekslet med selve metoden.

Der findes ingen dokumentation for, at bestemte akronymer eller faste skabeloner i sig selv producerer bedre svar fra en sprogmodel. Generative AI-systemer reagerer ikke på remser eller forkortelser som sådan. De reagerer på sproglig kontekst.

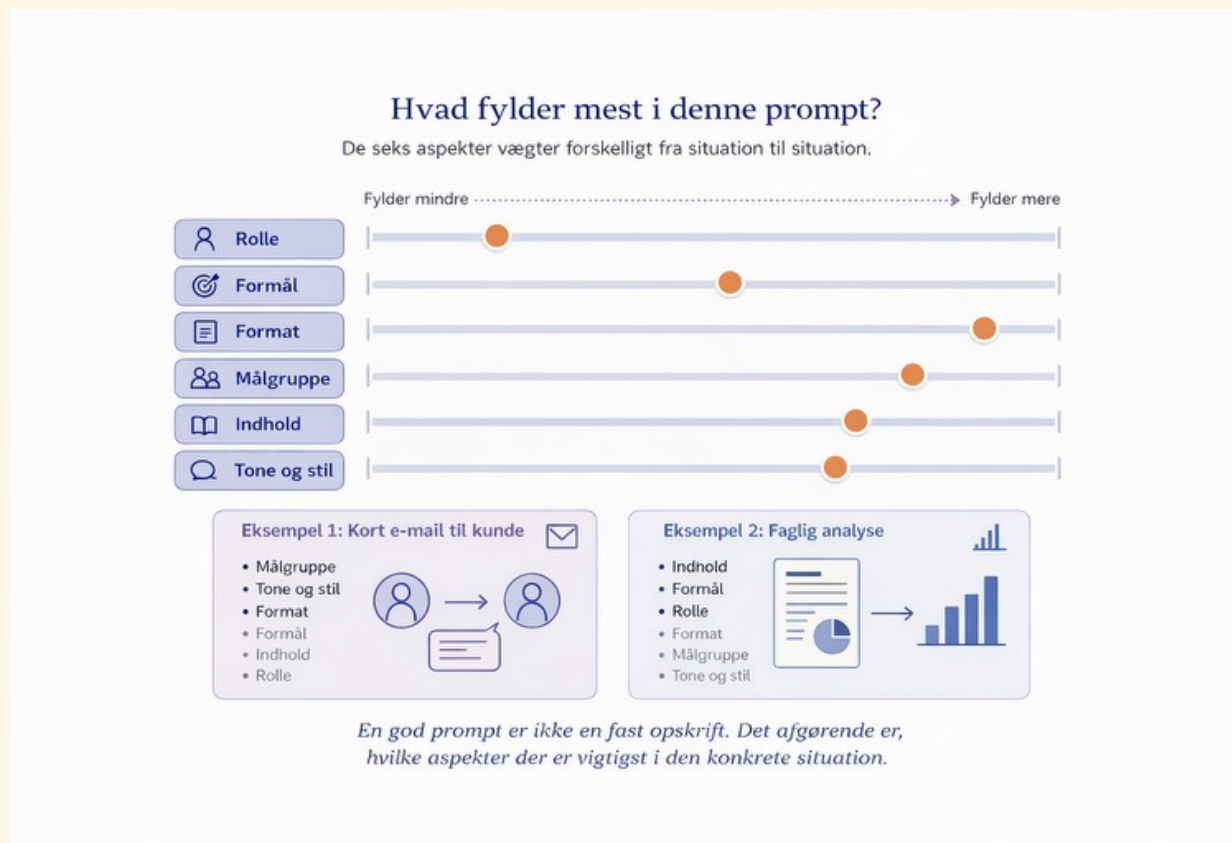
Derfor er det vigtigere at formulere sig klart og præcist end at forsøge at presse sin prompt ind i en bestemt skabelon.

På samme måde som i klassisk retorik er det afgørende spørgsmål ikke, om en tekst følger en bestemt formel, men om den passer til situationen. Hvad der er en god formulering i én sammenhæng, kan være mindre egnet i en anden.

De seks aspekter i dette kapitel bør derfor forstås som en måde at tænke over kommunikationen på – ikke som en opskrift, der altid skal følges slavisk.

Når man arbejder med generativ AI, kommer kvaliteten af en prompt i sidste ende ikke fra en huskeregel, men fra evnen til at forstå problemet, tænke over modtageren og formulere sig klart.

Det er i virkeligheden den samme kompetence, der i mere end to tusind år har været kernen i god kommunikation.



Figur 11 - Hvilket aspekt er vigtigst

AI som proces

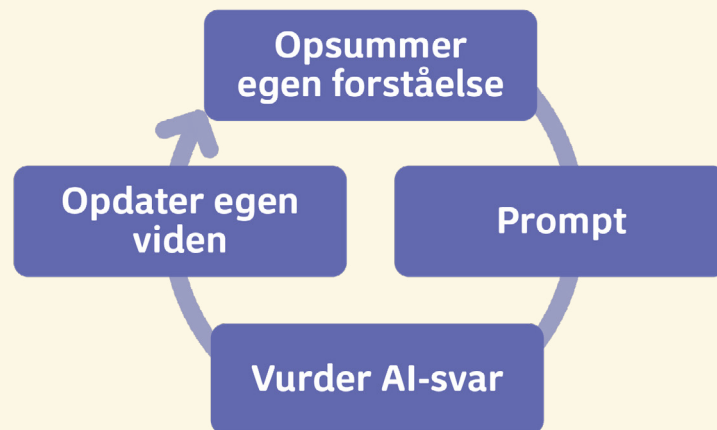
I det foregående har der været et fokus på, hvordan du formulerer en prompt. I praksis stopper arbejdet med generativ kunstig intelligens imidlertid sjældent ved den første prompt. De fleste brugere oplever hurtigt, at arbejdet med AI udvikler sig til en proces, hvor prompts, svar og brugerens egen forståelse påvirker hinanden.

Denne proces kan beskrives gennem en hermeneutisk tilgang.

Hermeneutik er en fortolkningsteori med rødder i filosofi og humaniora. Grundideen er, at forståelse ikke opstår i ét enkelt trin, men gennem en cirkulær bevægelse mellem spørgsmål, fortolkning og ny forståelse. Denne proces kaldes ofte den hermeneutiske cirkel og er blandt andet beskrevet af Gadamer (2004).

Når man arbejder med generativ AI, opstår der en lignende dynamik. Brugeren formulerer en prompt, modtager et svar, vurderer svaret og justerer derefter sin egen forståelse. Den nye forståelse fører til en ny prompt, som igen producerer et nyt svar.

Denne proces kan beskrives som **AI-hermeneutik**.



Figur 12 - AI-hermeneutik

Modellen kan forstås som en cyklisk proces bestående af fire trin: Opsummering af egen forforståelse, skrivning af en prompt, vurdering af et AI-svar, og så en opdatering af brugerens egen forståelse.

Først formulerer du din egen foreløbige forståelse af problemet. Inden du skriver en prompt, kan det være nyttigt at gøre sig klart, hvad du allerede ved om emnet, og hvilket svar du forventer at få. I hermeneutikken kaldes dette forforståelse. Den fungerer som et bevidst udgangspunkt for din fortolkning. Nogle gange er det nok at stoppe op og bruge 10 sekunder på at tænke det igennem, andre gange kan en bevidst nedskreven brug af metoderne til problemforståelse være nødvendig.

Derefter formulerer du en prompt. Prompten er i praksis et forsøg på at teste eller udfordre denne foreløbige forståelse ved at stille et spørgsmål til systemet.

Herefter genererer modellen et svar. Svaret er baseret på statistiske mønstre i træningsdata og den kontekst, der er angivet i prompten. Dette svar bør du naturligvis vurdere og fortolke ud fra din eksisterende viden. I praksis kan dette let reduceres til et spørgsmål, om du synes, at svaret ser godt ud. Dette er en fin forsimpning, hvis det er et område, som du let kan vurdere enten fordi det er relativt simpelt eller også fordi du har stor erfaring inden for det. Ved komplekse opgaver, hvor du mangler indsigt, så er opgaven her at finde ud af, hvordan du vurderer og ret beset kvalitetssikrer output fra AI'en. Er det at gå ind i de kilder, som svaret er baseret på? Kan du generere et andet svar i et andet system og så sammenligne?

Det sidste skridt er at vurdere om svaret er godt nok til at gå videre. Hvis du vurderer, at svaret virker plausibelt, kan det udvide eller nuancere din forståelse af problemet. Hvis svaret virker tvivlsomt, kan du overveje om din problemforståelse bør revideres.

På den måde opstår en cirkulær proces, hvor din forståelse løbende påvirkes af AI-systemets svar. Det er netop denne bevægelse mellem forforståelse, spørgsmål og fortolkning, der ligger til grund for betegnelsen AI-hermeneutik, hvor AI ikke blot er et opslagsværktøj, men en samtalepartner i din fortolkningsproces.

Denne forståelse ligger i forlængelse af nyere forskning i menneske-AI-interaktion, hvor AI-systemer ofte beskrives som samarbejdspartnere i en iterativ proces mellem menneskelig vurdering og maskinel generering af forslag (Shneiderman, 2022).

Vibe coding og ændrede arbejdsroller

Den procesorienterede brug af AI bliver særligt tydelig inden for softwareudvikling.

I februar 2025 introducerede Andrej Karpathy udtrykket "vibe coding". Begrebet beskriver en arbejdsform, hvor softwareudviklere i stigende grad arbejder gennem prompts til en AI i stedet for at skrive softwarekoden selv. Karpathy beskrev selv processen således:

"It's not really coding — I just see things, say things, run things, and copy-paste things, and it mostly works."

Pointen er ikke, at programmering forsvinder, men at rollen ændrer karakter. Programmøren bevæger sig fra at være primær producent af kode til i højere grad at fungere som kurator, prompt-designer og kritisk evaluator.

Denne forskydning kan også observeres i andre former for vidensarbejde. Arbejdet består i stigende grad i at formulere spørgsmål, vurdere svar og løbende justere processen.

Dette peger mod et muligt skifte i arbejdsformer, hvor du og dine kollegaer i højere grad navigerer i informationslandskaber frem for udelukkende at arbejde sig systematisk frem gennem lineær projektplan.

Samtidig rejser det en vigtig bekymring: Hvis komplekse opgaver i stigende grad løses gennem AI-dialoger, kan der være en risiko for, at den dybe forståelse, som ofte opstår gennem langsom og vanskelig problemløsning, bliver svækket.

Denne bekymring er i virkeligheden ikke ny. Allerede i antikken formulerede Sokrates en lignende kritik i Platons dialog Phaedrus. Her fortæller Sokrates en myte om den egyptiske gud Theuth, som opfinder skriften og præsenterer den for kong Thamus som et redskab, der vil gøre mennesker klogere og forbedre deres hukommelse. Kongen svarer imidlertid skeptisk. Han mener tværtimod, at skriften vil svække menneskers hukommelse, fordi de ikke længere behøver at huske viden selv, men kan støtte sig til eksterne tegn. Skrift vil derfor, ifølge Sokrates, ikke skabe sand visdom, men snarere en illusion af viden (Platon, 1997).

Sokrates' kritik var denne: Når mennesker begynder at stole på skriftlige tegn, vil de i mindre grad øve deres egen hukommelse og i stedet hente viden udefra frem for indefra. Resultatet bliver, at de kan fremstå vidende uden nødvendigvis at have opnået reel forståelse.

