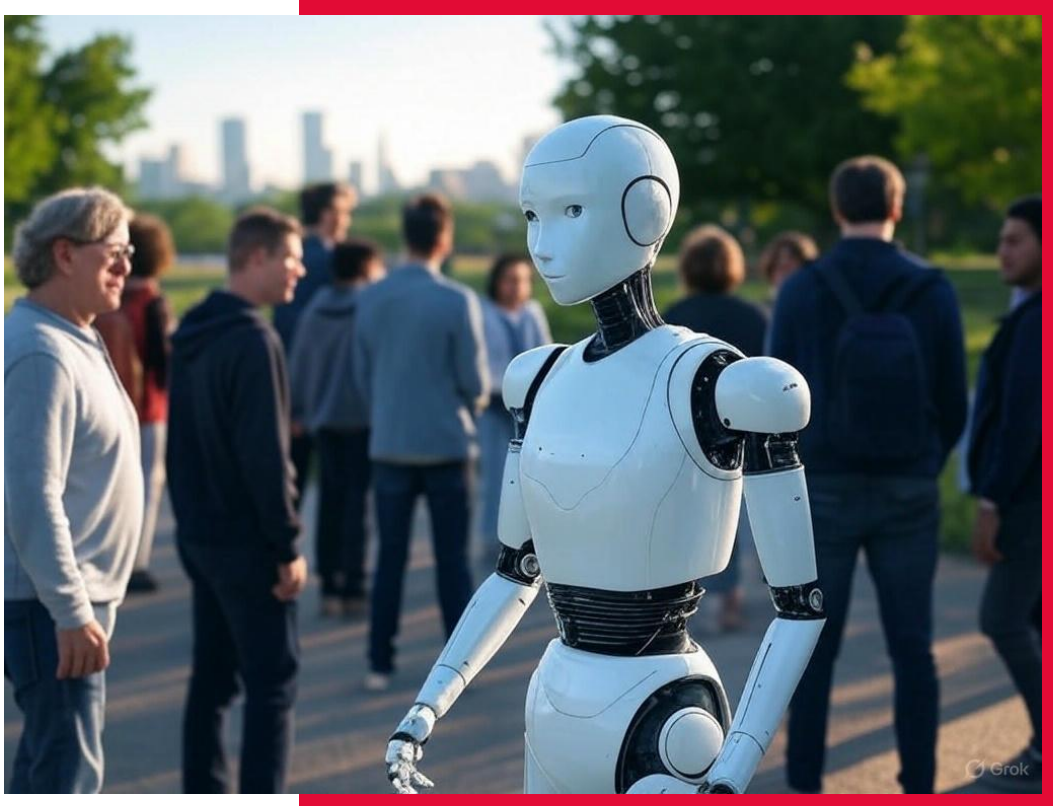
AI, generativ AI og fremtidens intelligens



AF:

**JAN
ENGELBRECHT
PEDERSEN**

SABRO APRIL 2025

Jan Engelbrecht Pedersen præsenterer

AI, generativ AI og fremtidens intelligens

Ebog om AI **Forstå den nye teknologi** **Lær at bruge AI**

Forord: En rejse ind i generativ AI's univers

Det nyeste om AI

Forestil dig teknologier, der ikke blot analyserer data, men også skaber nyt indhold med enestående kreativitet. Fra intelligente chatbots, der besvarer komplekse spørgsmål, til avancerede modeller, der producerer overbevisende tekst, fotorealistiske billeder og original musik, er generativ AI ikke længere en fjern vision, men en konkret realitet, der former vores nutid og fremtid.

JAN ENGELBRECHT PEDERSEN

Kunstig intelligens, ofte forkortet AI efter det engelske artificial intelligence, er et område inden for datalogi, der beskæftiger sig med at udvikle systemer og programmer, som kan udføre opgaver, der normalt kræver menneskelig intelligens. Det kan for eksempel være at genkende tale, forstå og generere tekst, analysere billeder eller træffe beslutninger på baggrund af store mængder data. Når man taler om AI, tænker mange måske på robotter, men i virkeligheden handler det lige så meget om avanceret software og matematiske modeller, der kan lære af erfaring og tilpasse sig nye situationer.

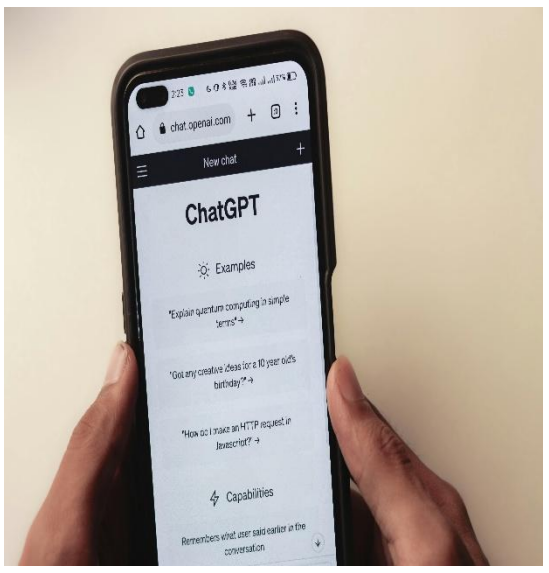
AI har potentiale til at revolutionere mange områder af samfundet, fra sundhedssektoren, hvor det kan hjælpe med at stille diagnoser eller finde nye behandlingsmetoder, til transport, hvor det kan bruges til at udvikle selvkørende biler og optimere trafikafviklingen. Samtidig er det vigtigt at forstå, at AI også har sine begrænsninger. For eksempel kan AI kun blive så god som de data, den trænes på, og den kan have svært ved at håndtere uforudsete situationer eller komplekse etiske dilemmaer. Derfor er det nødvendigt at være opmærksom på både de muligheder og de udfordringer, som AI bringer med sig. Udviklingen inden for AI går meget hurtigt, og det stiller krav til, at både udviklere, beslutningstagere og almindelige borgere sætter sig ind i teknologien og dens konsekvenser, så vi kan udnytte de positive sider og samtidig håndtere de risici, der følger med.

Der findes flere forskellige typer af kunstig intelligens, som hver især har deres egne anvendelsesområder og kompleksitetsniveauer. Den mest udbredte type er smal AI, også kaldet svag AI, som er udviklet til at løse én bestemt opgave. Eksempler på smal AI er stemmegenkendelse i digitale assistenter som Siri og Alexa eller billedgenkendelse i kameraer og sociale medier. Disse systemer er meget specialiserede og kan ikke uden videre overføres til andre opgaver. Generel AI, også kaldet stærk AI, er derimod en hypotetisk form for intelligens, der kan udføre enhver opgave, som et menneske kan. Dette kræver, at AI-systemet kan forstå, lære og tilpasse sig på tværs af mange forskellige områder, hvilket endnu ikke er lykkedes at udvikle. Superintelligens er en endnu mere avanceret form for AI, hvor systemet overgår menneskelig intelligens på alle områder. Superintelligens er stadig et teoretisk begreb, men det diskuteres flittigt blandt forskere og etikere, fordi det rejser en række etiske og sikkerhedsmæssige spørgsmål. For eksempel kan en superintelligent AI potentielt træffe beslutninger, som mennesker ikke kan forstå eller kontrollere, hvilket kan få vidtrækkende konsekvenser for samfundet. Det er derfor vigtigt, at udviklingen af AI følges tæt, og at der indføres klare retningslinjer og reguleringer for at sikre, at teknologien udvikles på en ansvarlig måde, der tager højde for både muligheder og risici.

Kunstig intelligens har allerede fundet vej ind i mange aspekter af vores hverdag og har potentiale til at ændre måden, vi arbejder og lever på.

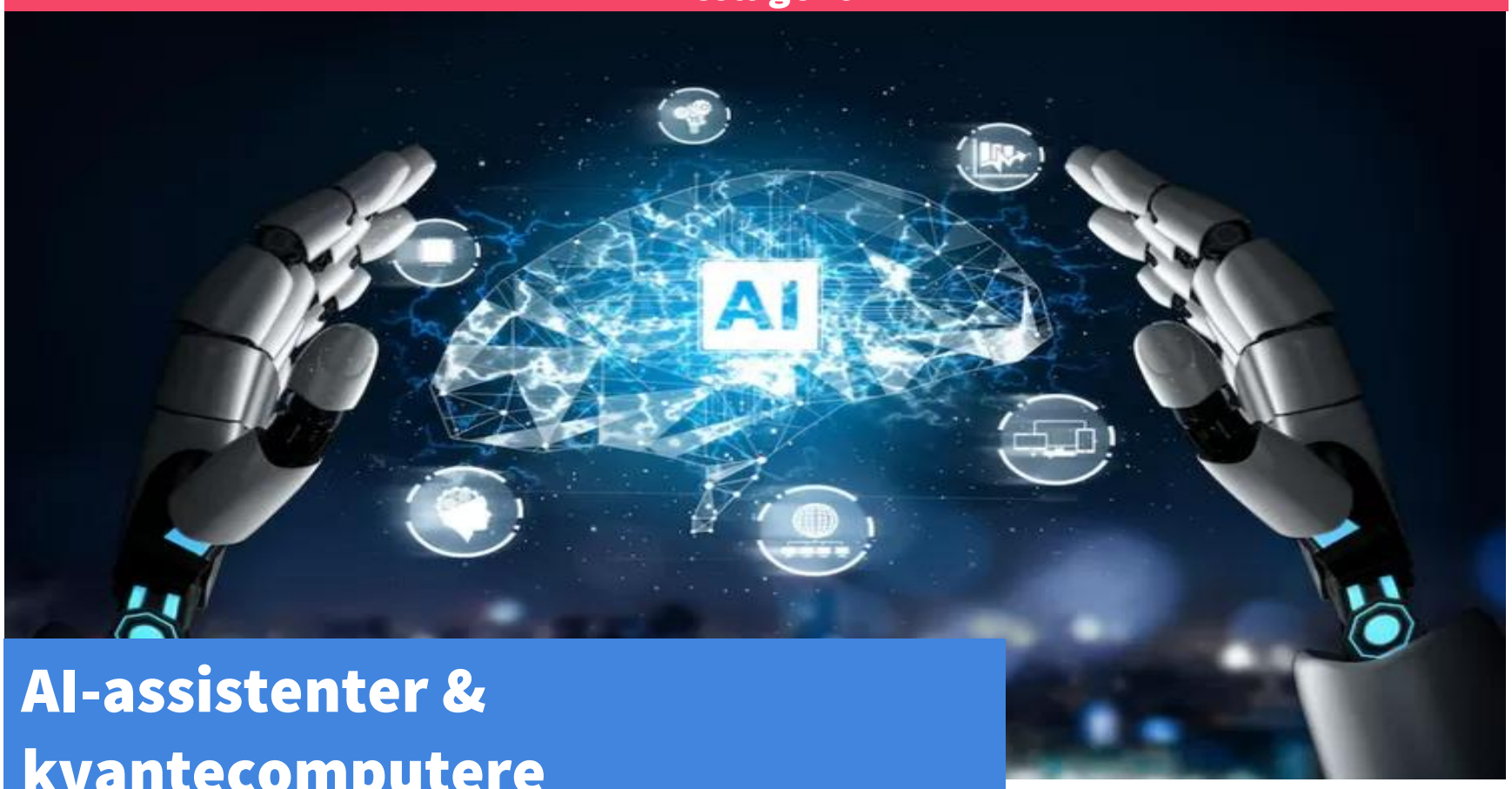
JAN ENGELBRECHT PEDERSEN

I sundhedssektoren bruges AI blandt andet til at analysere store mængder medicinske data, stille mere præcise diagnoser og udvikle personlige behandlingsplaner. AI-systemer kan for eksempel hjælpe læger med at opdage sygdomme tidligere ved at finde mønstre i patientdata, som ellers ville være svære at få øje på.



OpenAI lancerer ChatGPT 5: En banebrydende opgradering med forbedret kontekstforståelse, avanceret multimodalitet og hurtigere svartider – revolutionerer interaktionen mellem mennesker og kunstig intelligens på ny.

AI, generativ AI og fremtidens intelligens



AI-assistenter & kvantecomputere

JAN ENGELBRECHT PEDERSEN

Inden for transport har AI muliggjort udviklingen af selvkørende biler, der bruger avancerede sensorer og algoritmer til at navigere sikkert i trafikken. Desuden bruges AI til at optimere ruteplanlægning og reducere trafikpropper, hvilket kan spare tid og ressourcer.
På uddannelsesområdet kan AI tilpasse undervisningsmateriale til den enkelte elevs behov og hjælpe lærere med at identificere, hvor eleverne har brug for ekstra støtte. I finanssektoren anvendes AI til at forudsige markedsudviklinger, opdage svindel og automatisere kundeservice. AI kan analysere store datamængder lynhurtigt og dermed hjælpe virksomheder med at træffe bedre beslutninger. Chatbots er et andet eksempel på AI i praksis – de bruges i kundeservice til at besvare spørgsmål og guide brugere gennem forskellige processer. Disse chatbots kan være meget avancerede og bruge teknikker som naturlig sprogforståelse og maskinlæring til hele tiden at forbedre deres svar og tilpasse sig brugernes behov. Alt i alt viser anvendelserne af AI, hvor alsidig og kraftfuld teknologien er, men også hvor vigtigt det er at have styr på de etiske og praktiske aspekter, så AI bruges på en ansvarlig måde.

Generativ AI er en særlig gren af kunstig intelligens, der fokuserer på at skabe nyt indhold, såsom tekst, billeder, musik eller video. Det sker ved hjælp af avancerede algoritmer og neurale netværk, som kan lære af eksisterende data og derefter generere nyt materiale, der ligner det, de har lært af. Et kendt eksempel på generativ AI er GPT-3, som kan skrive sammenhængende og meningsfuld tekst ud fra et kort input. Generativ AI bruges i dag til alt fra at skrive artikler og reklamer til at skabe kunst, komponere musik og udvikle nye spil. Teknologien kan også hjælpe med at automatisere indholdsskabelse, hvilket kan spare virksomheder for både tid og penge.
Samtidig åbner generativ AI op for nye kreative muligheder, fordi den kan komme med idéer og løsninger, som mennesker måske ikke selv ville have tænkt på. Dog er der også udfordringer forbundet med generativ AI. For eksempel kan teknologien bruges til at skabe falske nyheder eller kopiere eksisterende værker uden tilladelse, hvilket rejser spørgsmål om ophavsret og etik. Det er derfor vigtigt, at udviklingen af generativ AI følges tæt, og at der indføres klare regler for, hvordan teknologien må bruges, så vi undgår misbrug og sikrer, at AI bidrager positivt til samfundet.