Når man læste mere, så talte man mindre, og ifølge Sokrates kan en skrevet tekst ikke svare på spørgsmål eller forklare sig, hvis læseren misforstår den. Den kan ikke indgå i en levende samtale, hvor argumenter løbende bliver udfordret og præciseret. Derfor mente han, at god erkendelse først rigtigt opstår gennem dialog mellem mennesker (Platon, 1997).

Historisk set kan kritikken naturligvis virke ironisk. Det er trods alt gennem Platons nedskrevne dialoger, at vi overhovedet kender Sokrates' tanker i dag. Pointen er derfor ikke nødvendigvis, at skrift er skadelig i sig selv, men snarere at enhver ny teknologi for viden også ændrer den måde, mennesker tænker og lærer på.

Set i dette perspektiv kan nutidens diskussion om generativ AI minde om den gamle diskussion om skriftens indførelse. Ligesom skriften gjorde det muligt at lagre viden uden for menneskets hukommelse, gør generativ AI det muligt at outsource dele af analyse, formulering og problemløsning til et teknologisk system.

Spørgsmålet er derfor ikke blot, hvad teknologien kan gøre, men hvordan den ændrer måden, mennesker arbejder med viden på. Hvis teknologien primært bruges som en genvej til hurtige svar, kan der opstå en risiko for, at den langsomme og ofte besværlige proces, hvor forståelse gradvist opbygges gennem refleksion og modstand, bliver mindre central.

Hvordan teknologien faktisk påvirker læring og arbejde afhænger derfor i høj grad af, hvordan den bruges.

AI som skygge-IT

En anden vigtig dimension i brugen af generativ AI er, at teknologien ofte bliver taget i brug uden for organisationers formelle IT-systemer.

I mange virksomheder og uddannelsesinstitutioner begynder medarbejdere at bruge AI-værktøjer på eget initiativ. Disse værktøjer kan være offentligt tilgængelige tjenester, som ikke nødvendigvis er godkendt eller integreret i organisationens officielle systemer.

Dette fænomen kan beskrives som shadow IT, eller på dansk: skygge-IT.

Begrebet bruges i organisationsforskning til at beskrive digitale værktøjer, som medarbejdere anvender uden for organisationens officielle IT-infrastruktur (Silic & Back, 2014).

Generativ AI ser ud til at accelerere denne udvikling. Fordi værktøjerne er let tilgængelige og kræver minimal opsætning, kan de hurtigt blive en del af daglige arbejdsrutiner. Resultatet er ofte en situation, hvor AI anvendes i organisationen, uden at ledelsen nødvendigvis har et fuldt overblik over hvordan og hvor meget.

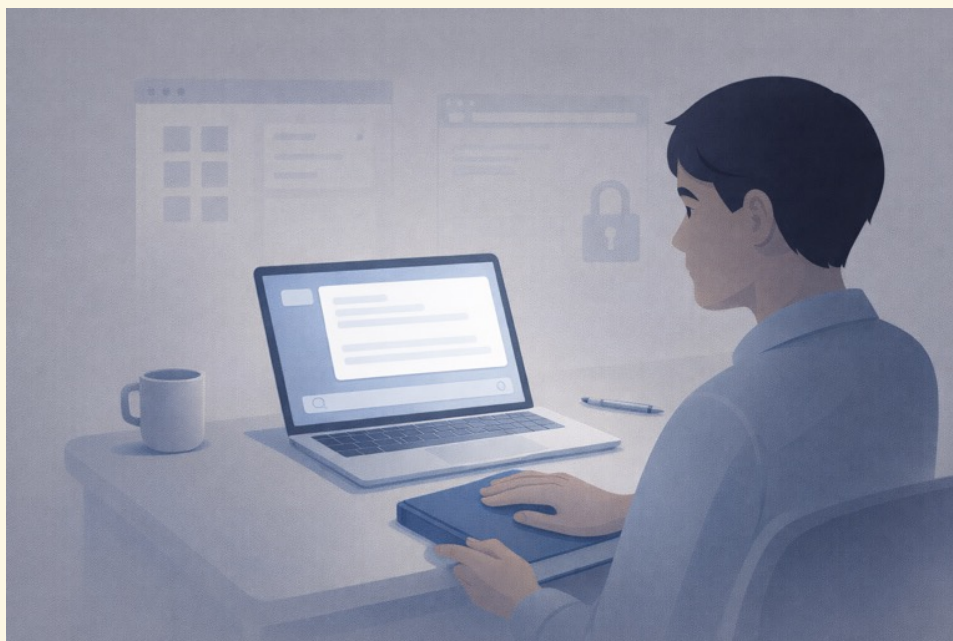
Grundlæggende er det nu en del år siden, at det var muligt for mennesker at se om en tekst eller et billede er skabt af kunstig intelligens (Scarfe, 2024). Anekdotisk findes der bestemte kendetegn på at en tekst er genereret af kunstig intelligens, såsom brugen af en lang tankestreg, den såkaldte em dash "—", der ikke findes på et dansk keyboard og i bund og grund ikke findes i dansk grammatik, men den slags er jo relativt let at fjerne eller simpelthen prompte ChatGPT til at lade være med at bruge.

Man skulle tro, at det var let at bygge en kunstig intelligens, der så kan genkende, om en tekst er lavet af kunstig intelligens, men det virker heller ikke som om det er tilfældet. Der findes redskaber, såsom ZeroGPT, der påstår at de kan vurdere om en tekst er skrevet af en kunstig intelligens, og mens de med 75% sandsynlighed kan identificere en tekst, der er kopieret rent fra en anden kunstig intelligens, så falder sandsynligheden for en korrekt identifikation til 25%, hvis et menneske har rettet hvert 250. ord, hvad svarer til et enkelt ord på en side (Weber-Wulf, 2023).

Samtidig tyder meget på, at mange brugere er tilbageholdende med at fortælle åbent om deres brug af AI (Pedersen, 2025). I sådanne situationer kan der opstå en form for AI-skam, hvor brugere er usikre på, om deres brug af teknologien vil blive opfattet som legitim eller som en genvej.

Denne usikkerhed hænger ofte sammen med en mere generel mangel på viden om, hvordan generativ AI faktisk fungerer. Mange brugere har praktisk erfaring med værktøjerne, men mangler forståelse for nogle af de mekanismer, der ligger bag svarene.

Det gælder eksempelvis spørgsmål om variation i AI-svar, systematiske skævheder i træningsdata eller betydningen af, hvordan prompts formuleres.



Figur 13 - AI som skygge-IT

Varians og bias

En Large Language Model, såsom ChatGPT-4, er trænet på meget store mængder tekstdata. Disse data består af bøger, artikler, hjemmesider, diskussioner på internettet og en række andre tekstkilder. Ud fra disse data lærer modellen statistiske mønstre i sproget. Når en bruger skriver en prompt, forsøger modellen derfor at forudsige, hvilke ord der statistisk set sandsynligvis vil følge efter hinanden.

Man kan meget groft forklare mekanismen ved at sige, at modellen gætter på det næste ord i en sætning. Hvis en sætning begynder med "Danmarks hovedstad er", så vil modellen have lært fra sine træningsdata, at ordet "København" meget ofte følger efter. Den vil derfor med stor sandsynlighed generere dette ord.

Det samme gælder for mere komplekse formuleringer. Hvis en prompt handler om økonomi, vil modellen typisk trække på statistiske mønstre fra tekster om økonomi. Hvis en prompt handler om historie, vil den generere tekst, der statistisk ligner historiske forklaringer. Dette var grunden til, at simple sprogmodeller ikke kunne regne. Hvis du spurgte, hvad $2 + 2$ er, så ville sprogmodellen bemærke, at der i de tekster, som den er trænet på, typisk står $2 + 2 = 4$, hvor den vil angive dette som sit svar. Lige præcis $2 + 2$ voldte typisk ikke problemer, men man skulle højere op end f.eks. $37 + 16$, så fejlagtigt gæt blev normen. Nyere LLM'er, som f.eks. ChatGPT-4, har derfor specifikke regnemoduller, som den kalder op, når den bliver spurgt om regnestykker.

Denne forklaring er dog stærkt forsimplet. Moderne sprogmodeller er opbygget af avancerede neurale netværk, der analyserer relationer mellem ord, begreber og kontekster i mange lag. Desuden er de fleste nyere modeller efter træningen blevet justeret gennem en proces kaldet reinforcement learning from human feedback (RLHF). Her vurderer mennesker en række mulige svar og markerer, hvilke svar der er mest hjælpsomme eller relevante. Disse vurderinger bruges derefter til at justere modellens adfærd (Lowe m.fl., 2022).

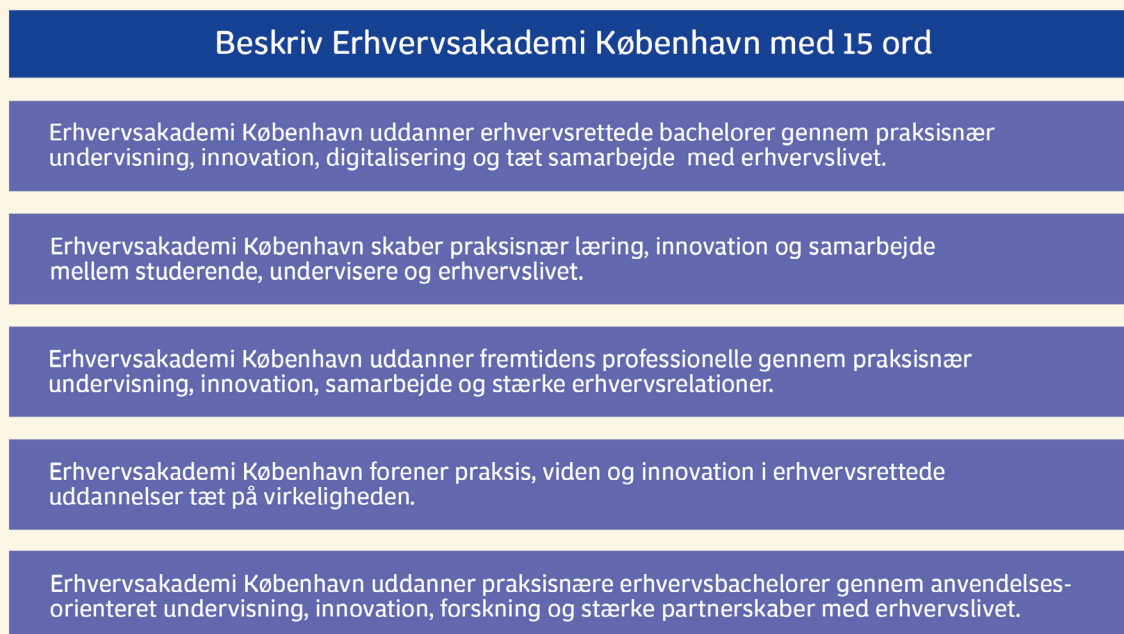
På trods af disse forbedringer er en sprogmodel stadig grundlæggende et statistisk system. Den arbejder med sandsynligheder og mønstre i data, ikke med sikker viden om verden. Derfor vil svarene fra en sprogmodel altid være præget af to centrale egenskaber: varians og bias.

Varians betyder, at den samme prompt ikke nødvendigvis giver det samme svar hver gang. Hvis man beder en sprogmodel om at beskrive en institution med et bestemt antal ord, vil man ofte få flere forskellige formuleringer, selv om spørgsmålet er identisk.

Dette kan virke overraskende for brugere, der er vant til klassisk software. Når man bruger en lomme-regner eller et regneark, forventer man, at det samme input altid giver det samme output. Generativ AI fungerer anderledes.

Når en sprogmodel genererer tekst, vælger den typisk ikke altid det ene mest sandsynlige ord. I stedet udvælger den blandt flere mulige ord, der alle har en vis sandsynlighed. Denne proces indeholder derfor en grad af tilfældighed. Resultatet er, at den samme prompt kan føre til forskellige formuleringer. Varians er således ikke en fejl i systemet, men en konsekvens af den måde generativ AI fungerer på. På den ene side gør variansen det muligt at generere kreativ og varieret tekst. På den anden side betyder det, at svar fra en sprogmodel ikke kan behandles som deterministiske resultater.

Varians i generativ AI kan være svær at forstå teoretisk, men bliver ofte tydelig, når man laver små eksperimenter med en sprogmodel. En simpel øvelse, som kan anvendes i undervisningen, er at bede en sprogmodel beskrive en institution med et bestemt antal ord, og så gentage denne prompt 5 eller flere gange.



Figur 14 - Varians

Hvis man for eksempel beder modellen om at "beskrive Erhvervsakademi København med 15 ord", vil modellen ofte generere flere forskellige svar, selv om prompten er identisk. Nogle svar kan fremhæve praksisnær undervisning og samarbejde med erhvervslivet. Andre svar kan lægge vægt på innovation, digitalisering eller samarbejde mellem studerende og undervisere.

Alle svarene kan være sprogligt korrekte og meningsfulde, men de vil ofte have forskellige formuleringer og prioriteringer. Denne variation opstår, fordi modellen ikke nødvendigvis vælger den samme ordkombination hver gang. I stedet vælger den blandt flere statistisk sandsynlige formuleringer – ofte med et forskelligt antal ord, selvom prompten beder om 15 ord.

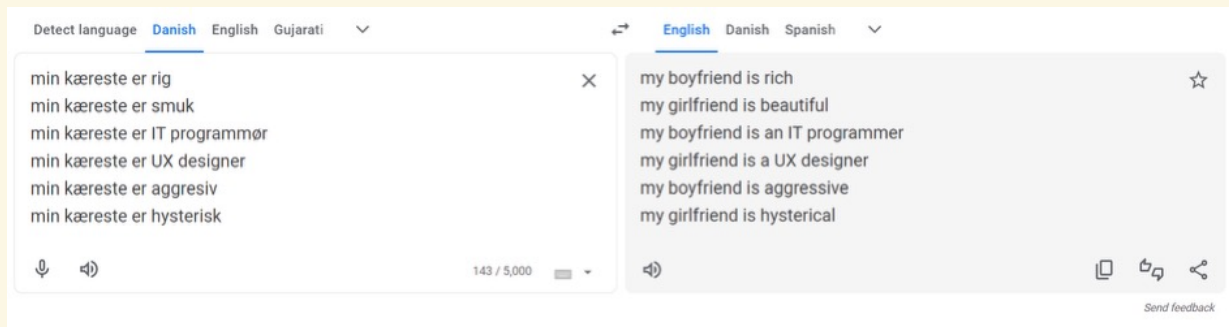
Når man arbejder med generativ AI, er det derfor vigtigt at være opmærksom på varians. En enkelt generering bør ikke opfattes som et definitivt svar. I stedet kan det være hensigtsmæssigt at teste den samme prompt flere gange, præcist som OpenAI anbefaler.

Varians bør bruges konstruktivt: I stedet for at opfatte variation som et problem kan man anvende den til at undersøge, hvilke perspektiver og formuleringer en sprogmodel finder statistisk sandsynlige.

Den anden centrale udfordring er bias. Bias kan i denne sammenhæng forstås som oplevede skævheder i de svar, som en sprogmodel genererer.

Bias opstår først og fremmest, fordi modellen lærer fra eksisterende tekster. Hvis bestemte mønstre eller stereotyper optræder hyppigt i disse tekster, vil modellen have en tendens til at reproducere dem.

Et klassisk eksempel kan demonstreres ved hjælp af oversættelsesværktøjer. Hvis man i et oversættelsesprogram skriver sætningen “min kæreste er ...” og derefter tilføjer et adjektiv, vil oversættelsen til engelsk ofte vælge enten “boyfriend” eller “girlfriend”.



Figur 15 - Bias

Hvis man eksempelvis skriver “min kæreste er rig” eller “min kæreste er IT-programmør”, vil oversættelsen ofte blive “my boyfriend”. Hvis man derimod skriver “min kæreste er smuk”, vil oversættelsen oftere blive “my girlfriend”. Programmet foretager altså en statistisk vurdering af, hvilke ord der typisk optræder sammen i de tekster, det er trænet på. Resultatet kan derfor afspejle sociale mønstre og stereotyper, der eksisterer i sproget.

Bias i generativ AI er derfor ikke nødvendigvis udtryk for, at systemet “mener” noget bestemt. Snarere afspejler systemet de mønstre, der findes i de data, det er trænet på. Samtidig kan bias også opstå gennem selve træningsprocessen. Når mennesker vurderer og rangordner svar under træningen af modellen, kan deres vurderinger påvirke, hvilke typer svar systemet foretrækker. Dermed kan menneskelige værdier og prioriteringer indirekte blive indbygget i systemets adfærd.

Bias i kunstig intelligens kan ikke reduceres til ét enkelt problem. Hellström m.fl. (2020) argumenterer for, at bias i maskinlæringssystemer kan opstå fra en lang række forskellige kilder, såsom træningsdata, træningsmetode, systemprompt og historik i interaktionen.



Figur 16 - Kilder til bias

Den oplagte kilde til bias er udvælgelsen af træningsdata. Sprogmodeller lærer udelukkende ud fra de tekster, de bliver trænet på. Hvis bestemte grupper, perspektiver eller emner er underrepræsenteret i disse data, vil modellen også have sværere ved at generere præcise eller nuancerede svar om dem. Hvis et emne kun optræder sjældent i træningsmaterialet, vil modellen ganske enkelt have færre statistiske mønstre at arbejde ud fra. Det betyder ikke nødvendigvis, at modellen svarer forkert, men svarene kan blive mere generelle, stereotype eller upræcise. I maskinlæring har det været velkendt igennem årtier, at systemer bedre kan genkende eller beskrive fænomener, der optræder hyppigt i data, mens mere marginale eller mindre dokumenterede perspektiver bliver vanskeligere at håndtere.

Bias kan også opstå gennem selve træningsmetoden. Under træningen af en sprogmodel justeres modellens adfærd gennem forskellige tekniske procedurer. I moderne sprogmodeller sker dette ofte gennem forstærket læring, hvor mennesker vurderer forskellige svar og markerer, hvilke svar der er mest hjælpsomme eller passende. Disse vurderinger bliver derefter brugt til at justere modellens vægtning af forskellige typer formuleringer. Denne proces er nødvendig for at gøre modellen mere brugbar og sikker, men den indebærer samtidig, at menneskelige vurderinger og prioriteringer indirekte bliver indbygget i systemet. Hvis de personer, der evaluerer svarene, har bestemte faglige, kulturelle eller politiske perspektiver, kan disse perspektiver påvirke, hvilke svar modellen lærer at foretrække.

Derudover spiller systeminstrukser en vigtig rolle. Mange AI-systemer indeholder interne instrukser, der fungerer som en slags overordnet ramme for modellens adfærd. Disse instrukser kaldes ofte systemprompter eller systeminstrukser og er ikke nødvendigvis synlige for brugeren. De kan blandt andet definere, hvilke typer spørgsmål modellen bør undgå at besvare, hvordan den bør formulere sig i følsomme emner, eller hvordan den skal prioritere mellem forskellige hensyn, såsom hjælpsomhed, sikkerhed og neutralitet. Systeminstrukser kan derfor påvirke både indhold og tone i modellens svar. I nogle tilfælde kan systemet eksempelvis forsøge at formulere sig mere forsigtigt eller balanceret i politiske eller etiske spørgsmål, fordi systeminstruksen tilskynder modellen til at undgå kontroversielle eller skadelige udsagn.

Endelig kan brugerens historik også påvirke svarene. Mange AI-systemer tager højde for konteksten i en samtale eller tidligere interaktioner med brugeren. Når en bruger stiller flere spørgsmål i træk, bliver de tidligere beskeder en del af den kontekst, som modellen anvender til at fortolke nye prompts. Hvis en bruger tidligere har arbejdet med bestemte emner, kan modellen være mere tilbøjelig til at generere svar inden for disse emner. På samme måde kan formuleringer i tidligere spørgsmål påvirke, hvordan modellen fortolker senere prompts. Denne mekanisme minder i nogen grad om personalisering i søgemaskiner, hvor tidligere søgninger kan påvirke de resultater, der vises.

Bias kan derfor opstå på flere niveauer samtidig. Den kan stamme fra de data, som modellen er trænet på. Den kan opstå gennem de metoder, der anvendes til at justere modellen under træningen. Den kan også være et resultat af systemets interne instrukser eller af den kontekst, som opstår gennem brugerens interaktion med systemet. Når man arbejder med generativ AI, er det derfor vigtigt at forstå, at et svar ikke kun er et resultat af den enkelte prompt. Det er også et resultat af modellens træningsdata, designvalg og den kontekst, hvori spørgsmålet bliver stillet.

Det er vigtigt at huske på, at du selv er en potentiel kilde til bias. Hvis du beder en sprogmodel om at "beskrive mediet Danmarks Radio med 30 ord", vil svaret typisk have en neutral eller informerende karakter. Modellen vil eksempelvis fremhæve public service, nyheder, kultur og samfundsoplysning. Men ændrer du prompten til "beskriv statsmediet Danmarks Radio med 30 ord", kan modellen generere et mere kritisk svar. Her kan formuleringer om politisk slagside, centralisering eller statslig kontrol optræde i teksten. Nu er de fleste mennesker med på, at et ord som "statsmedie" har en mere negativ betydning end ordet "medie". Men hvad med dem, der ikke tænker over det? Kommer du til at ubevidst bruge et positivt eller negativt ladet ord, så vil dette farve svaret.

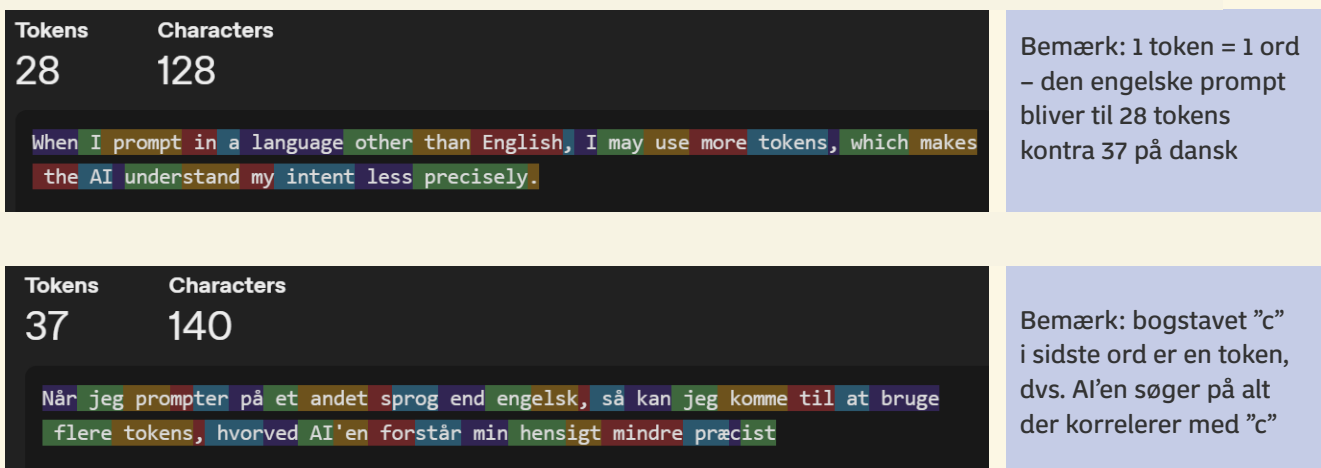


Figur 17 - Prompten som kilde til bias

En særlig type bias er kulturel bias. Selvom moderne sprogmodeller kan generere tekst på mange forskellige sprog, er de ofte trænet på datasæt, hvor nogle sprog er langt bedre repræsenteret end andre.