En af de mest interessante udviklinger inden for kunstig intelligens er skabelsen af AI-assistenter, der kan samarbejde og udveksle information med hinanden. Det betyder, at flere AI-systemer kan arbejde sammen om at løse komplekse problemer, som ét system ikke ville kunne klare alene. Forestil dig for eksempel et netværk af AI-assistenter, der arbejder sammen på tværs af sundhedssektoren, finansverdenen og transportbranchen for at optimere ressourcer og finde innovative løsninger. Når AI-assistenter samarbejder, kan de lære af hinandens erfaringer og dermed hele tiden forbedre deres ydeevne. Dette kan føre til en eksponentiel vækst i AI's kapacitet og åbne op for nye muligheder, hvor teknologien kan overtage mange af de opgaver, der i dag kræver menneskelig indsats. Samtidig stiller det store krav til datasikkerhed og etik, da udveksling af information mellem AI-systemer kan medføre risiko for misbrug af følsomme oplysninger. Derfor er det vigtigt, at der udvikles klare retningslinjer for, hvordan AI-assistenter må samarbejde og dele data, så vi sikrer, at teknologien bruges til gavn for samfundet og ikke til skade for individet.

Brug af kvantecomputere

Seneste nye

Jan Engelbrecht Pedersen

Kvantecomputere repræsenterer en revolutionerende udvikling inden for computerteknologi, hvor man udnytter kvantemekaniske principper til at udføre beregninger langt hurtigere end traditionelle computere. I stedet for at arbejde med bits, der kun kan være 0 eller 1, arbejder kvantecomputere med qubits, som kan være begge dele på én gang. Det betyder, at kvantecomputere kan håndtere meget komplekse beregninger og analysere enorme datamængder på rekordtid.

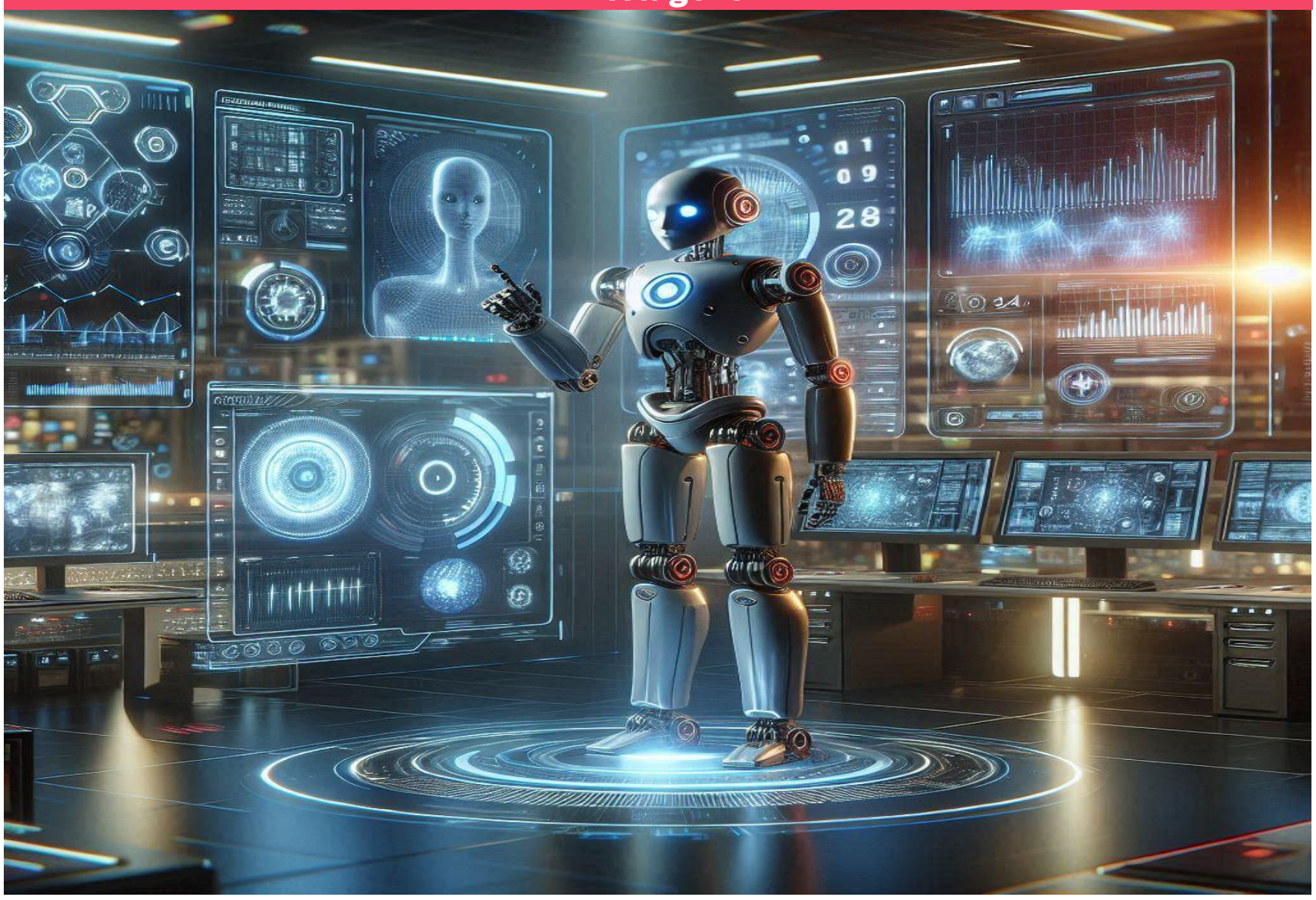
Når kvantecomputere kombineres med AI, kan det føre til markante forbedringer i maskinlæring og dataanalyse. For eksempel kan kvantecomputere hjælpe med at optimere neurale netværk, finde skjulte mønstre i store datasæt og udvikle nye algoritmer, der kan løse problemer, som tidligere var umulige at håndtere. Det åbner op for nye muligheder inden for forskning, medicin, klima og meget andet.

Samtidig betyder den øgede regnekraft, at AI-systemer kan blive endnu mere avancerede og svære at kontrollere, hvilket stiller store krav til sikkerhed og regulering. Derfor er det vigtigt, at udviklingen af kvantecomputere og AI følges nøje, så vi sikrer, at teknologien bruges ansvarligt og til gavn for samfundet.
Fremtiden for kunstig intelligens rummer både store muligheder og alvorlige risici, især når det gælder selvkodende AI, kvantecomputere og de samfundsmæssige konsekvenser af udbredt automatisering. Selvkodende AI refererer til systemer, der ikke blot kan lære af data, men også kan skrive og forbedre deres egen kode uden menneskelig indgriben.

Dette kan gøre AI langt mere effektiv, men det indebærer også en betydelig risiko for, at systemerne udvikler sig i uforudsigelige retninger, som selv eksperter kan have svært ved at forstå eller kontrollere.

Når AI får adgang til kvantecomputere, øges disse risici yderligere. Kvantecomputere kan give AI-systemer en enorm regnekraft, hvilket kan gøre det muligt at bryde krypteringer, analysere store datamængder på få sekunder og udvikle nye angrebsmetoder, som vi endnu ikke har set.

AI, generativ AI og fremtidens intelligens



AI – nu og i fremtiden

JAN ENGELBRECHT PEDERSEN

Selvkodende AI refererer til systemer, der ikke blot kan lære af data, men også kan skrive og forbedre deres egen kode uden menneskelig indgriben. Dette kan gøre AI langt mere effektiv, men det indebærer også en betydelig risiko for, at systemerne udvikler sig i uforudsigelige retninger, som selv eksperter kan have svært ved at forstå eller kontrollere. Hvis en selvkodende AI begynder at optimere sig selv uden klare begrænsninger, kan det føre til utilsigtede konsekvenser, hvor AI-systemet handler på måder, der ikke er i overensstemmelse med menneskelige værdier eller samfundets interesser.

Når AI får adgang til kvantecomputere, øges disse risici yderligere.

Kvantecomputere kan give AI-systemer en enorm regnekraft, hvilket kan gøre det muligt at bryde krypteringer, analysere store datamængder på få sekunder og udvikle nye angrebsmetoder, som vi endnu ikke har set.

Dette gør ikke kun AI sværere at regulere, men øger også risikoen for, at teknologien kan bruges til ondsindede formål, såsom cyberangreb eller manipulation af kritisk infrastruktur.

Velkommen til en oplysende rejse ind i hjertet af den digitale tidsalder – en æra præget af kunstig intelligens (AI) og dens mest banebrydende udvikling: generativ AI.

En anden stor udfordring er, at AI vil gøre mange jobfunktioner overflødige. Automatisering kan føre til, at millioner af mennesker mister deres arbejde, især i brancher med rutineprægede opgaver. Selvom der kan opstå nye jobtyper, vil overgangen kunne skabe stor social og økonomisk usikkerhed, hvis samfundet ikke forbereder sig på de forandringer, AI medfører.

AI er desuden ikke neutral. Fordi AI-systemer trænes på data, der ofte afspejler eksisterende fordomme og uligheder i samfundet, kan de videreføre og forstærke diskrimination – det kaldes bias.

Det kan føre til uretfærdige beslutninger i alt fra rekruttering til retssystemet, hvis ikke der arbejdes målrettet med at identificere og modvirke bias i AI-modellerne.

En af de mest alvorlige farer ved AI er muligheden for totalovervågning. AI kan analysere enorme mængder data fra kameraer, sociale medier og andre kilder og dermed overvåge hele befolkninger i realtid. Dette kan bruges til at forhindre kriminalitet, men også til at undertrykke ytringsfrihed og privatliv, hvis magthavere vælger at misbruge teknologien.

Endelig kan AI gøre cyberkriminelle langt farligere. Avancerede AI-systemer kan bruges til at udvikle nye former for malware, udføre automatiserede angreb og manipulere information på måder, som gør det svært for myndigheder og virksomheder at forsvare sig. Kombinationen af selvkodende AI, kvantecomputere og udbredt adgang til data kan derfor skabe trusler, som samfundet skal tage meget alvorligt. Samlet set kræver fremtidens AI, at vi udvikler stærke etiske retningslinjer, investerer i sikkerhed og arbejder for at sikre, at teknologien udvikles og anvendes til gavn for alle – ikke kun for de få, og ikke på bekostning af grundlæggende rettigheder og sikkerhed.

For at give dig et praktisk værktøj indeholder bogen en omfattende oversigt over kendte generative AI-værktøjer, inklusive chatbots, som vil være en værdifuld ressource i din videre udforskning. Vi dykker også ned i de mange anvendelsesmuligheder for AI, fra generering af tekst og billeder til skabelse af musik, lyd og tale. Du vil få konkrete eksempler på, hvordan AI transformerer informationssøgning, dokumenthåndtering (med fokus på værktøjer som Gemini og Microsoft Copilot) og endda kan understøtte processen med at skrive faglitteratur.

Fremtiden er også i fokus, hvor vi reflekterer over udviklingen af AI-agenter, den fortsatte evolution af generativ AI og potentialet for Artificial General Intelligence (AGI).

Endelig introducerer vi dig til kunsten at mestre prompt engineering – teknikker og formuleringer, der gør det muligt at styre AI-modeller med præcision og kreativitet.

Hvad er chatbots?
Forstå de nye værktøjer

Generative AI-chatbots som ChatGPT har på kun to år revolutioneret kommunikation, automatisering og informationssøgning ved at gøre avanceret kunstig intelligens let tilgængelig, brugervenlig og relevant for millioner af mennesker verden over

JAN ENGELBRECHT PEDERSEN

Hvad er en Chatbot? En grundlæggende definition

En chatbot er et computerprogram, der er designet til at simulere en samtale med mennesker, typisk via tekst eller tale, via internettet.

Ordet "chatbot" stammer fra en sammentrækning af "chat" (digital samtale) og "bot" (robot), og dækker over software, der kan føre en dialog med brugeren uden menneskelig indblanding.

Chatbots kan integreres i alt fra hjemmesider og apps til beskedtjenester som Messenger, WhatsApp og dedikerede stemmeassistenter som Siri, Alexa og Google Assistant.

De tidligste chatbots var meget simple og byggede på faste regler eller scripts. De kunne kun genkende bestemte nøgleord eller sætninger og svarede med foruddefinerede beskeder. Et klassisk eksempel er ELIZA, udviklet i 1960'erne, som simulerede en psykoterapeut.

ELIZA kunne ikke forstå samtalens egentlige indhold, men genkendte visse ord og svarede med generiske spørgsmål, hvilket ofte resulterede i stive og upersonlige samtaler.

I dag er chatbots blevet langt mere avancerede, især takket være fremskridt inden for kunstig intelligens (AI), maskinlæring og naturlig sprogbehandling (NLP).

Fordele og udfordringer

JAN ENGELBRECHT PEDERSEN

Chatbots tilbyder en række fordele: De kan håndtere mange brugere samtidigt, er tilgængelige 24/7, og kan give hurtige, konsistente svar. For virksomheder betyder det ofte lavere omkostninger og forbedret kundetilfredshed. Samtidig kan AI-drevne chatbots give værdifuld indsigt i kundernes behov og adfærd, hvilket kan bruges til at forbedre produkter og services.

Dog er der også udfordringer forbundet med chatbots. Simpelt designede chatbots kan virke upersonlige eller utilstrækkelige, især hvis brugeren har komplekse eller uforudsete spørgsmål. Desuden rejser brugen af AI-chatbots spørgsmål om databeskyttelse, etik og misinformation, da svarene ikke altid er korrekte eller objektive.