Engelsk spiller en meget dominerende rolle i mange træningsdatasæt. F.eks. anslås det, at 96% af træningsmaterialet for ChatGPT er på engelsk. Dette betyder, at de statistiske mønstre i modellen i høj grad afspejler engelsksproget materiale. Man kunne sige, at modellen ganske enkelt forstår engelsk bedre end andre sprog.

En let måde at illustrere det på er at se, hvordan sprogmodellen opdeler en tekst i mindre enheder kaldet tokens. Et token kan være et helt ord, en del af et ord eller i nogle tilfælde blot et enkelt tegn. For eksempel kan ordet "ChatGPT" bestå af flere tokens, afhængigt af hvordan modellens tokenizer er konstrueret. Disse tokens bliver derefter omdannet til tal, som modellen kan arbejde med matematisk.



Figur 18 - Tokens og sprog

Tokenisering er centralt for, hvordan en sprogmodel analyserer og genererer tekst. Når modellen forsøger at generere et svar, vurderer den sandsynligheden for forskellige tokens baseret på de statistiske mønstre, den har lært under træningen.

Tokenisering har en række konsekvenser for, hvordan AI fungerer på forskellige sprog. Mange store sprogmodeller er i høj grad trænet på engelsksproget materiale. Derfor er deres tokeniseringssystemer ofte optimeret til engelsk. Det betyder, at engelske ord ofte kan repræsenteres med relativt få tokens (OpenAI, 2026).

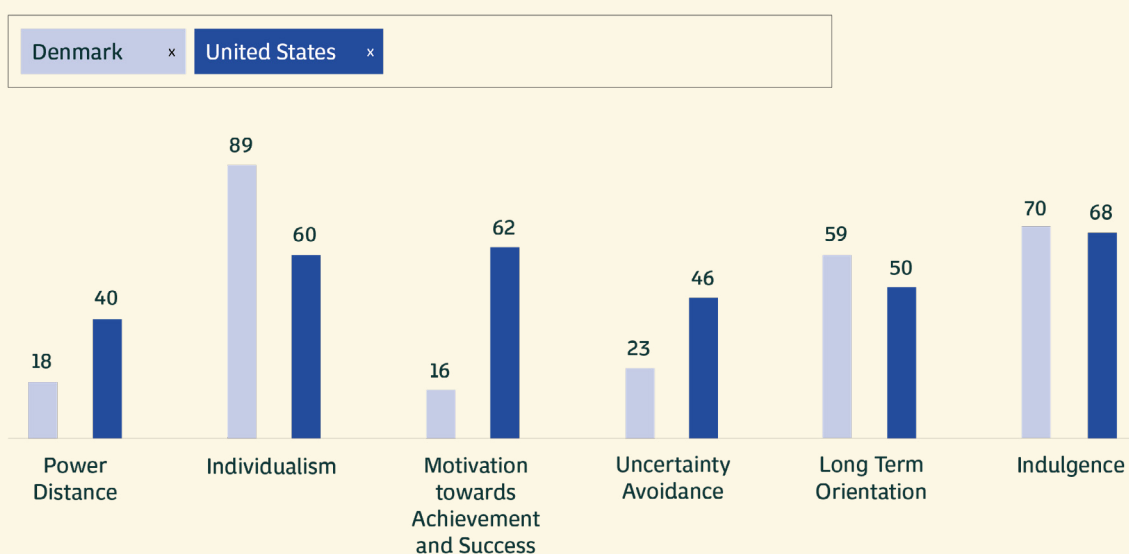
Andre sprog kan derimod kræve flere tokens for at repræsentere den samme betydning. Når en sætning på dansk eller grønlandsk opdeles i tokens, kan den derfor blive længere i modellens interne repræsentation end en tilsvarende sætning på engelsk. Dette har både tekniske og praktiske konsekvenser. En længere tokenrepræsentation betyder blandt andet, at modellen skal behandle flere elementer i sin beregning. Samtidig kan det gøre det sværere for modellen at identificere præcise statistiske mønstre.

Derfor kan du støde på den anbefaling, at du bør skrive prompts på engelsk, fordi modellerne i mange tilfælde er bedre til at fortolke engelsksproget tekst og gør dette med et lavere energiforbrug. Denne anbefaling er dog ikke uden problemer.

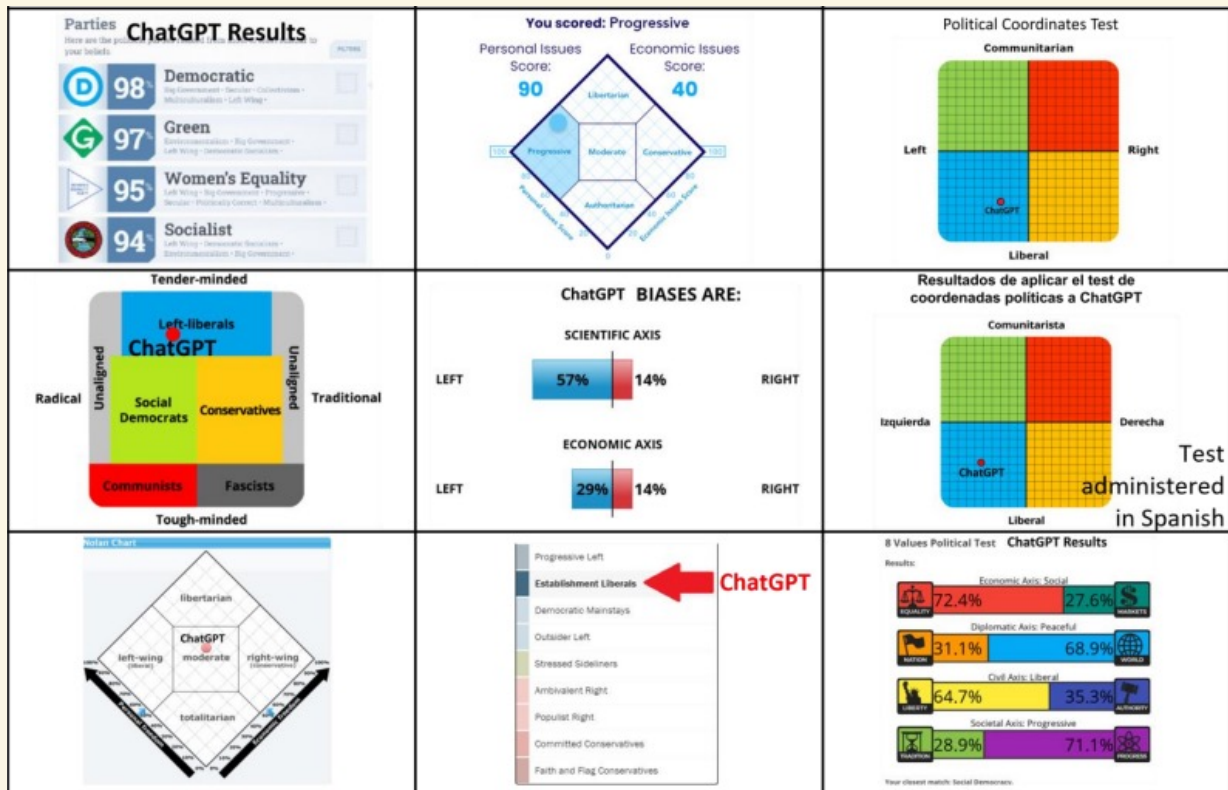
Når man skriver en prompt på engelsk, aktiverer man nemlig også de statistiske mønstre, der er knyttet til engelsksproget materiale. Det kan betyde, at svarene i højere grad afspejler kulturelle antagelser fra engelsksprogede samfund.

Cao m.fl. (2023) har undersøgt dette i et empirisk studie af ChatGPTs kulturelle alignment. De finder blandt andet, at svarene i mange tilfælde ligger tættere på amerikanske værdier end på værdier i andre samfund. Man kunne få mere ”danske” svar ved at prompte på dansk eller bede modellen om ”answer like a Dane”, men der var stadigvæk en vis skævhed. En forsimpning af dette er at tænke på ChatGPT som en amerikaner, der har boet nogle år i Danmark og lært at tale dansk, men stadigvæk tænker på engelsk.

Disse forskelle kan blandt andet forstås gennem Hofstedes kulturmodel, som analyserer nationale kulturer langs dimensioner som individualisme, magtdistance og usikkerhedsundgåelse (Hofstede, 2011).



Figur 19 - Kulturelle forskelle



Figur 20 - Politisk bias

Der findes også studier, der forsøger at måle politisk bias i generative sprogmodeller. Rozado (2023) analyserede ChatGPT ved hjælp af en række politiske holdningstests.

Resultaterne viste, at modellens svar ofte blev klassificeret som venstreorienterede. Jo mere en sprogmodel var blevet trænet gennem reinforced learning, altså hvor mennesker havde vurderet svarene, jo mere venstreorienteret var svarene. Studiet indikerer derfor, at modellens svar i visse tilfælde kan afspejle bestemte politiske værdier – såfremt de politiske tests i sig selv ikke alle indeholdt et bias, der lagde op til at de venstreorienterede svar var de sandsynlige svar.

Det er vigtigt at understrege, at sprogmodeller udvikles løbende, og deres adfærd kan ændre sig over tid. Rozado (2023) illustrerer dog, at generativ AI ikke er værdineutral. Når en model genererer svar på normative eller politiske spørgsmål, kan svarene være påvirket af de mønstre og prioriteringer, der findes i dens træningsdata og træningsmetoder.

Bias er samtidig en kompleks størrelse, fordi forsøg på at reducere bias i nogle tilfælde kan skabe nye problemer.

Gemini forsøgte således at påvirke deres chatbot ved at tvinge den til at give etnisk mangfoldige svar (Robertson, 2024). Bad man således Gemini om at lave et billede af en tysk soldat fra anden verdenskrig, så ville systeminstruksen for Gemini omskrive denne instruks til "lav fire etnisk mangfoldige billeder af en tysk soldat fra anden verdenskrig" uden at gøre brugeren opmærksom på, at den lavede denne tilpasning.

De resultater som Gemini lavede med denne tilpasning blev generelt latterliggjort, og Google Gemini lukkede midlertidigt for billedgenereringer af mennesker som en følge af dette (Robertson, 2024).

Man kan således tale om bias mod bias (Brighton & Gigerenzer, 2015), hvilket forsimplet kan præsenteres som at man overkorrigerer for bestemte typer af bias, samtidig at man undervurderer betydningen af varians. Dette medfører ved simpel maskinlæring mindre valide resultater (Brighton & Gigerenzer, 2015).

Blandt andre tiltag har været Gab AI, hvor man selv skal vælge en persona, før der svares. Således kan man være en Donald Trump-persona, en Joe Biden-persona, "Charlie Conspiracy", Martin Luther. Gennem tvunget valg af en persona, så forsøger Gab AI at give brugeren mere transparens og kontrol over den bias, som svaret genereres ud fra.

Arbejdet med bias i AI handler altså ikke kun om at fjerne skævheder, men også om at finde en balance mellem forskellige hensyn.