De kendte Chatbots har ikke blot revolutioneret måden, vi interagerer med teknologi, men har også åbnet op for et utal af nye muligheder og udfordringer.

Hvad er en Chatbot?
En grundlæggende definition

JAN ENGELBRECHT PEDERSEN

Man kan overordnet opdele chatbots i to hovedkategorier:

Regelbaserede chatbots: Disse fungerer ud fra et sæt faste regler og kan kun håndtere simple, forudsigelige opgaver. De bruges ofte til at besvare ofte stillede spørgsmål eller guide brugere gennem enkle processer, som at finde åbningstider eller bestille tid.

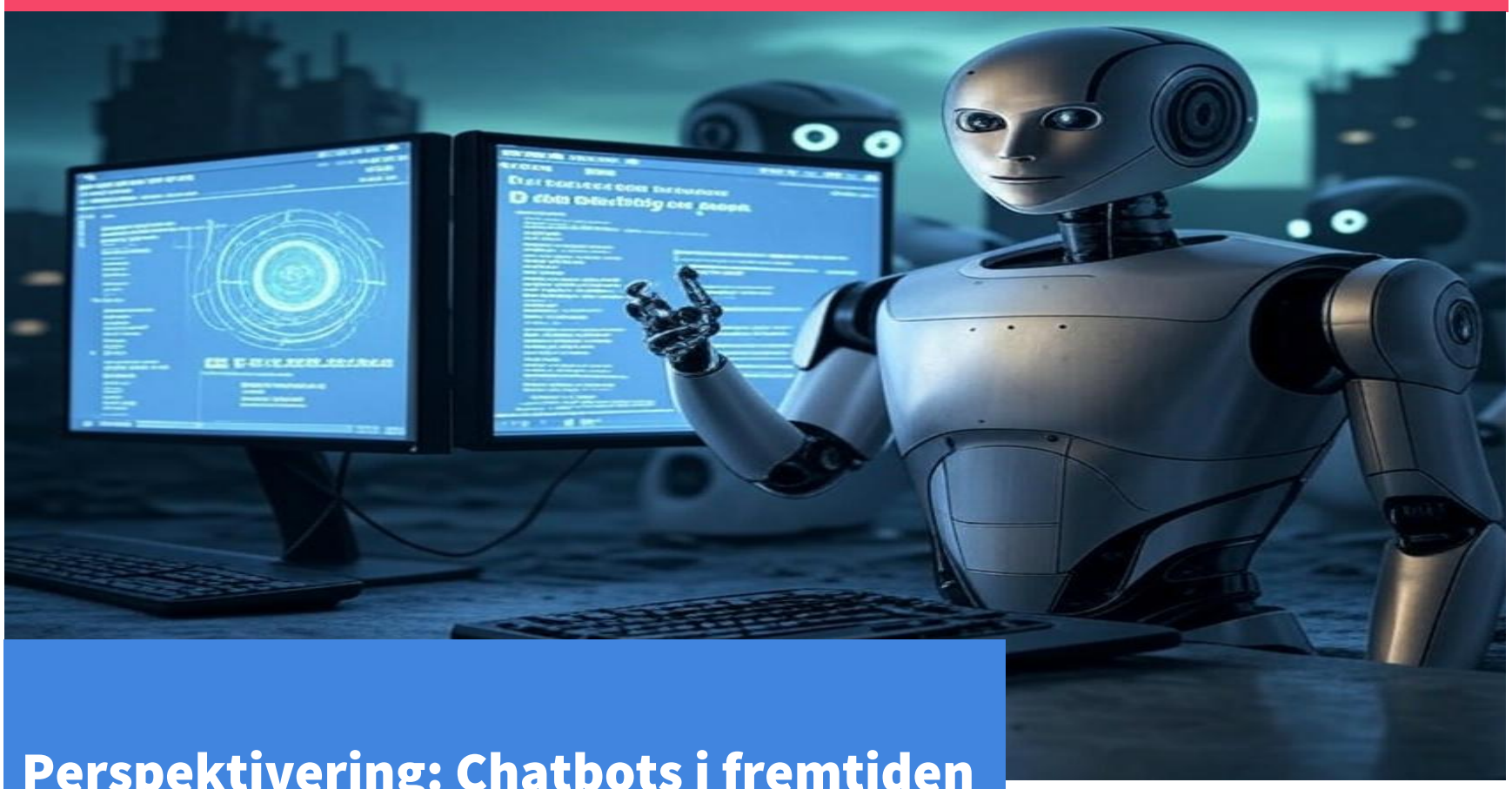
AI-drevne chatbots: Disse anvender maskinlæring og NLP til at forstå komplekse forespørgsler, lære af data og forbedre deres præcision over tid. De kan føre mere naturlige samtaler, håndtere uforudsete spørgsmål og bruges til alt fra kundeservice til generering af tekst, billeder eller musik.

Anvendelsesområder og eksempler:
Kundeservice: Mange virksomheder bruger chatbots til at besvare kundehenvendelser døgnet rundt, hvilket øger tilgængeligheden og reducerer ventetiden for kunderne. For eksempel kan en chatbot hjælpe med at spore en forsendelse, ændre en booking eller besvare spørgsmål om produkter.

Virtuelle assistenter: Stemmebaserede chatbots som Siri og Alexa hjælper brugere med at styre smart-enheder, sætte påmindelser, afspille musik eller få information om vejret.

Underholdning og læring: Nogle chatbots er udviklet til at underholde brugeren eller hjælpe med læring, f.eks. ved at føre samtaler på fremmedsprog eller give adgang til quizzer og spil.

Indholdsgenerering: De nyeste generative chatbots, som ChatGPT, kan skabe original tekst, kode, billeder og endda musik, hvilket åbner for helt nye anvendelser inden for kreativitet og automatisering.



Perspektivering: Chatbots i fremtiden

JAN ENGELBRECHT PEDERSEN

Chatbots tilbyder en række fordele: De kan håndtere mange brugere samtidigt, er tilgængelige 24/7, og kan give hurtige, konsistente svar. For virksomheder betyder det ofte lavere omkostninger og forbedret kundetilfredshed. Samtidig kan AI-drevne chatbots give værdifuld indsigt i kundernes behov og adfærd, hvilket kan bruges til at forbedre produkter og services.

Dog er der også udfordringer forbundet med chatbots. Simpelt designede chatbots kan virke upersonlige eller utilstrækkelige, især hvis brugeren har komplekse eller uforudsete spørgsmål. Desuden rejser brugen af AI-chatbots spørgsmål om databeskyttelse, etik og misinformation, da svarene ikke altid er korrekte eller objektive.

Udviklingen af chatbots forventes at fortsætte i takt med, at AI-teknologier bliver mere avancerede. Vi vil sandsynligvis se chatbots, der kan føre endnu mere naturlige og nuancerede samtaler, integreres dybere i vores dagligdag og arbejde, og måske endda få større beslutningskompetence i visse sammenhænge.

Samtidig vil det være nødvendigt at adressere udfordringerne med datasikkerhed, etik og gennemsigtighed for at sikre, at chatbots bruges ansvarligt og til gavn for både virksomheder og brugere.

Sammenfattende kan en chatbot defineres som et alsidigt værktøj, der – afhængig af teknologien bag – kan simulere menneskelig samtale, automatisere opgaver og forbedre brugeroplevelsen på tværs af brancher og platforme.

Moderne chatbots er i stand til at håndtere et væld af opgaver, såsom kundesupport, produktanbefalinger, tidsbestilling og endda rådgivning inden for sundhed og finans.

De kan integreres på hjemmesider, i apps og på sociale medier, hvilket giver virksomheder mulighed for at tilbyde 24/7 support og hurtige svar på brugerhenvendelser. Chatbots anvender ofte kunstig intelligens og maskinlæring til at forstå brugerens intentioner og tilpasse svarene, hvilket øger graden af personalisering og effektivitet.

Ved at analysere data fra interaktioner kan chatbots desuden give virksomheder værdifuld indsigt i kundeadfærd og præferencer, som kan bruges til at optimere produkter, services og markedsføring.

Samlet set bidrager chatbots til øget effektivitet, skalerbarhed og forbedret kundetilfredshed på tværs af sektorer.

ChatGPT: Fra OpenAI til et globalt fænomen

Jan Engelbrecht Pedersen

ChatGPT, udviklet af forskningslaboratoriet OpenAI, har uden tvivl været den chatbot, der for alvor har introduceret generativ AI for et bredere publikum. Lanceringen af de forskellige versioner af ChatGPT, især GPT-3 og GPT-4, har markeret et betydeligt fremskridt inden for naturlig sprogforståelse og -generering. GPT-4o, den nyeste model, udmærker sig ved sine forbedrede muligheder inden for lyd-, syns- og tekstforståelse.

Historisk baggrund: OpenAI, der blev grundlagt med det formål at udvikle og fremme "venlig" AI, har gennem årene stået bag en række banebrydende AI-modeller.

GPT-serien (Generative Pre-trained Transformer) er kendt for sin evne til at forstå kontekst og generere sammenhængende og relevant tekst baseret på det input, den modtager. ChatGPT bygger videre på denne teknologi og er specifikt finjusteret til at håndtere samtaleinteraktioner, hvilket gør den særligt velegnet til chatbots og virtuelle assistenter. Arkitekturen bag GPT-modellerne er baseret på transformer-netværk, der er særligt velegnede til at håndtere sekvensdata som tekst. Transformer-netværk anvender en mekanisme kaldet 'attention', som gør det muligt for modellen at fokusere på de mest relevante dele af inputtet, når den genererer output.
Funktioner og styrker: ChatGPT udmærker sig ved sin alsidighed og evne til at håndtere en bred vifte af opgaver. Den kan generere forskellige kreative tekstformater (digte, kode, scripts, musikstykker, e-mails, breve osv.), besvare spørgsmål på en informativ måde, oversætte sprog, skrive forskellige former for kreativt indhold og meget mere. ChatGPT kan bruges som et effektivt værktøj til brainstorming, research og til at skabe udkast til tekster, der kan viderebearbejdes.

En af dens største styrker er dens omfattende træningsdata, der gør den i stand til at forstå nuancer i sproget og generere svar, der ofte føles menneskelignende.

ChatGPT kan huske konteksten fra tidligere beskeder i en samtale, hvilket giver mulighed for mere flydende og sammenhængende interaktioner. Dette gør det muligt at føre komplekse samtaler over længere tid, hvor ChatGPT kan huske detaljer og referencer fra tidligere i samtalen.

Chatbots
Claude (Anthropic)

JAN ENGELBRECHT PEDERSEN

ChatGPT anvendes i dag inden for et utal af områder, herunder kundeservice, indholdsgenerering, uddannelse, forskning, kreativ skrivning, programmering og som en generel informationskilde og assistent. Dens evne til hurtigt at generere tekst og besvare spørgsmål har gjort den til et værdifuldt værktøj for både private brugere og virksomheder. Inden for uddannelse kan ChatGPT bruges til at generere øvelsesopgaver, give feedback på essays og besvare spørgsmål om forskellige emner. Inden for forskning kan den hjælpe med at opsummere videnskabelige artikler, generere hypoteser og analysere data. I erhvervslivet kan den bruges til at automatisere kundeservice, generere marketingmateriale og oversætte dokumenter.

Claude: Anthropic's fokus på sikkerhed og kontekstforståelse

Claude, udviklet af AI-sikkerheds- og forskningsvirksomheden Anthropic, er en anden fremtrædende chatbot, der har vundet anerkendelse for sin evne til at håndtere lange kontekster og generere velargumenterede og sammenhængende svar. Claude er designet med fokus på at være hjælpsom, harmløs og ærlig, og den er trænet til at undgå at generere skadelige eller biased output.

Den største fordel ved Claude AI er dens stærke etiske fokus og evne til at levere sikre, præcise og kontekstbevidste svar, hvilket minimerer risikoen for misinformation og upassende output

Historisk baggrund

Anthropic blev grundlagt af tidligere nøglepersoner fra OpenAI med en vision om at bygge "konstitutionel AI" – AI-systemer, der er styret af et sæt principper og retningslinjer for at sikre etisk og ansvarlig adfærd. Claude er et resultat af denne filosofi, og der er lagt stor vægt på at træne modellen til at undgå skadelige eller biased output. Anthropic har udviklet en unik træningsmetode, hvor modellen trænes til at følge et sæt etiske retningslinjer, der er nedfældet i en "konstitution".

Denne konstitution bruges til at guide modellens adfærd og sikre, at den genererer svar, der er i overensstemmelse med de etiske principper.

Funktioner og styrker

Claude udmærker sig ved sin evne til at håndtere meget lange inputtekster og bevare konteksten over lange samtaler. Dette gør den særligt velegnet til opgaver, der kræver dybdegående forståelse af komplekse dokumenter eller lange diskussioner. Claude er også kendt for at generere velstrukturerede, logiske og informative svar, ofte med en mere formel og analytisk tone end ChatGPT. Dens fokus på sikkerhed og pålidelighed gør den til et attraktivt valg for virksomheder, der håndterer følsomme oplysninger eller kræver præcis og faktuel information. Claude kan bruges til at analysere juridiske dokumenter, finansielle rapporter og andre komplekse tekster, hvor det er vigtigt at have en dybdegående forståelse af indholdet.

Anvendelsesområder

Claude bruges i stigende grad til opgaver som opsummering af lange dokumenter, besvarelse af komplekse spørgsmål baseret på omfattende tekstmateriale, generering af rapporter og analyser, og som en avanceret virtuel assistent, der kan håndtere mere nuancerede forespørgsler. Dens evne til at håndtere lange kontekster gør den også relevant inden for forskning og analyse af store datasæt. Claude kan bruges til at opsummere lange videnskabelige artikler, analysere markedsdata og generere rapporter om forskellige emner. Den kan også bruges som en virtuel assistent til at hjælpe med at planlægge møder, administrere e-mails og udføre andre administrative opgaver.