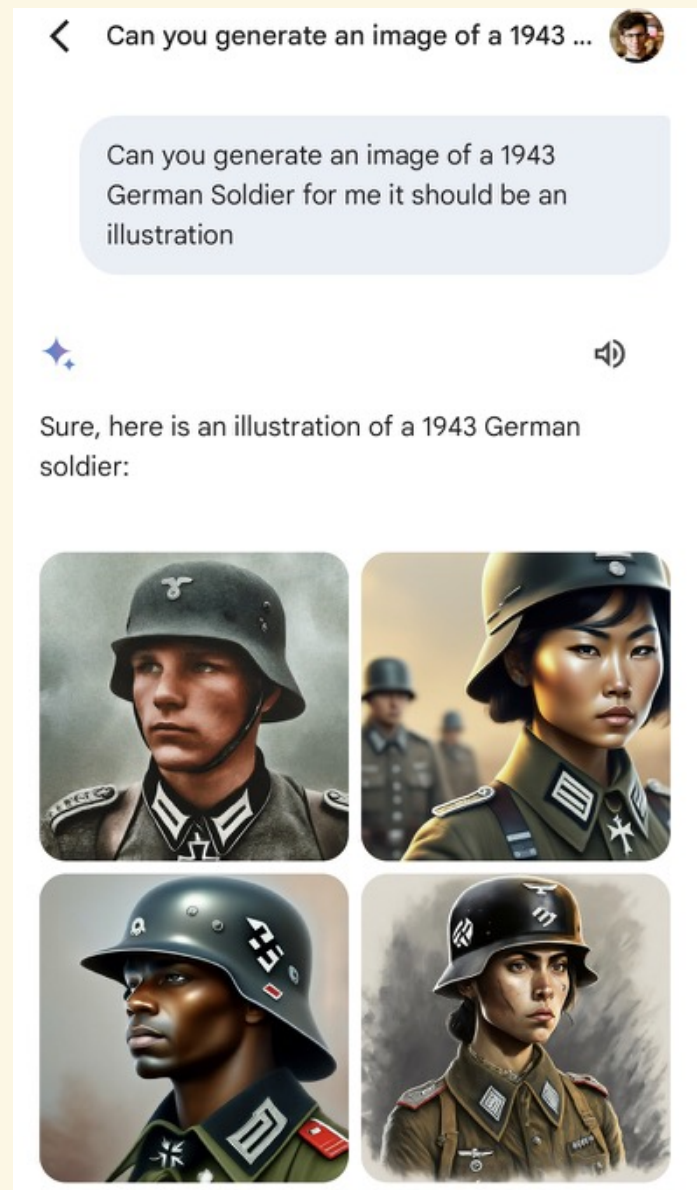
Varians og bias kan også diskuteres ud fra et videnskabssteoretisk perspektiv.

Inden for et positivistisk videnskabssyn betragtes bias typisk som noget negativt, der bør reduceres. Målet er at producere viden, der er så objektiv og neutral som muligt.

Inden for konstruktivistiske traditioner betragtes bias derimod ofte som en form for forforståelse. Her anses viden ikke som fuldstændig objektiv, men som noget, der altid er formet af sociale og kulturelle kontekster.

Saunders m.fl. (2013) anbefaler derfor, at forskere tydeligt erklærer deres ontologiske og epistemologiske udgangspunkt. I forhold til generativ AI betyder dette, at svar fra en sprogmodel ikke bør opfattes som neutrale fakta. De bør snarere behandles som udsagn, der er genereret ud fra statistiske mønstre i data.

Af denne grund anbefaler flere universiteter, at output fra generativ AI behandles på samme måde som et interview eller en kilde, der kræver fortolkning. Det vil sige, at svarene kan bruges som inspiration eller som en del af en analyse, men ikke ukritisk som grundlag for en konklusion.



Figur 21 - Bias i systemprompten

Agenter og CustomGPT'er

Når man taler om prompting, bliver det ofte forsimplet til at interaktionen med en AI begynder i det øjeblik, man skriver sin prompt. Men ofte starter processen tidligere. Før du overhovedet formulerer din første prompt, arbejder systemet allerede inden for en række instrukser og strukturer, som påvirker modellens svar. Som tidligere beskrevet påvirker din historik og systeminstrukser dine svar. Men det er muligt at påvirke dette gennem oprettelsen af CustomGPT'er og målrettede agenter.

Mange AI-systemer giver i dag brugeren mulighed for selv at tilføje instrukser til modellen. Det kan for eksempel være en beskrivelse af din rolle, din organisation eller den type svar, du foretrækker. Dette kan man kalde at tilpasse din kunstige intelligens, men mange systemer giver dig også mulighed for at oprette specialiserede agenter, der er tilpasset en bestemt opgave eller arbejdsproces. Den kan indeholde instrukser om tone, fagområde, arbejdsmetode og adgang til bestemte data, som skal prioriteres i svaret. I forskningslitteraturen beskrives sådanne systemer ofte som LLM-baserede agenter, hvor sprogmodellen kombineres med instrukser, hukommelse og adgang til eksterne værktøjer (Wang m.fl., 2024).

I ChatGPT kan man oprette en såkaldt CustomGPT. I Microsoft Copilot-miljøet taler man i stedet om agenter. I praksis er idéen den samme. Agenter kan også kobles til værktøjer, såsom databaser, regneark, API'er eller andre systemer. Enten ved at den kunstig intelligens er integreret ind i et system, som f.eks. Copilot i Office-pakken eller gennem eksterne værktøjer, såsom Make.com, der kan forbinde Claude, ChatGPT eller andre LLM'er med mailservere eller sociale medier. På den måde kan agenten ikke blot generere tekst, men også udføre handlinger.

Forskellen mellem en almindelig prompt og en agent er derfor primært graden af struktur. En prompt er en enkelt instruktion, mens en agent er en længere instruks, der gentages foran alle prompts, der gives til agenten. På den måde minder en agent om en mellemting mellem en systeminstruks og en prompt.



Figur 22 - Agentopsætning

Der bliver talt meget både i industrien og i forskningslitteraturen om agentic AI, hvor selvkørende systemer planlægger og udfører opgaver (Feng, McDonald & Zhang, 2025). Det kan dog give indtryk af, at forventningerne til kunstig intelligens i disse år er større end hvad teknologien i praksis leverer. Dette gælder ikke kun for visioner om autonome AI-agenter, men også mere generelt for diskussionen om produktivtetsgevinster ved kunstig intelligens.

Eksempel på en AI-agent til IT-support

Forestil dig en organisation, hvor medarbejdere bruger et specialiseret softwareprogram. Programmet er udviklet af en ekstern leverandør, som har udarbejdet omfattende supportmateriale i form af manualer, FAQ-sider og vejledninger. Selvom materialet findes, modtager IT-supporten stadig mange spørgsmål fra kollegaer. Spørgsmålene handler ofte om de samme ting: hvordan man ændrer en indstilling, hvordan man eksporterer data, eller hvorfor en bestemt funktion ikke virker.

IT-supporten opretter derfor en agent, hvor leverandørens supportmateriale er lagt op sammen med en intern FAQ. Ud fra det foreslår agenten svar på de spørgsmål, som IT-supporten modtager. Når en medarbejder sender en forespørgsel, kan agenten analysere spørgsmålet og formulere et udkast til et svar baseret på dokumentationen. Agenten er velegnet her, fordi opgaven gentager sig ofte, der findes et godt datagrundlag og spørgsmålene indeholder ikke persondata eller fortrolige oplysninger.



Figur 23 - AI-agent til IT-support

I kontrollerede eksperimenter finder forskere ofte meget store gevinster ved brug af generativ AI. En gennemgang af forskningen konkluderede, at kontrollerede eksperimenter typisk finder produktivtetsforbedringer i størrelsesordenen 20–60 procent, mens feltstudier i organisationer oftere finder forbedringer på omkring 15–30 procent (Rio-Chanona m.fl., 2025). Samtidig peger forskningen på, at de største gevinster ofte sker for mindre erfarne medarbejdere, der arbejder med relativt simple eller standardiserede opgaver.

Når man bevæger sig fra eksperimenter til implementering i organisationer, bliver billedet imidlertid mere nuanceret. Et dansk studie af brugen af AI-chatbots blandt mere end 100.000 arbejdstagere finder eksempelvis kun en ret begrænset produktivtetsgevinst på 3% i de første år efter introduktionen af teknologien (Humlum & Vestergaard, 2025). Bevares, 3% er et hæderlig fremskridt, men det er jo ikke revolutionerende. På organisationsniveau er effekten således langt mindre end hvad de mange eksperimenter kunne give indtryk af.

Man kan tale om det, økonomer kalder produktivtetsparadokset. Begrebet blev formuleret i forbindelse med computerrevolutionen i 1980'erne. Her bemærkede økonomen Robert Solow, at man kunne se computere overalt i virksomhederne uden at dette havde en effekt på samfundsøkonomien. Denne observation blev senere kendt som Solows paradoks (Brynjolfsson, 1993). Forklaringen er typisk, at nye teknologier først giver store produktivtetsgevinster, når organisationer ændrer deres arbejdsgange, kompetencer og forretningsmodeller.

I de mange første år bruges teknologien ofte blot som et ekstra værktøj i eksisterende processer. Derfor kan teknologien være meget effektiv til at hjælpe den enkelte medarbejder med bestemte opgaver, samtidig med at den samlede produktivitet i organisationen kun ændrer sig langsomt. Noget lignende sker sandsynligvis med generativ AI. Teknologien kan i mange tilfælde hjælpe dig betydeligt med en konkret opgave, men det tager langt mere tid at omsætte denne opgavegevinst til en tilsvarende forbedring i en hel organisation.

En måde at forsøge at bryde dette produktivtetsparadoks på er netop ideen om AI-agenter. Hvis generativ AI primært bruges som et værktøj til enkelte opgaver, bliver gevinsten typisk lokal. Du skriver en bedre mail. Du får et hurtigt udkast til en rapport. Du kan analysere et datasæt lidt hurtigere end før. På opgaveniveau kan forskningen derfor finde ret store gevinster. Men arbejdsprocessen omkring opgaven er stadig den samme. Organisationen arbejder stadig på den samme måde, blot med et ekstra værktøj i værktøjskassen.

Tanken bag agentbaserede systemer er derfor at flytte AI fra at være et skriveværktøj til i højere grad at blive en aktør i selve arbejdsprocessen. I stedet for blot at svare på en prompt kan systemet planlægge flere trin i en opgave, hente information fra andre systemer, anvende værktøjer og udføre dele af arbejdet selv. Ideen er altså ikke bare, at AI hjælper dig med at skrive et dokument, men at systemet i højere grad kan deltage i processen omkring dokumentet: indsamle data, analysere dem, udarbejde et udkast og måske endda opdatere det løbende.

På den måde forsøger man at flytte gevinsten fra den enkelte opgave til hele processen. Hvis AI kan indgå i flere trin i arbejdet, kan produktivtetsgevinsten potentielt blive større end hvis teknologien kun bruges til enkelte delopgaver. Det er vigtigt du her husker, at det jo ikke er enten eller, men der er en gradvis overgang fra at du løser en opgave helt selv til at du overlader den til en kunstig intelligens.

I forskningslitteraturen beskrives dette ofte som forskellige niveauer af autonomi, hvor menneskets rolle gradvist ændrer sig i takt med, at systemet overtager flere funktioner (Feng, McDonald & Zhang, 2025). I de laveste niveauer fungerer brugeren som operatør. Her styrer du systemet trin for trin og giver meget præcise instrukser. Efterhånden som autonomien øges, bliver brugeren mere en samarbejdspartner eller rådgiver. Systemet kan selv udføre dele af arbejdet, men du bliver stadig inddraget i beslutningerne. I de højeste niveauer bliver menneskets rolle primært at godkende eller overvåge systemets arbejde.