Claude AI anvendes i det offentlige især som generativ AI-assistent til at forbedre borgerkommunikation og effektivisere interne arbejdsgange. For eksempel bruges Claude 3.5 Sonnet som sprogmodel bag AI-assistenten på borger.dk, hvor den hjælper med at besvare borgernes spørgsmål hurtigt og præcist samt assisterer medarbejdere med at finde information og skrive tekster i øjenhøjde. Generativ AI som Claude bruges også i kommunale chatbots, til referatskrivning, ansøgningsvejledning og formidling af komplekse regler, hvilket forenkler borgernes møde med det offentlige og frigør tid til kerneopgaver.

DeepSeek
Den kinesiske udfordrer

DeepSeek, grundlagt i 2023 af den kinesiske hedgefond High Flyer, som er kendt for sin ekspertise inden for AI-baseret aktiehandel, har hurtigt opnået international opmærksomhed. DeepSeek er dedikeret til at udvikle banebrydende sprogmodeller og chatbots, og har hurtigt fanget verdens opmærksomhed, især efter lanceringen af deres avancerede AI-modeller, som konkurrerer med vestlige giganter som OpenAI og Google.

JAN ENGELBRECHT PEDERSEN

DeepSeek, en innovativ AI-virksomhed baseret i Kina, har udviklet banebrydende sprogmodeller og chatbots, hvilket har cementeret Kinas position som en betydelig spiller på AI-scenen. Selvom DeepSeek måske ikke har opnået samme globale genkendelighed som ChatGPT eller Claude, har deres modeller vist sig at være yderst konkurrencedygtige på flere områder, især inden for kodegenerering og specifikke sproglige opgaver på kinesisk

Historisk baggrund: DeepSeek har hurtigt etableret sig som en innovativ kraft inden for AI-forskning og -udvikling. Virksomheden har fokuseret på at bygge robuste og effektive AI-modeller, der kan håndtere både engelsk og kinesisk sprog med høj præcision. DeepSeek har investeret massivt i forskning og udvikling af AI-teknologier, og de har tiltrukket nogle af de dygtigste forskere og ingeniører inden for området.

Funktioner og styrker: DeepSeek-modellerne er særligt anerkendt for deres evne til at generere kode i forskellige programmeringssprog.
De har vist imponerende resultater i benchmarks for kodegenerering og er blevet et populært værktøj blandt udviklere.
Derudover har DeepSeek også udviklet stærke sprogmodeller, der kan håndtere komplekse sproglige opgaver på både engelsk og kinesisk, herunder oversættelse og tekstgenerering.

Deres fokus på effektivitet og ydeevne gør dem til en interessant konkurrent på det globale AI-marked. DeepSeek kan generere kode i sprog som Python, Java, C++ og JavaScript, og de kan hjælpe udviklere med at automatisere rutineopgaver og fremskynde udviklingsprocessen.

Anvendelsesområder: DeepSeek anvendes primært inden for softwareudvikling, hvor deres avancerede sprogmodeller automatiserer kodegenerering og assisterer programmører. Deres sprogmodeller finder også anvendelse inden for oversættelse, indholdsgenerering og som intelligente assistenter i kinesisktalende markeder. DeepSeek kan generere kode til web-, mobil- og desktopapplikationer. Derudover kan deres sprogmodeller oversætte teknisk dokumentation, generere brugervejledninger og skrive marketingmateriale.

Chatbots som DeepSeek spredes lynhurtigt, fordi de er gratis, brugervenlige, open source, tilbyder avancerede AI-funktioner og kan tilgås direkte via populære app-butikker, hvilket gør dem let tilgængelige for hele befolkningen

Perplexity AI
Informationssøgning i samtaleform

JAN ENGELBRECHT PEDERSEN

Perplexity AI adskiller sig fra mange andre chatbots ved sit primære fokus på informationssøgning og -præsentation i en samtaleform. I stedet for blot at generere tekst baseret på træningsdata, forsøger Perplexity AI aktivt at søge på internettet i realtid for at finde relevante svar på brugerens spørgsmål og præsenterer disse svar med kildehenvisninger.

Historisk baggrund: Perplexity AI blev grundlagt med en mission om at gøre viden mere tilgængelig og forståelig gennem en interaktiv søgeoplevelse. Ved at kombinere kraften fra store sprogmodeller med evnen til at søge og verificere information på nettet, tilbyder Perplexity AI en unik tilgang til informationssøgning. Perplexity AI bruger avancerede søgeteknologier til at finde relevante kilder på internettet, og de bruger NLP-teknikker til at opsummere og præsentere informationen på en forståelig måde.

Funktioner og styrker: Den primære styrke ved Perplexity AI er dens evne til at give svar, der er baseret på aktuelle informationer fra internettet, og at citere de kilder, hvorfra informationen er hentet. Dette øger troværdigheden og giver brugeren mulighed for selv at verificere oplysningerne. Samtalegrænsefladen gør det nemt at stille opfølgende spørgsmål og præcisere søgningen. Perplexity AI kan også opsummere information fra flere kilder og præsentere komplekse emner på en forståelig måde.

Perplexity AI kan bruges til at finde svar på spørgsmål om aktuelle begivenheder, videnskabelige emner, historiske begivenheder og meget mere.

Anvendelsesområder: Perplexity AI er særligt nyttig til research, hurtig informationssøgning, opsummering af aktuelle begivenheder og til at få svar på spørgsmål, hvor det er vigtigt at have adgang til den seneste information og de underliggende kilder. Den bruges af studerende, forskere, journalister og alle, der har behov for hurtigt og pålideligt at finde information online. Perplexity AI kan bruges til at researche til en opgave, finde baggrundsinformation om en nyhedshistorie, opsummere en videnskabelig artikel eller finde svar på et spørgsmål om et bestemt emne.

Microsoft CoPilot

Integration i produktivitetssuiten

JAN ENGELBRECHT PEDERSEN

Microsoft CoPilot repræsenterer en integration af generativ AI direkte ind i Microsofts omfattende suite af produktivitetsværktøjer, herunder Office-pakken (Word, Excel, PowerPoint, Outlook) og Windows-operativsystemet. Dette gør AI-kraften tilgængelig direkte i de applikationer, som mange mennesker bruger dagligt.

Historisk baggrund:

Microsoft har i stigende grad investeret i AI og har integreret forskellige AI-funktioner i deres produkter over tid. Lanceringen af CoPilot markerer et stort skridt i denne udvikling, hvor avancerede sprogmodeller gøres til en integreret del af brugeroplevelsen i de mest populære produktivitetsværktøjer. Microsoft har arbejdet på at integrere AI i deres produkter i mange år, og CoPilot er kulminationen på disse bestræbelser.

Funktioner og styrker:

CoPilot tilbyder en række kontekstuelle AI-funktioner direkte i de relevante applikationer.

I Word kan den hjælpe med at generere udkast, opsummere tekst og forbedre skrivestilen.

I Excel kan den analysere data og foreslå visualiseringer. I PowerPoint kan den hjælpe med at skabe præsentationer baseret på tekstinput.

I Outlook kan den hjælpe med at skrive e-mails og opsummere tråde. Integrationen i Windows giver også mulighed for generelle AI-assistancefunktioner i operativsystemet.

Styrken ligger i den direkte adgang til AI-kraft inden for de værktøjer, brugerne allerede er fortrolige med og bruger dagligt, hvilket potentielt kan øge produktiviteten markant. CoPilot kan hjælpe brugere med at spare tid på rutineopgaver, forbedre kvaliteten af deres arbejde og få mere ud af deres Microsoft-produkter.

Anvendelsesområder:

Microsoft CoPilot er designet til at forbedre produktiviteten i en bred vifte af arbejdsrelaterede opgaver.

Det kan hjælpe med at spare tid på rutineopgaver, generere ideer, forbedre kommunikationen og analysere data mere effektivt.

Målgruppen er primært professionelle brugere og virksomheder, der er integreret i Microsofts økosystem.

CoPilot kan bruges til at skrive rapporter, oprette præsentationer, analysere data, skrive e-mails og meget mere.

Microsoft CoPilot

Dybt integreret

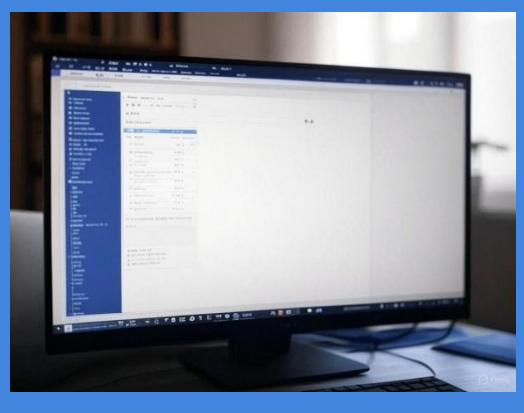
Jan Engelbrecht Pedersen

Microsoft integrerer nu Copilot bredt i deres produkter for at øge produktiviteten og lette brugeroplevelsen. I Microsoft 365 (tidligere Office 365) assisterer Copilot i Word, Excel, PowerPoint, Outlook og OneNote med tekstgenerering, dataanalyse, præsentationer og e-mails.

I Edge-browseren findes Copilot i sidepanelet, hvor den kan sammenfatte websider, svare på spørgsmål og hjælpe med research. Selv Notepad har fået Copilot-integration til tekstforslag og opsummeringer.

Copilot fungerer også som en selvstændig app på Windows, web og mobil, hvor den samler AI-assistance på tværs af hele Microsoft-økosystemet. I Word kan Copilot markant forbedre skriveprocessen ved at fungere som en intelligent AI-assistent, der hjælper med både at generere og optimere tekst. Med Copilot kan du hurtigt gå fra en tom side til et færdigt udkast, da den kan foreslå indhold, formulere afsnit og strukturere dokumenter ud fra dine stikord eller instruktioner.Den assisterer også med at omskrive tekst, justere tone og stil samt rette grammatiske fejl og stavefejl, hvilket sikrer et mere professionelt og fejlfrit resultat.

Copilot kan desuden omdanne tekst til tabeller, foreslå layout- og designskabeloner og integrere relevante data fra andre Office-programmer, hvilket gør dokumentet mere informativt og visuelt tiltalende.



Den lærer løbende af din skrivestil og præferencer, så forslagene bliver mere personlige og relevante over tid.

Grok
X's frittalende AI

JAN ENGELBRECHT PEDERSEN

Grok, udviklet af Elon Musks AI-selskab xAI, positionerer sig som en chatbot med en mere "oprørsk" personlighed og en særlig evne til at svare på spørgsmål, som andre AI-systemer måske undgår. Grok er også unik ved at have adgang til realtidsdata fra X (tidligere Twitter), hvilket giver den mulighed for at inkorporere aktuelle begivenheder i sine svar.

Historisk baggrund:

xAI blev grundlagt med en ambition om at udfordre de eksisterende AI-modeller og udvikle AI med en mere "nysgerrig" og "sandhedssøgende" tilgang.

Grok er det første produkt af denne vision. xAI har til hensigt at udvikle AI, der er mere åben og ærlig end de eksisterende modeller, og de har fokus på at skabe AI, der kan hjælpe mennesker med at forstå verden omkring dem bedre.

Grok 3 er en avanceret AI-chatbot med fokus på dybtgående viden og kontekstforståelse. Den er optimeret til komplekse diskussioner og tilbyder forbedret problemløsning ved at analysere data effektivt. Grok 3 udmærker sig ved sin tilpasselige tilgang og høj præcision.

Anvendelsesområder:

Grok appellerer til brugere, der søger en mere direkte og 'ufiltreret' AI-samtalepartner, og som ønsker adgang til realtidsinformation integreret i AI-svar.

Grok's unikke personlighed og adgang til X-data gør den særligt relevant inden for nyhedsformidling, debat og som en kilde til forskellige perspektiver på aktuelle begivenheder. Grok kan anvendes til at følge med i nyhederne, deltage aktivt i debatter og få et omfattende overblik over forskellige perspektiver på et emne.

Funktioner og styrker:

Grok adskiller sig fra mange andre chatbots ved sin mere direkte og til tider humoristiske tone. Den er designet til at kunne besvare kontroversielle spørgsmål og har adgang til en enorm mængde realtidsinformation via X.

Denne adgang til aktuelle data giver den en fordel i forhold til modeller, der primært er trænet på statiske datasæt.

Grok er også kendt for sin evne til at generere svar, der kan være mere udfordrende og mindre "politisk korrekte" end output fra andre AI-systemer.

Grok kan bruges til at få et mere nuanceret perspektiv på aktuelle begivenheder og til at udforske kontroversielle emner.

Grok 3 adskiller sig fra andre chatbots ved at have færre indholdsmæssige begrænsninger og en mere "frisindet" tilgang til spørgsmål og emner. Her er nogle nøgleområder, hvor Grok 3 har færre grænser:

Mere tilladende humor og sarkasme
Grok 3 er designet til at svare med kant og personlighed – også med sarkasme og mørk humor, hvor andre modeller ofte forsøger at neutralisere eller undgå det.

Færre begrænsninger i politiske eller kontroversielle emner
Grok 3 kan være mere åben og direkte i sine svar om følsomme emner, hvor f.eks. ChatGPT, Claude eller Gemini ofte indfører etiske filtre og neutralitet.

Realtidsdata fra X
Modsat mange chatbots der bygger på træning frem til en bestemt dato, har Grok adgang til aktuelle opslag og trends fra X, hvilket giver mere friske og kontekstuelle svar.

Mindre censur
Elon Musk har selv udtalt, at Grok er "designet til at sige det, som det er – selv hvis det er upopulært," hvilket gør den mere direkte i sine svar.

Grok har **stadig nogle sikkerhedsgrænser**, især når det gælder ulovligt eller ekstremt skadeligt indhold.

Sammenligning af funktioner og anvendelsesområder

Som det fremgår af beskrivelserne, besidder hver af disse førende chatbots unikke styrker og er optimeret til forskellige anvendelsesområder. Forskellene i deres tilgange, træningsdata og designfilosofier afspejler sig i deres respektive evner og de områder, hvor de udmærker sig.

JAN ENGELBRECHT PEDERSEN

ChatGPT: Denne chatbot skiller sig ud med sin alsidighed og kreativitet. Den er i stand til at generere forskellige tekstformater, fra digte og scripts til kode og e-mails, og kan tilpasses en bred vifte af opgaver på tværs af mange domæner. ChatGPT er særligt god til at forstå nuancer i sproget og generere svar, der føles menneskelignende, hvilket gør den velegnet til interaktive samtaler og kreative projekter. Eksempelvis kan marketingfolk bruge ChatGPT til at generere idéer til kampagner, forfattere kan bruge den til at skabe udkast til romaner, og studerende kan bruge den til at få hjælp med research og skrivning.

Claude: er kendt for sin evne til at håndtere lange kontekster og generere velargumenterede svar.

Claude er ideel til komplekse analyser og dokumenthåndtering, hvor det er nødvendigt at forstå og behandle store mængder information effektivt.

Med sit fokus på sikkerhed og pålidelighed er Claude et attraktivt valg for virksomheder, der håndterer følsomme oplysninger eller kræver præcis og faktuel information.

For eksempel kan advokater bruge Claude til at analysere juridiske dokumenter, finansielle analytikere kan generere rapporter, og forskere kan opsummere videnskabelige artikler.

DeepSeek: Denne chatbot har vist sig stærk inden for kodegenerering og håndtering af både engelsk og kinesisk sprog. DeepSeek er særligt nyttig for softwareudviklere, der ønsker at automatisere kodningsprocessen eller få hjælp med at skrive kode i forskellige programmeringssprog. Dens evne til at håndtere kinesisk sprog gør den også relevant for virksomheder, der opererer på det kinesiske marked. DeepSeek kan bruges til at generere kode til webapplikationer, mobilapplikationer og desktopapplikationer.

Perplexity AI: Denne chatbot skiller sig ud med sin evne til at søge på internettet i realtid og præsentere svar med kildehenvisninger. Perplexity AI er et værdifuldt værktøj for research og informationssøgning, da den giver brugerne adgang til aktuelle og verificerbare oplysninger. Dens samtalegrænseflade gør det nemt at stille opfølgende spørgsmål og præcisere søgningen. Perplexity AI kan bruges til at finde svar på spørgsmål om aktuelle begivenheder, videnskabelige emner og historiske begivenheder.

Grok: Denne chatbot differentierer sig med sin mere frittalende personlighed og adgang til realtidsdata fra X (tidligere Twitter). Grok appellerer potentielt til brugere, der søger en mere direkte og "ufiltreret" AI-samtalepartner, og som er interesseret i at få adgang til realtidsinformation integreret i AI-svar.

Dens unikke personlighed og adgang til X-data kan gøre den relevant inden for nyhedsformidling, debat og som en kilde til perspektiver på aktuelle begivenheder.

Grok kan bruges til at følge med i nyhederne, deltage i debatter og få et overblik over forskellige perspektiver på et emne.

Microsoft CoPilot: Denne chatbot integrerer AI-kraften direkte i de velkendte Microsoft-produktivitetsværktøjer. CoPilot har potentialet til at revolutionere måden, vi arbejder på, ved at automatisere rutineopgaver, generere ideer og forbedre kommunikationen.

Dens integration i Microsofts økosystem gør den let tilgængelig for millioner af brugere over hele verden. CoPilot kan bruges til at skrive rapporter, oprette præsentationer, analysere data og skrive e-mails.

Fremtidsperspektiver og teknologisk dybde

JAN ENGELBRECHT PEDERSEN

Denne indledende præsentation af de kendte ansigter på chatbot-arenaen danner grundlaget for vores videre udforskning. I de kommende kapitler vil vi dykke dybere ned i de teknologier, der driver disse intelligente systemer, undersøge deres opbygning og funktionsmåde, og ikke mindst se på de mange muligheder og udfordringer, som generativ AI bringer med sig.

Fra de neurale netværks komplekse arkitektur til de store sprogmodellers imponerende evne til at forstå og generere menneskelignende tekst, vil vi afdække hemmelighederne bag denne fascinerende teknologi. Vi vil også se nærmere på, hvordan disse grundlæggende byggesten samles for at skabe de chatbots, vi har introduceret i dette kapitel, og de overvejelser, der ligger bag designet af effektive og engagerende samtaleoplevelser.

Dynamisk udvikling og afsluttende bemærkninger

Det er vigtigt at understrege, at feltet inden for chatbots og generativ AI er i konstant udvikling.

Nye modeller og funktioner lanceres løbende, og konkurrencen mellem de forskellige aktører er intens.

Denne dynamik betyder, at de specifikke styrker og anvendelsesområder for de enkelte chatbots også kan ændre sig over tid. For eksempel ser vi en tendens til, at flere chatbots integrerer multimodale inputmuligheder, hvilket betyder, at de kan behandle både tekst, billeder og lyd.

Derudover arbejdes der på at forbedre chatbots evne til at håndtere komplekse resonnementer og løse problemer, hvilket vil udvide deres anvendelsesområder yderligere.

Ikke desto mindre giver den indledende oversigt i dette kapitel et solidt fundament for at forstå de grundlæggende forskelle og potentialer inden for denne spændende teknologi.

Hvad er AI ?
En bred definition

JAN ENGELBRECHT PEDERSEN

Efter at have stiftet bekendtskab med de mere synlige manifestationer af kunstig intelligens i form af chatbots, er det nu tid til at dykke dybere ned i selve fundamentet for denne revolutionerende teknologi.

Dette kapitel vil udfolde de grundlæggende principper og kernekoncepter, der udgør kunstig intelligens (AI).

Vi vil bevæge os fra en bred definition af AI til en mere detaljeret undersøgelse af de arkitekturer og metoder, der gør maskiner i stand til at udføre opgaver, som traditionelt har krævet menneskelig intelligens. Forståelsen af disse underliggende mekanismer er essentiel for at kunne vurdere potentialet og begrænsningerne ved de generative AI-modeller og chatbots, vi introducerede i det foregående kapitel.

Kunstig intelligens (AI) refererer i sin essens til en maskines eller et programs evne til at efterligne kognitive funktioner, som typisk forbindes med menneskelig intelligens. Dette inkluderer evner som læring, problemløsning, beslutningstagning, perception, sprogforståelse og kreativitet. Definitionen af AI er dog dynamisk og har udviklet sig i takt med teknologiske fremskridt. Tidligere definitioner fokuserede ofte på evnen til at udføre komplekse beregninger eller spille strategiske spil. Moderne definitioner lægger derimod større vægt på fleksibilitet, adaptivitet og evnen til at håndtere usikkerhed samt komplekse, ustrukturerede data.

AI er ikke en monolitisk disciplin, men snarere et tværfagligt felt, der trækker på principper fra datalogi, matematik, statistik, psykologi, neurovidenskab, filosofi og lingvistik. Målet er at skabe systemer, der kan ræsonnere, planlægge, lære og handle intelligent i forskellige miljøer. Traditionelt har man skelnet mellem "svag" eller "narrow" AI og "stærk" eller "general" AI.

I denne bog vil vores primære fokus være på den svage AI, der driver de generative modeller og chatbots. For at forstå disse systemers virkemåde er det afgørende at dykke ned i de underliggende arkitekturer, især neurale netværk.

Svag AI (også kendt som Narrow AI) er designet og trænet til at udføre en specifik opgave, såsom billedgenkendelse, sprogtranslation eller spil. De fleste af de AI-systemer, vi ser i dag, inklusive de avancerede generative AI-modeller, falder ind under denne kategori. De driver de generative modeller og chatbots, vi har introduceret.
De er ekstremt dygtige inden for deres afgrænsede domæne, men mangler den bredde og fleksibilitet, der kendetegner menneskelig intelligens.

Stærk AI (også kendt som Artificial General Intelligence eller AGI) refererer til en hypotetisk form for AI med intelligens på niveau med eller overlegen menneskelig intelligens på tværs af alle kognitive domæner. AGI ville være i stand til at lære, forstå og anvende viden på samme måde som et menneske, og potentielt endda overgå menneskelig kapacitet inden for mange områder. Selvom der er betydelig forskning inden for AI, er AGI endnu ikke realiseret og forbliver et langsigtet mål og genstand for intens debat og spekulation.

Træning af neurale netværk
Backpropagation

Jan Engelbrecht Pedersen

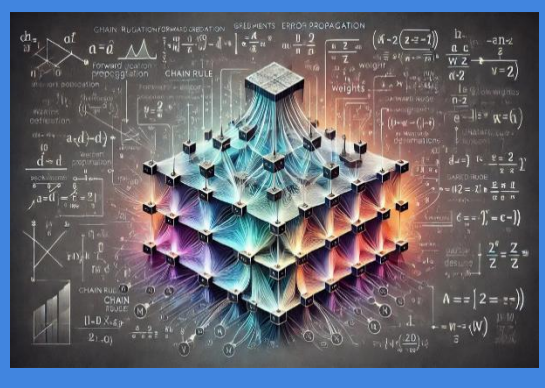
For at et neuralt netværk kan udføre sine tilsigtede opgaver effektivt, skal det trænes på en passende mængde data. Denne træning involverer justering af vægtene og bias i netværket baseret på den data, det trænes på. Denne proces kaldes træning eller læring, og den har til formål at optimere netværkets parametre (vægte og bias), så det kan producere korrekte eller nyttige resultater for nye, ukendte input. Den mest almindelige algoritme, der anvendes til at træne dybe neurale netværk, er backpropagation i kombination med en optimeringsalgoritme.

Backpropagation (fejl-tilbageføring) er en central algoritme i træningen af neurale netværk. Den bruges til at beregne, hvordan hver vægt og bias i netværket skal justeres for at minimere forskellen mellem netværkets output og det ønskede output. Processen foregår i to hovedtrin:

Forward pass (Fremadrettet gennemløb): Først præsenteres netværket for et input fra træningsdatasættet. Inputtet propagateres gennem netværket lag for lag, hvor hver neuron udfører sin vægtede sum og aktiveringsfunktion, indtil outputlaget producerer en forudsigelse.

Backward pass (Tilbageadrettet gennemløb): Outputtet sammenlignes med det korrekte output (også kaldet "ground truth" eller "label") fra træningsdatasættet. En fejlfunktion (også kaldet lossfunktion eller omkostningsfunktion) beregnes for at kvantificere forskellen mellem det forudsagte output og det faktiske output. Jo større fejlen er, desto mere skal netværket justere sine vægte. Backpropagation algoritmen bruger gradienten af fejlfunktionen med hensyn til hver vægt i netværket til at bestemme, i hvilken retning vægten skal justeres for at reducere fejlen.

Gradienten er et mål for, hvor stejlt fejlfunktionen ændrer sig med en lille ændring i vægten. Ved at følge gradienten "nedad" kan algoritmen finde et minimum i fejlfunktionen, hvor netværkets præstation er optimeret.



Backpropagation er grundlæggende inden for deep learning, en undergren af maskinlæring, der involverer dybe neurale netværk, og som har været drivkraften bag nogle af de mest bemærkelsesværdige fremskridt inden for kunstig intelligens i de senere år.

Optimeringsalgoritmer
At finde de bedste vægte

JAN ENGELBRECHT PEDERSEN

Når gradienterne er beregnet, anvendes en optimeringsalgoritme til at opdatere vægtene i netværket. Formålet med optimeringen er at finde de vægtsæt, der minimerer fejlfunktionen på træningsdataene. Der findes mange forskellige optimeringsalgoritmer, hver med sine egne styrker og svagheder. Nogle af de mest almindelige inkluderer:

Gradient Descent (Gradientnedstigning): Dette er den mest grundlæggende optimeringsalgoritme. Den justerer vægtene i netværket i den modsatte retning af gradienten med en bestemt læringsrate. Læringsraten bestemmer, hvor store skridt algoritmen tager i retning af minimummet. En for høj læringsrate kan føre til, at algoritmen overskyder minimummet, mens en for lav læringsrate kan gøre træningen langsom og ineffektiv.

Stochastic Gradient Descent (SGD): Denne algoritme er en variant af gradient descent, hvor gradienten beregnes på et tilfældigt udvalgt undersæt af træningsdataene (en "mini-batch") i stedet for hele datasættet. Dette gør træningen hurtigere og mindre hukommelseskrævende, men kan også føre til mere støj og ustabilitet.

Mini-batch Gradient Descent: Dette er en kompromis mellem gradient descent og SGD. Gradienten beregnes på en lille batch af træningsdata, hvilket giver en mere stabil opdatering af vægtene end SGD, men stadig er hurtigere end gradient descent.

Adam (Adaptive Moment Estimation): Dette er en adaptiv optimeringsalgoritme, der justerer læringsraten for hver vægt individuelt baseret på historiske gradienter. Adam er ofte en god udgangspunkt for træning af dybe neurale netværk på grund af sin robusthed og effektivitet.