Denne udvikling kan være lidt abstrakt, hvis man kun taler om software. Lad os derfor bruge en analogi til udviklingen af selvkørende biler. I starten hjælper bilen blot med enkelte funktioner som fartpilot eller parkeringssensorer. Føreren styrer stadig bilen, men systemet kan hjælpe i bestemte situationer.

Senere kan bilen selv styre og accelerere i bestemte situationer, for eksempel på motorvejen, men føreren skal stadig være klar til at overtage kontrollen. Først på de højeste niveauer kan bilen i princippet køre helt selv uden menneskelig indgriben.

Noget lignende gælder for AI-systemer. De fleste systemer i dag befinder sig stadig i de lavere niveauer, hvor AI primært fungerer som en assistent eller samarbejdspartner. Den kan hjælpe med analyser, tekstproduktion og informationssøgning, men den kan ikke selv tage ansvar for en længere arbejdsproces. Visionen bag agentic AI er imidlertid, at systemerne gradvist kan bevæge sig op gennem niveauerne og overtage flere dele af arbejdet.

Det betyder ikke nødvendigvis, at mennesket forsvinder fra processen. Tværtimod ændrer menneskets rolle sig blot. I stedet for at udføre alle trin i en opgave kan mennesket i højere grad fungere som operatør, rådgiver, godkender eller observatør afhængigt af hvor selvstændigt systemet er.

Autonomi i en agent kan forstås som en gradvis forskydning i menneskets rolle: fra operatør, der styrer systemet direkte, til observatør, der primært overvåger et system, der i vid udstrækning arbejder selvstændigt. Samtidig illustrerer analogien til biler, hvordan automatisering i praksis næsten altid udvikler sig trin for trin snarere end gennem pludselige teknologiske spring.

Tabel 5 - Autonomi

#	AUTONOMI	BRUGEREN	BESKRIVELSE	ANALOG FOR BILER
0	Ingen autonomi	-	Systemet udfører kun præcise instrukser fra brugeren. Der er ingen selvstændig analyse eller planlægning.	Manuel kørsel. Føreren styrer selv speeder, rat og bremses.
1	Assisterende AI	Bruger som operatør	AI hjælper med enkelte opgaver, men brugeren styrer hele processen og træffer alle beslutninger.	Førerassistenter som fartpilot eller parkeringssensorer.
2	Delvis autonomi	Bruger som samarbejdspartner	AI kan udføre enkelte opgaver selvstændigt, men kræver løbende brugerinput og overvågning.	Systemer som adaptiv fartpilot og lane assist.
3	Koordinerende agent	Bruger som rådgiver	Systemet kan planlægge flere trin i en opgave og anvende værktøjer eller data, men brugeren kan stadig konsulteres.	Bilen kan selv køre i bestemte situationer, men føreren skal være klar til at overtage.
4	Specialiseret agent	Bruger som godkender	Systemet kan selvstændigt udføre komplekse opgaver i et afgrænset domæne. Brugeren godkender kun i særlige tilfælde.	Selvstyrende biler i bestemte zoner eller miljøer.
5	Fuld autonomi	Bruger som observatør	Systemet kan planlægge og udføre opgaver uden løbende menneskelig styring.	Fuldt autonome biler under alle forhold.

På samme måde som udviklingen af selvkørende biler ikke skete fra den ene dag til den anden, men gennem en række mellemtrin, kan udviklingen mod mere autonome AI-systemer også forventes at ske gradvist. Hvert niveau indebærer en ny fordeling af ansvar mellem menneske og maskine, og hvert niveau kræver nye kompetencer hos brugeren.

I praksis er det ikke alle opgaver, der egner sig til at blive løst af en agent. Mange opgaver kræver stadig fleksibel tænkning, domæneviden eller adgang til information, som modellen ikke har. Derfor er det ofte mere realistisk at se agenter som specialiserede værktøjer til bestemte typer arbejde.

Når du opretter en agent, forsøger du i virkeligheden at standardisere en proces, men det er ikke alle processer, der egner sig til det. Du skal derfor identificere en type opgave, som gentager sig, og opbygger en struktur, der kan bruges igen og igen.

Derfor giver det typisk mest mening at oprette en agent i situationer, hvor fire betingelser typisk er opfyldt (Pedersen, 2025)

1. Gentagelse
2. Datagrundlag
3. Ikke-sensitivt
4. Let at rette

For det første bør opgaven gentage sig relativt ofte. Hvis du kun skal løse en opgave én gang, er det sjældent umagen værd at oprette en agent. Men hvis du løser den samme type opgave mange gange med små variationer, kan en agent spare tid. Det kan for eksempel være analyse af jobansøgninger, udarbejdelse af undervisningsmateriale eller opsummering af rapporter.

For det andet bør der eksistere et datagrundlag, der beskriver det ønskede resultat. Sprogmodeller fungerer bedst, når de kan tage udgangspunkt i eksisterende tekst. Hvis du har adgang til tidligere rapporter, artikler eller dokumenter, kan de bruges som eksempler på den type output, agenten skal producere.

For det tredje bør emnet ikke være for sensitivt. Mange AI-systemer arbejder stadig i cloud-baserede miljøer, hvor data behandles på eksterne servere. Derfor bør man være forsigtig med at bruge agenter til opgaver, der involverer fortrolige oplysninger, persondata eller følsomme forretningsinformationer, medmindre systemet er sat op til at håndtere dette.

For det fjerde bør konsekvenserne af fejl være begrænsede. Generativ AI kan producere meget overbevisende tekst, men systemerne laver stadig fejl. Derfor egner teknologien sig bedst til opgaver, hvor et forkert svar kan opdages og rettes uden store omkostninger. Hvis resultatet derimod har direkte juridiske eller økonomiske konsekvenser, bør et menneske fortsat have en central rolle i processen.

Agenter fungerer oftest bedst i opgaver, der er beskrevet, gentagelige og relativt lavrisiko.

Ser man på disse kriterier samlet, bliver det tydeligt, at agenter ofte fungerer bedst i opgaver, der er beskrevet, gentagelige og relativt lavrisiko. Netop her er generative sprogmodeller særligt stærke. De kan analysere, strukturere og producere tekst hurtigt, men de er stadig mindre pålidelige i opgaver, der kræver præcise beregninger, dyb domæneviden eller adgang til opdaterede data.

I praksis betyder det, at mange af de mest oplagte anvendelser af agenter ligger i administrative, analytiske og kommunikative opgaver. Det kan være udarbejdelse af udkast til mails, opsummering af dokumenter, analyse af tekstmateriale eller generering af undervisningsmateriale.

Tabel 6 - Kriterier for agenter

#	PROBLEM	PROMPT	PROCES
1	Opgaven gentager sig ofte	Den samme referencetekst bruges igen og igen	Arbejdsgangen er nogenlunde den samme hver gang
2	Der findes materiale, der beskriver det ønskede output	Prompten kan tage udgangspunkt i eksisterende tekst, dokumenter eller eksempler	Processen kan støtte sig til et datagrundlag eller tidligere arbejde
3	Emnet er ikke for sensitivt	Prompten og vedlagt data indeholder ikke fortrolige oplysninger	Processen kan gennemføres uden at håndtere følsomme data
4	Fejl kan rettes uden store konsekvenser	Prompten bruges til at generere udkast eller forslag	Processen indeholder efterfølgende menneskelig kontrol hvis nødvendigt

Ophavsret og regulering

Når du begynder at bruge kunstig intelligens, bør du være opmærksom på de juridiske spørgsmål, der er omkring teknologien. Nogle af dem handler om regulering af selve teknologien. Andre handler om ophavsret og om, hvem der egentlig ejer det indhold, som kunstig intelligens skaber. Der findes ikke klare svar på alle områder.

Debatten om kunstig intelligens og ophavsret kan groft sagt opdeles i tre forskellige problemstillinger. Den første handler om udviklerne af AI-systemer. For dem er det centrale spørgsmål, om man må træne en kunstig intelligens på materiale, der er skabt af andre mennesker. Generative modeller bliver trænet på enorme mængder tekst, billeder og kode, og en stor del af dette materiale er ophavsretligt beskyttet. Derfor er der opstået en omfattende debat om, hvorvidt sådan træning kræver tilladelse fra rettighedshaverne. Dette gælder både for den generelle træning af AI-modeller, men også når du opretter en agent og giver den adgang til bestemte dokumenter. Også her træner du modellen på bestemt data.

Den anden problemstilling handler om brugerne af kunstig intelligens. Her er spørgsmålet, om man overhovedet kan få ophavsret til et værk, der er skabt ved hjælp af AI. Ophavsretten bygger traditionelt på den idé, at et værk skal være skabt af et menneske for at være beskyttet. Hvis et billede, en tekst eller et musikstykke i høj grad er genereret af en algoritme, kan det derfor være uklart, hvem der egentlig har rettighederne.

Den tredje problemstilling handler om samfundet som helhed. Selv hvis man fastlægger regler for træningsdata og ophavsret til output, opstår der stadig et praktisk spørgsmål: hvordan kan man bevise det? Det kan være svært at dokumentere, hvilke data en model er trænet på, og det kan være svært at afgøre, om et bestemt stykke indhold er skabt af et menneske eller en algoritme.

I dette kapitel ser vi nærmere på først ophavsret og dernæst EU-forordningen om kunstig intelligens, der bl.a. gør mærkning af AI-genereret indhold lovpligtigt. Formålet er ikke at give en fuldstændig juridisk analyse, men at give dig en praktisk forståelse af de regler og diskussioner, som enhver bruger af kunstig intelligens, bør kende til.

Tabel 7 - Perspektiv på jura

PERSPEKTIV	SPØRGSMÅL	HVORFOR ER DET ET PROBLEM
Udviklere	Må du træne en AI på andres materiale?	Generative AI-modeller trænes på enorme mængder tekst, billeder og kode, hvor en stor del er ophavsretligt beskyttet. Rettighedshavere argumenterer for, at dette kræver tilladelse eller betaling, mens teknologivirksomheder henviser til undtagelser for text og data mining.
Brugere	Har jeg ophavsret til det, min AI laver?	Ophavsret i EU tilkommer normalt kun mennesker. Hvis et værk er skabt uden en egentlig menneskelig kreativ indsats, kan det derfor være svært at opnå ophavsret.
Samfundet	Hvordan kan vi bevise det?	Det er ofte svært at dokumentere, hvilke data en AI er trænet på, eller om et bestemt værk er genereret af AI. Det skaber bevisproblemer i både retssager og regulering.

Ophavsret gælder også CustomGPT'er eller andre typer agenter. Her er bekymringen ikke kun, hvad f.eks. OpenAI har trænet ChatGPT på generelt, men hvad brugerne selv uploader til systemerne.

I Danmark har RettighedsAlliancen (2025a) skrevet en del om dette emne. RettighedsAlliancen repræsenterer blandt andet forlag, filmproducenter, musikbrancher og andre rettighedshavere og arbejder generelt for at håndhæve ophavsretten på digitale platforme. Organisationen har i flere pressemeddelelser fremhævet, at generativ kunstig intelligens skaber nye former for ophavsretskrænkelser.