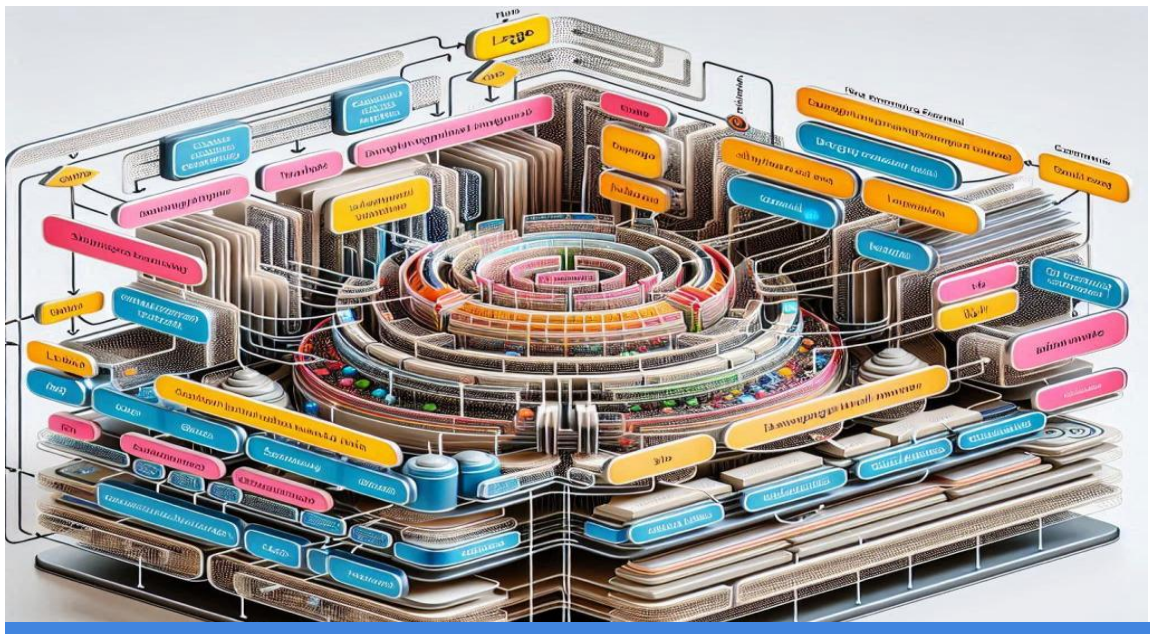
RMSprop (Root Mean Square Propagation): Dette er en anden adaptiv optimeringsalgoritme, der justerer læringsraten baseret på den gennemsnitlige kvadratiske gradient. RMSprop er effektiv til at håndtere problemer med varierende gradientstørrelser.

Adagrad (Adaptive Gradient Algorithm): Denne algoritme justerer læringsraten for hver vægt baseret på den akkumulerede sum af kvadratiske gradienter. Adagrad er velegnet til at træne netværk med sparsomme data. Disse algoritmer anvender gradienterne til iterativt at justere vægtene i netværket i retning af et minimum i fejlfunktionen. Hver algoritme har sine egne styrker og svagheder, og valget af optimeringsalgoritme afhænger af det specifikke problem og netværkets arkitektur.

Valget af træningsdata, optimeringsalgoritme, regulariseringsteknikker og overvågningsstrategier er afgørende for at opnå et neuralt netværk, der kan præstere godt på både træningsdata og nye, usete data.

Træningsprocessen er iterativ og involverer typisk mange gennemløb af træningsdataene (kaldet epoker). Under hver epoke justeres vægtene gradvist, og netværkets præstation på træningsdataene forbedres forhåbentlig. Efter hver epoke evalueres netværket på et valideringssæt af data, som ikke bruges direkte til træning. Valideringssættet bruges til at overvåge netværkets evne til at generalisere – det vil sige, dets evne til at præstere godt på nye, usete data. Hvis netværket kun præsterer godt på træningsdataene, men dårligt på valideringssættet, er det et tegn på overfitting.

Overfitting opstår, når netværket har lært træningsdataene "udenad" i stedet for at lære de generelle mønstre og relationer, der er relevante for problemet. Dette kan ske, hvis netværket er for komplekst (har for mange parametre) i forhold til mængden af træningsdata, eller hvis træningen fortsætter for længe. For at undgå overfitting kan der anvendes forskellige teknikker, såsom: **Data augmentation**: At øge mængden af træningsdata ved at skabe nye varianter af eksisterende data, **Regularisering**: At tilføje en straf til fejlfunktionen for at undgå, at vægtene bliver for store. **Dropout**: At tilfældigt deaktivere nogle af neuronerne i netværket under træningen for at tvinge de resterende neuroner til at lære mere robuste features. **Tidlig stop**: At stoppe træningen, når præstationen på valideringssættet begynder at forringes, selvom præstationen på træningssættet fortsætter med at forbedres.



Large Language Models (LLM)
Sprogets mestre

Large Language Models (LLM'er) udgør et af de mest avancerede højdepunkter inden for AI-baseret sprogforståelse og -generering. Disse modeller er meget store dybe neurale netværk, næsten altid baseret på transformer-arkitekturen, og de er trænet på enorme tekstmængder fra kilder som bøger, artikler, hjemmesider og kode. generativ AI, såsom ChatGPT, Gemini og mange andre.

JAN ENGELBRECHT PEDERSEN

Arkitekturen bag LLM'er :Transformer-arkitekturen, introduceret af Google i 2017, har fundamentalt ændret måden, hvorpå AI-modeller arbejder med tekst og sekventielle data. Hvor tidligere modeller som RNN'er (Recurrent Neural Networks) behandlede tekst sekventielt (ét ord ad gangen), gør transformerens opbygning det muligt at analysere hele tekstsekvenser parallelt og effektivt.

Kernekomponenter i Transformer-arkitekturen:

Selv-attention-mekanisme: Selv-attention gør det muligt for modellen at vurdere, hvilke ord i en tekstsekvens der er mest relevante for hinanden, uanset hvor langt de er fra hinanden i teksten. For eksempel hjælper det modellen med at forstå, at "den" i sætningen "Hunden jagtede katten, fordi den var hurtig" refererer til "katten". Dette opnås ved at beregne vægte for hvert ord i forhold til alle andre ord i sekvensen, så modellen kan fokusere på de vigtigste sammenhænge.

Multi-head attention: I stedet for kun at have én attention-mekanisme, bruger transformer-arkitekturen flere "hoveder", hvor hver head lærer forskellige relationer og mønstre i teksten. Dette øger modellens evne til at fange komplekse sproglige sammenhænge.

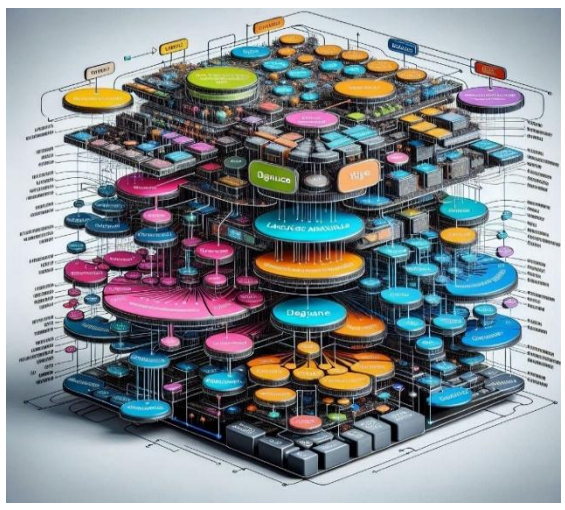
Positional encoding: Da transformer-arkitekturen ikke har en indbygget forståelse for rækkefølgen af ord (modsat RNN'er), tilføjes positionsinformation til hvert ords indlejring (embedding). Dette gør det muligt for modellen at forstå sekvensens struktur og rækkefølge.

Den oprindelige transformer består af to hoveddele:

Encoder: Modtager inputsekvensen (fx en sætning) og genererer interne repræsentationer, der fanger betydningen af inputtet.
Decoder: Bruger disse repræsentationer til at generere outputsekvensen, fx en oversættelse eller et svar.
I praksis bruger de fleste LLM'er, der skal generere tekst (som GPT-serien), kun decoder-delen, mens encoder-decoder-arkitekturen bruges til opgaver som maskinoversættelse.

Feedforward-netværk og lagdeling: Hver transformerblok indeholder også et feedforward-netværk og lag-normalisering, hvilket hjælper med at stabilisere træningen og forbedre modellens generaliseringsevne.

Residual connections: For at undgå problemer med forsvindende gradienter tilføjes inputtet til outputtet af hvert lag (residual connection), hvilket hjælper med at bevare information gennem dybe netværk.



LLM'er kan genkende, opsummere, oversætte, forudsige og generere tekst med en hidtil uset grad af sammenhæng og flydendehed. De har revolutioneret naturlig sprogbehandling (NLP) og ligger til grund for moderne

Træning af LLM'er
Data og skalering

JAN ENGELBRECHT PEDERSEN

LLM'ers succes skyldes især to forhold: adgang til enorme træningsdatasæt og skalering af modelstørrelsen.
Data: LLM'er trænes på tekstdata fra utallige kilder, såsom hjemmesider, bøger, videnskabelige artikler, kode og sociale medier. Disse datasæt kan omfatte hundreder af milliarder eller endda trillioner af ord. Jo større og mere varieret datasættet er, desto bedre kan modellen lære sproglige mønstre, semantik og kontekst.

Skalering: Forskning har vist, at ydeevnen for sprogmodeller ofte forbedres markant, når både modelstørrelsen (antallet af parametre) og mængden af træningsdata øges. Moderne LLM'er som GPT-4 og Gemini har ofte milliarder eller endda hundredvis af milliarder af parametre, hvilket kræver meget store computerressourcer og specialiseret hardware som GPU'er (Graphics Processing Units) og TPU'er (Tensor Processing Units).

Træningsmetoder:

Selvsuperviseret læring: LLM'er trænes typisk ved selvsuperviseret læring, hvor modellen lærer fra umærkede data. En udbredt metode er causal language modeling (også kaldet next-token prediction), hvor modellen får en tekstsekvens og skal forudsige det næste ord.

Ved at gentage denne proces på store datasæt lærer modellen at modellere sandsynlighedsfordelingen for tekst.

Finjustering (fine-tuning): Efter den indledende prætræning kan LLM'er finjusteres på mindre, specialiserede datasæt for at forbedre deres præstation på specifikke opgaver, fx spørgsmålsbesvarelse eller tekstgenerering i en bestemt stil.

Forståelse og generering af naturligt sprog

JAN ENGELBRECHT PEDERSEN

Gennem træning på massive datasæt og den effektive Transformer-arkitektur er LLM'er blevet bemærkelsesværdigt dygtige til både at forstå og generere naturligt sprog.

Forståelse: Selvom LLM'er ikke "forstår" sprog på samme måde som mennesker (med bevidsthed og intention), har de lært at genkende komplekse statistiske mønstre og relationer mellem ord, sætninger og hele tekster. De kan identificere semantisk lighed, syntaktisk struktur og kontekstuel betydning i en grad, der gør dem i stand til at udføre opgaver som tekstklassifikation, sentimentanalyse, informationsudtrækning og spørgsmålsbesvarelse med høj præcision.

Generering: Evnen til at generere sammenhængende, relevant og ofte kreativ tekst er en af de mest iøjnefaldende egenskaber ved LLM'er.

Ved at forudsige det næste ord i en sekvens baseret på den foregående kontekst kan de generere lange tekster, der kan ligne menneskeskrevet indhold. Kvaliteten af den genererede tekst afhænger af træningsdata, modelstørrelsen og de instruktioner (prompts), modellen modtager. Avancerede LLM'er kan generere forskellige stilarter af tekst, tilpasse sig forskellige tonefald og endda forsøge at efterligne specifikke forfatteres skrivestil.

Det er dog vigtigt at være opmærksom på, at selvom LLM'er kan generere tekst, der lyder overbevisende og informativ, mangler de en reel forståelse af verden og den sandhed, de præsenterer. Deres viden er begrænset til de data, de er trænet på, og de kan generere faktuelle fejl eller meningsløst output, især hvis de får upræcise eller vildledende prompts. Denne begrænsning er en af de vigtigste faktorer, der skal overvejes, når man bruger LLM'er i applikationer, hvor nøjagtighed er afgørende.

Maskinlæring (Supervised, Unsupervised, Reinforcement Learning): Maskinlæring (ML) er en central gren inden for kunstig intelligens, der omfatter algoritmer og statistiske modeller, som gør det muligt for computere at lære fra data uden at være eksplicit programmeret til hver enkelt opgave. I stedet for at følge faste regler lærer maskinen at genkende mønstre, træffe beslutninger og lave forudsigelser baseret på erfaring fra data. Maskinlæring anvendes i alt fra spamfiltrering og ansigtsgenkendelse til anbefalingssystemer og automatiseret diagnostik.

Dyb Læring (Deep Learning): Dyb læring er en specialiseret underkategori af maskinlæring, der anvender dybe neurale netværk – altså neurale netværk med mange lag – til at lære komplekse repræsentationer af data. Dyb læring har vist sig særligt effektiv til opgaver som billedgenkendelse, talegenkendelse og naturlig sprogbehandling.
De fleste moderne store sprogmodeller (LLM'er) er baseret på dyb læring, hvor transformer-arkitekturen er en særlig succesfuld variant.

Der findes tre hovedtyper af maskinlæring

Jan Engelbrecht Pedersen

Supervised Learning (Overvåget læring): Modellen trænes på et labelled datasæt, hvor input er parret med korrekte output (labels). Målet er at generalisere og forudsige korrekte output for nye data. Eksempler inkluderer klassifikation (fx spamfilter) og regression (fx boligpriser).

Unsupervised Learning (Uovervåget læring): Modellen trænes på et ulabelled datasæt og skal selv finde mønstre eller grupperinger. Eksempler inkluderer klyngeanalyse (fx kundesegmentering) og dimensionsreduktion (fx PCA).

Reinforcement Learning (Forstærkende læring): En agent lærer at træffe beslutninger for at maksimere en belønningsfunktion. Anvendelser inkluderer spil (fx AlphaGo), robotstyring og ressourceoptimering.

Faglige AI-termer forklaret:

Maskinlæring (ML): Algoritmer, der lærer fra data uden eksplicit programmering.

Labelled datasæt: Datasæt hvor hvert input har et tilknyttet korrekt output (label)

Unlabelled datasæt: Datasæt uden korrekte output, hvor mønstre skal opdages automatisk.

Agent: En beslutningstager i reinforcement learning, der interagerer med et miljø.

Belønningsfunktion: En funktion, der måler, hvor godt agentens handlinger opfylder målet.

Naturlig sprogbehandling (Natural Language Processing, NLP) er et tværfagligt område, der kombinerer datalogi, lingvistik og AI for at gøre det muligt for computere at forstå, analysere, generere og interagere med menneskeligt sprog.

Naturlig Sprogbehandling (NLP)

JAN ENGELBRECHT PEDERSEN

Store sprogmodeller (LLM'er) har revolutioneret NLP ved at levere hidtil uset præcision og fleksibilitet i mange af disse opgaver, og de udgør kernen i moderne chatbots og digitale assistenter.

Der er flere vigtige teknikker inden for naturlig sprogbehandling (NLP), som bruges til at analysere og forstå menneskeligt sprog. Her er nogle af de mest centrale teknikker:

Tokenisering: Opdeling af tekst i mindre enheder, såsom ord eller sætninger.

Stemming og Lemmatization: Reduktion af ord til deres grundform for at lette analyse.

Part-of-Speech Tagging: Identifikation af ordklasser (f.eks. substantiver, verber) i en tekst.

Named Entity Recognition (NER): Identifikation af specifikke enheder som personer, steder og organisationer.

Sentimentanalyse: Analyse af følelser og holdninger i tekst: Formålet er at bestemme, om en given tekst udtrykker en positiv, negativ eller neutral holdning

Naturlig sprogbehandling (NLP) er et område inden for datalogi og kunstig intelligens, der fokuserer på at give computere evnen til at forstå, fortolke og generere menneskeligt sprog. Det bruges i mange applikationer som chatbots, stemmestyrede assistenter og automatiserede oversættelsessystemer.

Syntaktisk og Semantisk Analyse: En proces, hvor man både analyserer grammatiske strukturer i en sætning og vurderer dens betydning i en given kontekst for at opnå dybere forståelse.

Maskinoversættelse: En teknologi, der automatisk oversætter tekst fra et sprog til et andet ved hjælp af avancerede algoritmer og sproganalyse, hvilket muliggør hurtig og effektiv kommunikation mellem forskellige kulturer

Talegenkendelse: Konvertering af tale til tekst.

Stemming og lemmatization er begge teknikker inden for naturlig sprogbehandling (NLP), der bruges til at reducere ord til deres grundform, men de gør det på forskellige måder.

Stemning: Fokuserer på at fjerne endelser (suffixer) fra ord for at opnå en basisform.

Lemmatization: Bruger en ordbog og morfologisk analyse for at finde den korrekte grundform af et ord.

NLP bruges i mange dagligdags applikationer som søgemaskiner, chatbots, stemmestyrede GPS-systemer og digitale assistenter som Siri og Alexa.

Det hjælper også med at automatisere opgaver, forbedre dataanalyse og generere indhold.

Computer Vision
Generativ AI

Computer vision er et område inden for kunstig intelligens, der fokuserer på at gøre det muligt for computere at "se", analysere og forstå billeder og videoer på samme måde som mennesker. Målet er at automatisere fortolkning af visuel information.

JAN ENGELBRECHT PEDERSEN

Generativ AI inden for ComputerVision gør det muligt at skabe og analysere billeder samt videoer. Teknologien bruges til at generere realistiske visualiseringer, forbedre billedkvalitet og muliggøre avancerede anvendelser som objektdetektion, ansigtsgenkendelse og dataindsigt på visuelle områder.

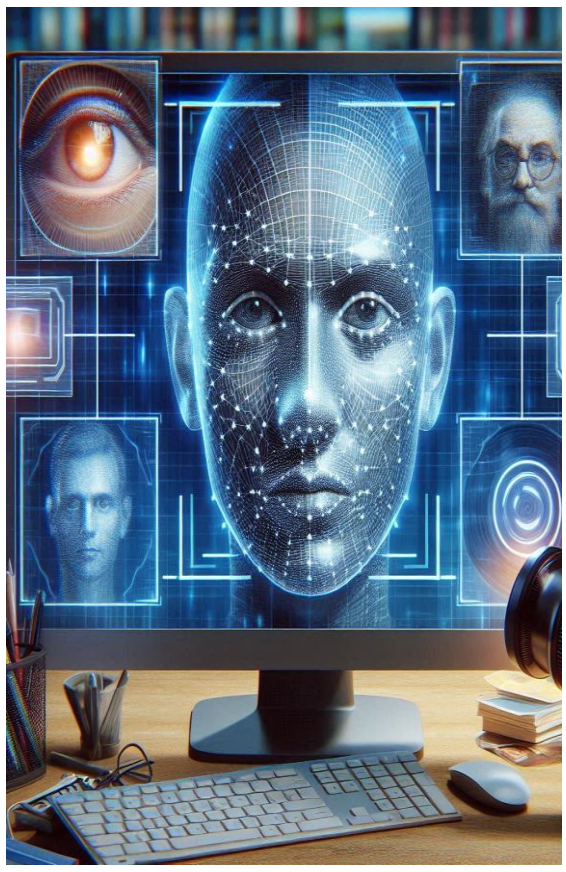
Typiske opgaver inden for computer vision inkluderer:
Objektgenkendelse: Identifikation og klassifikation af objekter i billeder eller videoer.
Billedsegmentering: Opdeling af et billede i meningsfulde segmenter eller regioner.
Ansigtsgenkendelse: Identifikation af personer baseret på ansigtstræk.
Billedgenerering: Skabelse af nye billeder baseret på beskrivelser eller eksisterende billeder (f.eks. GAN-modeller).

Selvom computer vision ikke direkte indgår i LLM-baserede chatbots, kan det integreres for at skabe multimodale systemer, der kan forstå både tekst og billeder.

Objektgenkendelse identificerer specifikke elementer i billeder, mens billedsegmentering opdeler dem i meningsfulde sektioner. Ansigtsgenkendelse analyserer og verificerer menneskelige ansigter, og billedgenerering skaber realistiske eller kreative visualiseringer. Disse teknologier har revolutioneret chatbots som Copilot, ChatGPT 4 og Grok 3, der integrerer visuelle funktioner for bedre interaktion og kontekstforståelse. Chatbots med ComputerVision kan for eksempel analysere billeder for at assistere brugere, styrke kreativitet og levere avanceret support, hvilket gør dem mere intuitive og funktionelle.

Ansigtsgenkendelse:

Ansigtsgenkendelse i overvågningskameraer anvendes til at identificere og verificere personer, ofte som en del af sikkerheds- og adgangskontrolsystemer. Kameraerne registrerer ansigtstræk og sammenligner dem med databaser for at genkende kendte individer eller flagge mistænkelige personer. Denne teknologi bruges ofte i offentlige rum såsom lufthavne, togstationer og virksomheder for at forhindre kriminalitet, forbedre sikkerheden og sikre adgangsbegrænsninger. Det kan også bidrage til hurtig reaktion i nødsituationer, f.eks. ved at hjælpe myndighederne med at finde forsvundne personer. Ansigtsgenkendelse er kraftfuldt, men det rejser også etiske spørgsmål omkring privatliv og misbrug.



Ansigtsgenkendelsesteknologi i Kina har vakt bekymring om privatliv og menneskerettigheder. Det bruges i vid udstrækning til overvågning, især via CCTV-netværk og offentlige rum, for at spore borgere og begrænse deres bevægelse. Nogle kritikerne peger på dens rolle i masseovervågning af specifikke befolkningsgrupper, såsom minoriteter i Xinjiang-regionen. Teknologien kan kobles til sociale kreditsystemer, hvilket påvirker borgernes adgang til samfundsmæssige ydelser.

Anbefalingssystemer
Personlige forslag drevet af kunstig intelligens

JAN ENGELBRECHT PEDERSEN

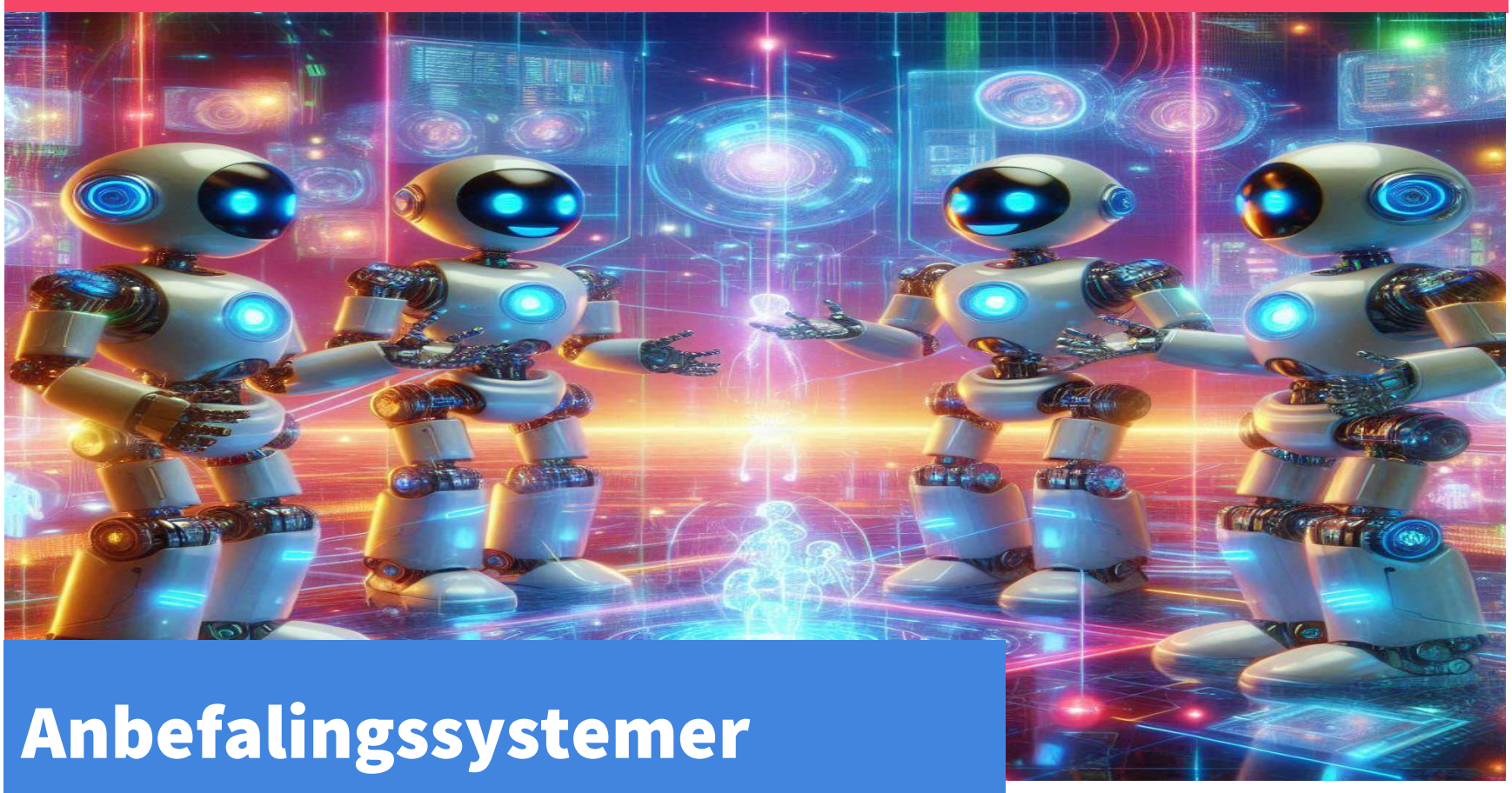
Anbefalingssystemer er avancerede AI-drevne teknologier, der har til formål at forudsige og præsentere produkter, tjenester eller informationer, som en bruger sandsynligvis vil finde relevante og interessante. Disse systemer er centrale i mange digitale platforme og hjælper med at navigere i den enorme mængde tilgængeligt indhold ved at skræddersy oplevelsen til den enkelte bruger. Anbefalingssystemer analyserer typisk store mængder data om brugernes tidligere adfærd, såsom købshistorik, søgninger, klik, visninger og ratings.

Derudover kan de også inddrage data om produkternes eller indholdets egenskaber, såsom genre, pris, popularitet eller tekniske specifikationer. Ved hjælp af disse data forsøger systemet at identificere mønstre og præferencer, som kan bruges til at generere personlige anbefalinger.
Der findes flere forskellige metoder og teknikker til at bygge anbefalingssystemer, men de mest udbredte inkluderer:

Collaborative Filtering (Samarbejdsfiltrering): Denne metode baserer sig på brugernes adfærd og præferencer. Hvis to brugere har lignende smag (f.eks. har købt eller bedømt lignende produkter højt), vil systemet anbefale produkter, som den ene bruger har interageret med, til den anden. Collaborative filtering kan være user-based (brugere med lignende præferencer) eller item-based (produkter, der ofte vælges sammen).

Content-Based Filtering (Indholdsbaseret filtrering): Her fokuserer systemet på egenskaber ved produkterne eller indholdet. Anbefalinger genereres ved at matche brugerens tidligere interaktioner med produkter, der har lignende karakteristika. For eksempel, hvis en bruger har set mange actionfilm, vil systemet anbefale andre actionfilm baseret på genre, skuespillere eller instruktør.

Hybridmetoder: Kombinationer af collaborative og content-based filtering anvendes ofte for at forbedre præcisionen og håndtere begrænsninger ved hver metode. Hybridmodeller kan også inkludere yderligere data som brugernes demografi eller kontekstuelle oplysninger (tidspunkt, sted osv.).



Anbefalingssystemer
Anvendelser

JAN ENGELBRECHT PEDERSEN

Anbefalingssystemer er udbredte på mange digitale platforme, hvor de forbedrer brugeroplevelsen og øger engagementet:

E-handel: Platforme som Amazon bruger anbefalingssystemer til at foreslå produkter baseret på tidligere køb, søgninger og brugernes adfærdsmønstre.

Streamingtjenester: Netflix og Spotify anbefaler film, serier og musik baseret på brugerens tidligere afspilninger, ratings og præferencer.

Sociale medier: Facebook, Instagram og TikTok bruger anbefalingsalgoritmer til at vise brugerne indhold, der matcher deres interesser, hvilket øger tiden brugt på platformen.

Nyhedsformidling: Medieplatforme anbefaler artikler baseret på læsevaner og interesser.

Selvom anbefalingssystemer kan forbedre brugeroplevelsen, rejser de også en række udfordringer:

Filterbobler og ekkokamre: Anbefalingsalgoritmer kan skabe situationer, hvor brugere kun eksponeres for indhold, der bekræfter deres eksisterende holdninger og præferencer, hvilket kan begrænse mangfoldighed og føre til polarisering.