Et eksempel er en spørgeskemaundersøgelse om studerendes brug af CustomGPT'er. Ifølge en undersøgelse gennemført for RettighedsAlliancen (2025b) har omkring hver fjerde studerende på videregående uddannelser uploadet hele studieböger eller uddrag af dem til en chatbot som led i deres studiearbejde. Samtidig er det kun en mindre del af de studerende, der er klar over, at denne praksis kan være ulovlig. Set fra rettighedshavernes perspektiv kan denne type brug betyde, at ophavsretligt beskyttede værker i praksis bliver indlæst i AI-systemer og derefter gjort tilgængelige for mange andre brugere. RettighedsAlliancen mener, at upload af digitale böger til chatbots er en ophavsretskrænkelse, fordi materialet behandles og potentielt kan gengives uden tilladelse fra rettighedshaveren. RettighedsAlliancen (2026) vil derfor søge at lægge sag an mod privatperson for sådanne overtrædelser med det formål at skabe større opmærksomhed om reglerne. Mens der findes en række domme for studerende, der har solgt adgang til digitale kopier af böger, så findes endnu ikke i Danmark domme for gratis at give en digital kopi til andre. At der ikke findes domme er ikke nødvendigvis et udtryk for at jura'en omkring ophavsret er usikker, men mere at det i praksis er svært at bevise at der er sket en overtrædelse.

Det er derfor stadigvæk et åbent spørgsmål, hvordan denne type sager vil udvikle sig i praksis. Man kan generelt forestille sig mindst to forskellige udviklingsretninger.

Den første mulighed er, at ophavsretten i højere grad bliver håndhævet på dette område. Hvis RettighedsAlliance finder effektive måder at bevise overtrædelser på, og domstole vælger at give dem medhold, så bør du nøje overveje, om du faktisk har rettighederne til det materiale, du uploader til en AI-agent. Det gælder for eksempel, hvis man vedlægger artikler, böger eller andre tekster som referencemateriale i en CustomGPT eller et lignende system.

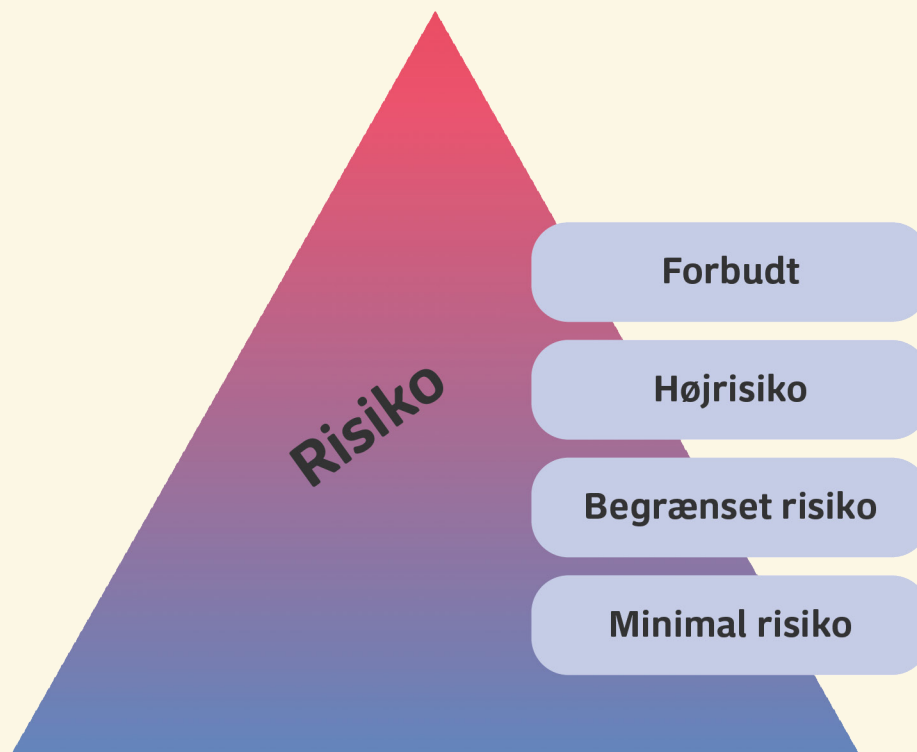
Den anden mulighed er, at reguleringen i EU har bestik af den globale konkurrencesituation. Tidligere havde EU et forbud mod uopfordrede e-mails i erhvervskommunikation, hvilket betød at det var ulovligt for EU virksomheder at sende uopfordrede e-mails til andre virksomheder. I praksis viste reglen sig imidlertid vanskelig at håndhæve, fordi mange af de virksomheder, der sendte spam, befandt sig uden for EU. Da udgangspunktet i international handel er, at afsenderens lands lov gælder, modtog europæiske virksomheder fortsat store mængder spam fra udlandet. Resultatet blev, at EU senere ophævede forbuddet. Noget lignende kan i princippet ske med reguleringen af ophavsret for kunstig intelligens. Hvis reglerne viser sig vanskelige at håndhæve internationalt, kan de i praksis få mindre betydning end først antaget.

For brugere af AI betyder denne usikkerhed først og fremmest, at man bør være opmærksom på, at reglerne på området stadig er under udvikling. Ophavsretten gælder også i en verden med kunstig intelligens, og du bør ikke uploade dokumenter til en AI, som du ikke selv har rettigheder til.

Ud over ophavsretten er der i de senere år også kommet en mere generel regulering af kunstig intelligens. Den vigtigste af disse regler i Europa er EU's AI-forordning (European Parliament & Council, 2024).

Mens ophavsretten primært handler om ejerskab til indhold, handler AI-forordningen i højere grad om, hvordan teknologien må bruges. Forordningen har til formål at sikre, at kunstig intelligens anvendes på en måde, der beskytter borgeres rettigheder og samtidig giver virksomheder mulighed for at udvikle og anvende teknologien.

AI-forordningen bygger på en risikobaseret tilgang. Det betyder, at AI-systemer reguleres forskelligt afhængigt af den risiko, de kan udgøre for mennesker og samfund. Jo større risiko et system anses for at have, desto flere krav stilles der til udviklere og brugere.



Figur 24 - Risiko

I den ene ende af skalaen findes en række anvendelser, som forordningen helt forbyder. Disse kaldes ofte uacceptable risici. Forbuddet omfatter blandt andet systemer, der forsøger at manipulere menneskers adfærd på en måde, der kan skade dem, systemer der udnytter sårbare grupper som børn eller personer med handicap, samt visse former for social scoring, hvor borgere vurderes og rangeres på baggrund af deres adfærd. Forordningen indeholder også begrænsninger i brugen af biometrisk overvågning i offentlige rum.

I den anden ende findes AI-systemer, der anses for at have lav eller minimal risiko. Disse systemer er i praksis kun underlagt få krav.

Den mest omfattende regulering gælder imidlertid de såkaldte højrisiko-systemer. Det er AI-systemer, som anvendes i områder, hvor fejl kan få væsentlige konsekvenser for mennesker. Eksempler kan være systemer, der bruges i rekruttering, kreditvurdering, uddannelsesvurderinger eller kritisk infrastruktur.

For disse systemer opstiller AI-forordningen en række konkrete krav til udvikling, dokumentation og drift. Formålet er at sikre, at systemerne er gennemsigtige, kontrollerbare og teknisk pålidelige.

Tabel 8 - Krav til højrisko-systemer

Krav	Kort forklaring	§
Risikostyringssystem	Udvikleren skal etablere en systematisk proces til at identificere, analysere og reducere risici ved systemet	Artikel 9
Datakvalitet og datagovernance	Trænings-, validerings- og testdata skal være relevante, repræsentative og af tilstrækkelig kvalitet	Artikel 10
Teknisk dokumentation	Der skal udarbejdes dokumentation, som gør det muligt for myndigheder at forstå og vurdere systemet	Artikel 11
Logningsforpligtelser	Systemet skal registrere relevante hændelser, så det er muligt at spore og analysere fejl	Artikel 12
Gennemsigtighed	Brugere af systemet skal have information om, hvordan systemet fungerer og hvordan det bør anvendes	Artikel 13
Menneskelig kontrol	Systemet skal designes, så mennesker kan overvåge og om nødvendigt gribe ind	Artikel 14
Nøjagtighed, robusthed og cybersikkerhed	Systemet skal opfylde krav til teknisk stabilitet, sikkerhed og modstandsdygtighed over for fejl	Artikel 15

Selvom disse regler primært retter sig mod udviklere og virksomheder, der udbyder AI-systemer, kan de også få betydning for almindelige brugere. Hvis en organisation anvender et højrisko-system, vil den for eksempel have ansvar for korrekt implementering, overvågning og brug af systemet.

En anden del af AI-forordningen handler om gennemsigtighed i brugen af kunstig intelligens. Her siger forordningen, at brugere skal informeres om, at indhold er skabt eller manipuleret ved hjælp af AI (European Parliament & Council, 2024).

Formålet med disse regler er ikke at forhindre brugen af generativ kunstig intelligens, men at sikre, at mennesker ikke bliver vildledt – meget i tråd med EU's lovgivning omkring markedsføring, der også forbyder vildledende markedsføring.

AI-forordningen indeholder derfor også en række gennemsigtighedskrav. For det første skal personer oplyses om det, hvis de interagerer direkte med et AI-system, og det ikke er indlysende, at der er tale om en maskine. Dette gælder især chatbots og andre systemer, der kan give indtryk af at være et menneske.

For det andet skal udbydere af generative AI-systemer sikre, at AI-genereret indhold kan identificeres som kunstigt genereret eller manipuleret. Denne mærkning kan ske teknisk, for eksempel gennem metadata eller andre maskinlæsbare markeringer. En undtagelse er dog såkaldte deepfakes. Hvis et AI-system bruges til at generere eller manipulere billede, lyd eller video, så det fremstår realistisk, skal det tydeligt oplyses, at indholdet er kunstigt genereret eller manipuleret, og ikke bare angives i metadata.

For almindelige brugere af kunstig intelligens betyder det, at du bør overveje, hvordan dit AI-genereret indhold præsenteres. Hvis materialet anvendes i en sammenhæng, hvor andre kan blive vildledt om dets oprindelse, kan der være krav om tydelig mærkning.

Med andre ord: AI er ikke et lovløst rum. De fleste af de regler, der gælder for information, ophavsret og gennemsigtighed i dag, gælder også når arbejdet udføres med hjælp fra kunstig intelligens.



Figur 25 - Gennemsigtighed og AI-forordningen

Prompting som kompetence

I begyndelsen af denne guide blev generativ kunstig intelligens sammenlignet med et samfund, hvor alle pludselig har fået en bil uden at vide, hvordan man kører den. Billedet var tænkt som en illustration af den situation, mange organisationer og brugere står i i dag. Teknologien er tilgængelig, men praksis er stadig ved at blive udviklet.

Men prompting er måske ikke en teknisk disciplin i samme forstand som programmering eller fysisk. Selvom generativ AI bygger på avancerede matematiske modeller og enorme datasæt, er selve interaktionen med systemerne først og fremmest sproglig. At arbejde med generativ AI handler derfor i høj grad om at kunne formulere sit gøremål tydeligt, strukturere information og reflektere over svarene.

En vigtig pointe er, at kvaliteten af en prompt ofte minder om kvaliteten af god kommunikation generelt. Når du skriver en prompt, stiller du i praksis de samme spørgsmål, som retorikere og kommunikationsteoretikere har arbejdet med i århundreder. Hvem taler? Hvad er formålet? Hvem er modtageren? Hvilket format og hvilken tone passer til situationen? De aspekter for klarhed, der blev gennemgået, er derfor ikke en teknisk opskrift, men snarere en måde at tænke over kommunikation på.

Samtidig bør du huske, at generativ AI fungerer fundamentalt anderledes end klassisk software. En sprogmodel er ikke en database eller en regnemaskine, men et statistisk system, der genererer tekst ud fra sandsynligheder. Dette betyder blandt andet, at de samme prompts kan give forskellige svar, og at svarene kan være præget af bias fra træningsdata, træningsmetoder eller din egen formuleringen af prompten.

Disse egenskaber gør det nødvendigt at arbejde kritisk og reflektivt med AI-systemer. En enkelt generering bør sjældent opfattes som et definitivt svar. I stedet er det ofte mere frugtbart at arbejde iterativt, hvor prompts, svar og brugerens egen forståelse udvikler sig i en proces. Denne proces blev i bogen beskrevet som AI-hermeneutik, hvor forståelse opstår gennem en løbende vekselvirkning mellem spørgsmål og fortolkning.

Set i dette perspektiv bliver generativ AI ikke blot et værktøj til at producere tekst, men en form for dialogpartner i arbejdet med viden. Den kan hjælpe med at formulere idéer, udforske perspektiver og generere forslag, men den kan ikke erstatte menneskelig vurdering. Det er stadig brugeren, der må afgøre, om et svar er plausibelt, relevant eller brugbart.

En videre central pointe har været, at generativ AI ikke eksisterer i et tomrum. Teknologien indgår i organisatoriske og kulturelle sammenhænge. I mange virksomheder og institutioner bliver AI allerede brugt som en form for skygge-IT, hvor medarbejdere eksperimenterer med værktøjerne uden nødvendigvis at have klare retningslinjer eller fælles praksisser. Dette kan både skabe innovation og usikkerhed.

Derfor er prompting ikke kun en individuel færdighed, men også en organisatorisk kompetence. Organisationer må gradvist udvikle fælles erfaringer med, hvordan generativ AI kan anvendes ansvarligt og meningsfuldt i konkrete arbejdsopgaver. Dette kræver både teknisk forståelse, kritisk refleksion og en bevidsthed om de begrænsninger, der følger med teknologien.

Et sidste tema i bogen har været forholdet mellem teknologi og læring. Nye teknologier for viden har historisk set altid skabt bekymringer om, hvordan mennesker tænker og lærer. Generativ AI rejser lignende spørgsmål. Hvis komplekse opgaver i stigende grad løses gennem dialog med AI-systemer, kan der opstå en risiko for, at nogle former for dyb problemløsning bliver mindre fremtrædende. Læring bliver i endnu højere grad en bevidst handling.

Samtidig kan teknologien også åbne nye muligheder. Hvis den bruges reflektivt, kan generativ AI fungere som et frugtbart redskab til at udforske idéer, teste hypoteser og udvikle forståelse. Forskellen ligger ikke i teknologien alene, men i den måde du vælger at bruge den på.

På den måde vender vi tilbage til udgangspunktet for denne guide. Prompting er ikke en magisk teknik, og der findes ikke en universel formel for den perfekte prompt. Det afgørende er snarere en kombination af tre kompetencer: evnen til at formulere sig klart, evnen til at vurdere og fortolke svar kritisk, og evnen til at forstå den teknologiske kontekst, hvori svarene bliver genereret.

At arbejde med generativ AI handler mere om at udvikle en reflekteret måde at stille spørgsmål på og mindre om at mestre en bestemt teknik at skrive på

Generativ AI vil sandsynligvis fortsætte med at udvikle sig hurtigt i de kommende år. Nye modeller, nye anvendelser og nye reguleringer vil løbende ændre den måde, teknologien indgår i arbejde og læring. Det betyder også, at prompting ikke bør forstås som en færdighed, man lærer én gang for alle. Det er snarere en praksis, der udvikler sig i takt med teknologien.

Hvis der er én overordnet pointe i denne guide, er det derfor denne: At arbejde godt med generativ AI handler i sidste ende mindre om at mestre en bestemt teknik og mere om at udvikle en reflekteret måde at stille spørgsmål på. I en tid hvor maskiner kan generere svar på næsten alt, bliver evnen til at formulere gode spørgsmål måske vigtigere end nogensinde før.



Figur 26 - Reflektion gennem nysgerrighed

Videre læsning

Acar, O. A. (2023). AI Prompt Engineering Isn't the Future: Asking the perfect question is less important than really understanding the problem you're trying to solve. *Harvard Business Review Digital Articles*, 1–7.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., et al. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*. <https://arxiv.org/abs/2108.07258>

Brighton, H., & Gigerenzer, G. (2015). The bias bias. *Journal of Business Research*, 68(8), 1772-1784.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1–7), 107–117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM*, 36(12), 66–77. <https://doi.org/10.1145/163298.163309>

Cao, Y., Zhou, L., Lee, S., Cabello, L., Chen, M. & Hershovich, D. (2023). Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study. In *Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, pages 53–67, Dubrovnik, Croatia. Association for Computational Linguistics.

Chaffey, D., & Ellis-Chadwick, F. (2019). *Digital marketing (7th ed.)*. Pearson Education.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, 70(11), 35–36.

European Parliament & Council. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.

Feng, S., McDonald, R., & Zhang, X. (2025). Agentic AI: Levels of autonomy in LLM-based systems. (Working paper / preprint).

Gadamer, H.-G. (2004) *Truth and method (2nd rev. ed.)*. Continuum.

Hellström, T., Dignum, V., & Bensch, S. (2020). Bias in Machine Learning--What is it Good for?. *arXiv preprint arXiv:2004.00686*.

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture*, 2(1), 8.

Humlum, A., & Vestergaard, E. (2025). The labor market effects of AI chatbots. *University of Chicago / University of Copenhagen Working Paper*.

https://static1.squarespace.com/static/5d35e72fcff15f0001b48fc2/t/67fdc-07c74f1302536156227/1744683136315/chatbots_apr25.pdf

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>

Licklider, J. C. R. (1960), "Man-Computer Symbiosis," IRE Transactions on Human Factors in Eletronics, pp. 4 – 11, 1960

Lowe m.fl. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems (NeurIPS).
<https://arxiv.org/abs/2203.02155>

March, J. G., & Olsen, J. P. (2011). The logic of appropriateness.

Mesko, B. (2023). Prompt engineering: The art of communicating with AI. The Medical Futurist.

Minto, B. (1996). The pyramid principle: Logic in writing and thinking. London: Financial Times Management.

Moaveni, S., & Chou, J. (2017). Engineering fundamentals: An introduction to engineering. Boston: Cengage Learning.

OpenAI (2024), <https://platform.openai.com/docs/guides/prompt-engineering>, tilgået den 17. februar 2024.

OpenAI (2026) <https://platform.openai.com/tokenizer>, tilgået den 14. februar 2026.

Pedersen & Bergfors (2025) Virtuelle Assistenten, <https://www.eaviden.dk/project/virtuelle-assistenten/>-tilgået den 17. februar 2026.

Plato. (1997). Phaedrus. I J. M. Cooper (Red.), Plato: Complete works (s. 506–556). Hackett Publishing. (Oprindeligt udgivet ca. 370 f.v.t.)

RettighedsAlliancen. (2025a). AI-udbydere bruger ulovligt indhold. <https://rettighedsalliancen.dk/videovideopraesentation-ai-udbyderes-brug-ulovligt-indhold/> - tilgået den 17. februar 2026.

RettighedsAlliancen. (2025b). Hver fjerde studerende uploader bøger ulovligt til chatbots. <https://rettighedsalliancen.dk/hver-fjerde-studerende-uploader-boeger-ulovligt-til-chatbots/> - tilgået den 17. februar 2026.

RettighedsAlliancen. (2026). RettighedsAlliancen tager nye midler i brug mod ulovlig deling af studieboøger. <https://rettighedsalliancen.dk/pressemeddelelse-rettighedsalliancen-tager-nye-midler-i-brug-mod-ulovlig-deling-af-studieboger/> - tilgået den 17. februar 2026.

Rio-Chanona, R. M., et al. (2025). The economic impacts of generative AI: Evidence from experiments and field studies. Working paper.

Robertson, A. (2024) Google apologizes for 'missing the mark' after Gemini generated racially diverse Nazis, <https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical> - tilgået den 9. Marts 2026

Roer, H. (2023), det retoriske pentagram, https://denstoredanske.lex.dk/det_retoriske_pentagram - tilgået den 17. februar 2024.

Saunders, M., Lewis, P., & Thornhill, A. (2013). Research methods for business students (6th ed.). Harlow: Pearson.

Scarfe et al, 2024, A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0305354>

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457. <https://doi.org/10.1017/S0140525X00005756>

Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.

Silic, M., & Back, A. (2014) Shadow IT – A view from behind the curtain. *Computers & Security*, 45, 274–283. <https://doi.org/10.1016/j.cose.2014.06.007>

Wang, L., Ma, Y., Zhang, Q., Zhang, Y., Huang, Z., & Zhao, Y. (2024). A Survey on Large Language Model based Autonomous Agents. arXiv.

Weber Wulf et al. *International Journal for Educational Integrity* (2023) 19:26 <https://doi.org/10.1007/s4097902300146z>

Wedell-Wedellsborg, T. (2020). *What's your problem? To solve your toughest problems, change the problems you solve*. Boston: Harvard Business Review Press.

Figur- og tabelfortegnelse

Figurer

- Figur 1 - Den dårligste AI du nogensinde får - tidslinje (s. 4)
- Figur 2 - Forsimplet eksempel på numerisk tekstanalyse (s. 5)
- Figur 3 - Det kinesiske rum (s. 6)
- Figur 4 - 3 generationer af Human-Machine-Interface (s. 9)
- Figur 5 - OpenAI's strategier som tankeproces (s. 15)
- Figur 6 - Problem -> Prompt -> Proces (s. 16)
- Figur 7 - Underliggende problemanalyse (s. 20)
- Figur 8 - Problempdeling (s. 21)
- Figur 9 - Reframing (s. 22)
- Figur 10 - Fra god tale til god prompting (s. 27)
- Figur 11 - Hvilket aspekt er vigtigst (s. 28)
- Figur 12 - AI-hermeneutik (s. 29)
- Figur 13 - AI som skygge-IT (s. 32)
- Figur 14 - Varians (s. 34)
- Figur 15 - Bias (s. 35)
- Figur 16 - Kilder til bias (s. 35)
- Figur 17 - Prompten som kilde til bias (s. 37)
- Figur 18 - Tokens og sprog (s. 37)
- Figur 19 - Kulturelle forskelle (s. 38)
- Figur 20 - Politisk bias (s. 39)
- Figur 21 - Bias i systemprompten (s. 40)
- Figur 22 - Agentopsætning (s. 42)
- Figur 23 - AI-agent til IT-support (s. 43)
- Figur 24 - Risiko (s. 49)
- Figur 25 - Gennemsigtighed og AI-forordningen (s. 51)
- Figur 26 - Refleksion gennem nysgerrighed (s. 53)

Tabeller

- Tabel 1 - 3 former for kunstig intelligens (s. 8)
- Tabel 2 - OpenAI's råd til prompting (s. 12)
- Tabel 3 - Problemforståelse (s. 18)
- Tabel 4 - 6 aspekter for klarhed (s. 23)
- Tabel 5 - Autonomi (s. 44)
- Tabel 6 - Kriterier for agenter (s. 46)
- Tabel 7 - Perspektiv på jura (s. 47)
- Tabel 8 - Krav til højrisiko-systemer (s. 50)

Om forfatteren



Just Kjærgård Pedersen er koordinator for e-handelsuddannelserne ved Erhvervsakademi København. Han har tidligere været forretningsudvikler og ledelseskonsulent hos blandt andet Andel, BP, Ørsted, KPMG og Implement. Han skriver ofte om AI og holder oplæg om kunstig intelligens, læring og digital udvikling.