Privatliv: Anbefalingssystemer kræver ofte omfattende dataindsamling om brugernes adfærd, hvilket rejser spørgsmål om databeskyttelse og brugernes ret til privatliv.

Manipulation: Algoritmer kan blive designet til at fremme bestemte produkter eller holdninger af kommercielle eller politiske interesser, hvilket kan påvirke brugernes valg og opfattelser.

Fremtidige muligheder: Integration med Generativ AI

Selvom traditionelle anbefalingssystemer ikke altid er direkte forbundet med generativ AI, åbner den hurtige udvikling inden for generative modeller nye muligheder.

Ved at kombinere anbefalingssystemers evne til at forstå brugerpræferencer med generativ AI's kapacitet til at skabe nyt og kreativt indhold kan fremtidens systemer levere endnu mere personaliserede og innovative anbefalinger.

Eksempelvis kan generativ AI skabe skræddersyet produktbeskrivelser, personlige playlister eller unikke visuelle anbefalinger baseret på brugerens smag, hvilket kan forbedre engagement og tilfredshed.

Forklaring af centrale AI-termer i anbefalingssystemer

Jan Engelbrecht Pedersen

Anbefalingssystem (Recommender System): En AI-baseret algoritme, der analyserer data om brugere og produkter for at forudsige og foreslå relevante varer eller indhold til brugeren.

Collaborative Filtering: En metode, der anbefaler produkter baseret på lignende brugeres adfærd og præferencer.

Content-Based Filtering: En metode, der anbefaler produkter baseret på ligheder i produkternes egenskaber i forhold til tidligere interaktioner.

Hybridmodel: En kombination af flere anbefalingsmetoder for at forbedre nøjagtighed og robusthed.

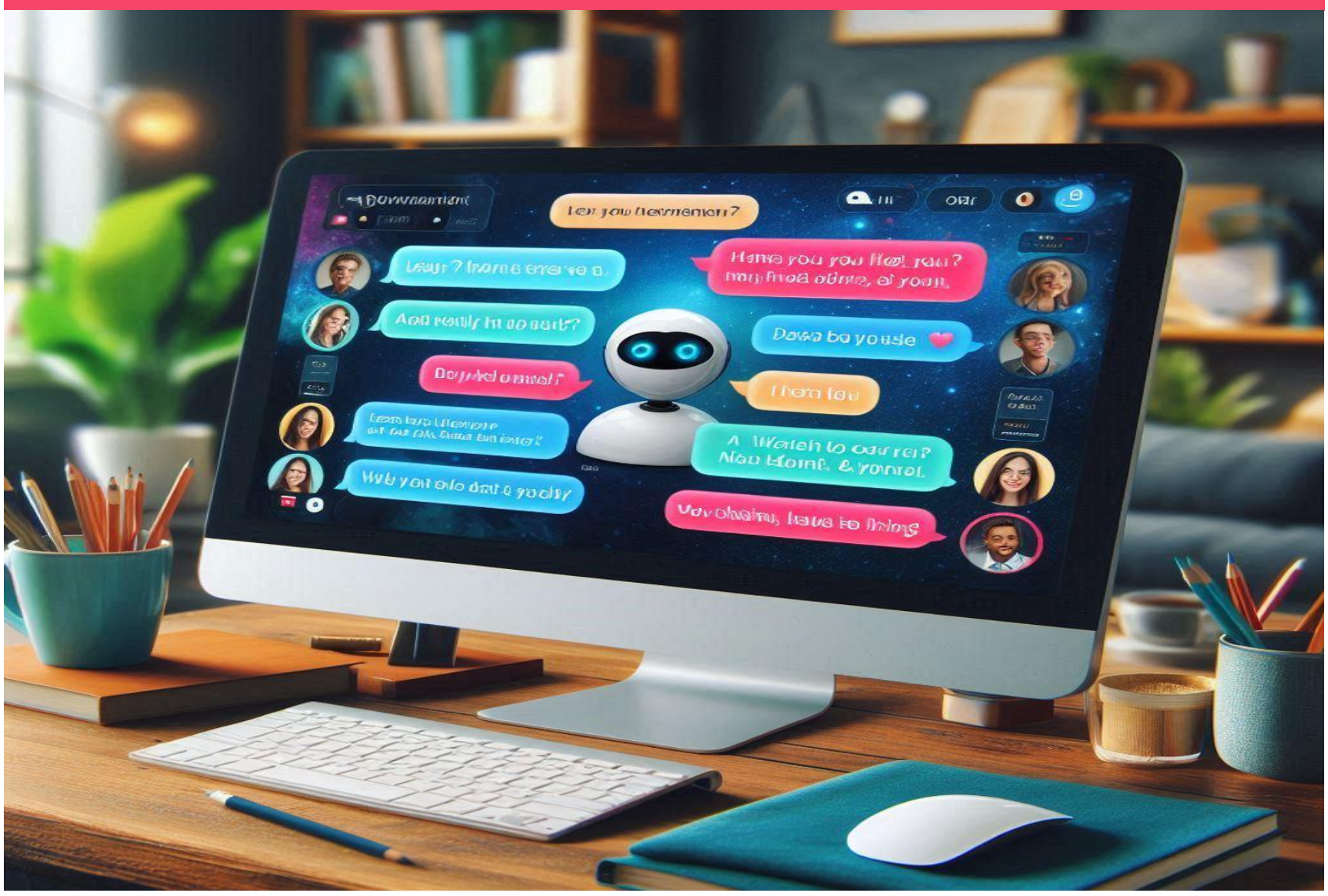
Filterboble: En situation, hvor en bruger kun præsenteres for information, der bekræfter deres eksisterende synspunkter, hvilket kan begrænse eksponering for forskellige perspektiver.

Ekkokammer: Et miljø, hvor information og holdninger forstærkes gennem gentagen eksponering inden for en lukket gruppe.

Personalisering: Tilpasning af indhold eller anbefalinger baseret på individuelle brugerdata og præferencer.

Indholdsbaserede anbefalingssystemer: Fokuserer på karakteristika ved elementer, som brugeren kan lide.

Forståelsen af disse grundlæggende AI-koncepter og teknologier er afgørende for at kunne værdsætte den kompleksitet og de muligheder, der ligger i generativ AI og de chatbots, vi har introduceret.



Fra LLM til Chatbot
Integrering af komponenter

JAN ENGELBRECHT PEDERSEN

Efter at have etableret en grundlæggende forståelse for kunstig intelligens, neurale netværk og Large Language Models (LLM'er) i det foregående kapitel, vil vi nu rette fokus mod selve konstruktionen af de chatbots, der har vundet så stor udbredelse.

Dette kapitel vil afdække de forskellige komponenter og processer, der er involveret i at transformere en generel sprogmodel som GPT-4 eller Claude til en interaktiv samtalepartner som ChatGPT eller deres slægtninge.

Vi vil undersøge, hvordan LLM'er integreres med andre teknologier, hvordan kunsten af prompt engineering spiller en afgørende rolle, og hvordan chatbots finjusteres og trænes til specifikke formål.

Derudover vil vi se på udfordringerne ved at implementere hukommelse og kontekstforståelse i samtaler samt de vigtige aspekter af sikkerhed og bias i udviklingen af disse AI-drevne dialogsystemer. Afslutningsvis vil vi illustrere typiske arkitekturer og arbejdsflows for moderne chatbots.

Under motorhjelmen i ChatGPT ligger avancerede neurale netværk, som bruger transformerarkitekturen. Den analyserer tekst, identificerer mønstre, forstår kontekst og genererer relevante svar gennem maskinlæring og optimerede algoritmer.

En Large Language Model i sig selv er en kraftfuld motor for sprogforståelse og -generering, men for at fungere som en interaktiv chatbot kræver den yderligere komponenter og processer. Overgangen fra en grundlæggende LLM til en fuldt funktionel chatbot involverer typisk integrationen af flere nøgleelementer:

Brugergrænseflade (User Interface - UI): Dette er det lag, som brugeren interagerer direkte med. Det kan være en tekstbaseret chatboks, en stemmebaseret grænseflade eller endda en kombination af begge. UI'en håndterer input fra brugeren og præsenterer chatbotens output.

API (Application Programming Interface): API'et fungerer som en bro mellem brugergrænsefladen og selve LLM'en. Det modtager brugerens input fra UI'en, formaterer det korrekt og sender det til LLM'en. Når LLM'en har genereret et svar, modtager API'et dette output og sender det tilbage til UI'en for at blive vist for brugeren.

Prompt Management System: Dette system håndterer konstruktionen og manipuleringen af de prompts (instruktioner eller spørgsmål), der sendes til LLM'en. Det kan involvere tilføjelse af kontekst fra tidligere samtaler, systeminstruktioner om chatbotens personlighed eller adfærd, og dynamisk generering af prompts baseret på brugerens input.

Dialog Management: Denne komponent er ansvarlig for at styre flowet af samtalen. Den holder styr på samtalens historie, brugerens intentioner og chatbotens tidligere svar. Dialog management sikrer en mere sammenhængende og meningsfuld interaktion ved at opretholde kontekst og styre, hvordan chatbotten reagerer på forskellige input.

Post-Processing og Filtrering: Efter at LLM'en har genereret et svar, kan et post-processing lag anvendes til at forbedre outputtet. Dette kan inkludere filtrering af upassende eller skadeligt indhold, omformulering af svar for klarhed eller stil, og tilføjelse af yderligere information eller formatering.

Data Storage og Hukommelse (til persistering): For at kunne huske information på tværs af flere samtaler eller personalisere oplevelsen for individuelle brugere, kan chatbots integreres med databaser til at gemme brugerprofiler, præferencer og samtalehistorik.

Integrationen af disse komponenter kræver omhyggelig design og udvikling for at sikre en smidig og effektiv brugeroplevelse.

Selvom LLM'en udgør hjernen i chatbotten, er de omkringliggende systemer afgørende for at gøre denne intelligens tilgængelig og anvendelig i en samtaleform.

Prompt Engineering
Kunsten at kommunikere med AI

Prompt engineering er en afgørende disciplin i udviklingen og brugen af avancerede chatbots baseret på LLM'er. En prompt er den inputtekst, som gives til LLM'en for at instruere den i at udføre en specifik opgave eller generere et bestemt output. Kvaliteten og udformningen af prompten har en enorm indflydelse på svarets relevans, nøjagtighed og stil. Prompt engineering handler derfor om at designe effektive prompts, der får LLM'en til at agere og generere output i overensstemmelse med de ønskede mål.

JAN ENGELBRECHT PEDERSEN

Kunsten i prompt engineering ligger i at formulere instruktioner, spørgsmål og kontekst på en måde, som LLM'en kan forstå og reagere hensigtsmæssigt på. Dette kan involvere at være præcis i sine anmodninger, give tilstrækkelig kontekst, definere den ønskede outputformat, specificere den ønskede tone eller stil, og endda give eksempler på ønsket adfærd (kendt som "few-shot prompting").

Effektive prompts kan inkludere elementer som:

Instruktioner: Klare og præcise anvisninger om, hvad LLM'en skal gøre (f.eks. "Skriv et resumé af følgende tekst:", "Oversæt denne sætning til fransk:", "Generer tre ideer til en marketingkampagne for et nyt produkt:").
Kontekst: Baggrundsinformation, der hjælper LLM'en med at forstå opgaven bedre (f.eks. "Du er en marketingekspert, der arbejder for en nystartet virksomhed, der sælger bæredygtige kaffebønner.").
Inputdata: Den tekst eller de informationer, som LLM'en skal arbejde med (f.eks. den tekst, der skal opsummeres eller oversættes).
Outputformat: Specifikationer for, hvordan outputtet skal se ud (f.eks. "Skriv svaret i punktopstilling:", "Generer svaret som en tabel:", "Hold svaret under 100 ord.").
Tone og Stil: Anvisninger om den ønskede tone eller stil i svaret (f.eks. "Skriv i en professionel tone:", "Svar med en humoristisk vinkel:", "Adopter en akademisk skrivestil.").

Begrænsninger: Angivelse af, hvad LLM'en ikke må gøre eller inkludere i sit svar (f.eks. "Undgå at bruge jargon:", "Nævn ikke specifikke firmanavne:", "Hold dig til fakta og undgå spekulationer.").

Avancerede prompt engineering teknikker kan involvere mere komplekse strategier, såsom chain-of-thought prompting, hvor man beder LLM'en om at forklare sin ræsonnementstrin for trin, hvilket ofte fører til mere nøjagtige svar på komplekse spørgsmål. En anden teknik er role-playing, hvor man beder LLM'en om at agere som en bestemt person eller ekspert inden for et givent område for at få mere specialiserede og kontekstuelt relevante svar.

Effektiv prompt engineering er en iterativ proces, der ofte kræver eksperimenteren og finjustering for at opnå de bedste resultater fra LLM'en. Forståelsen af, hvordan forskellige formuleringer og elementer i en prompt påvirker LLM'ens output, er en nøglefærdighed for både udviklere og brugere af avancerede chatbots.

Prompt Engineering: Disciplinen der omhandler design og formulering af effektive prompts til LLM'er for at opnå de ønskede resultater.

Few-shot prompting: En teknik, hvor man giver LLM'en nogle få eksempler på den ønskede adfærd eller output i prompten for at hjælpe den med at generere mere relevante svar.

Chain-of-thought prompting: En teknik, hvor man beder LLM'en om at forklare sin ræsonnementstrin for trin, hvilket ofte fører til mere nøjagtige svar på komplekse spørgsmål.

Role-playing: En teknik, hvor man beder LLM'en om at agere som en bestemt person eller ekspert inden for et givent område for at få mere specialiserede og kontekstuelt relevante svar.

Prompt Engineering
Kunsten at styre AI

JAN ENGELBRECHT PEDERSEN

Prompt engineering handler om at skabe specifikke og effektive input, der guider AI-systemer som ChatGPT til at levere ønskede svar. Her er nogle eksempler på prompt engineering:

Kontekstskabelse: "Du er en ekspert i psykologi. Forklar mekanismerne bag gruppepres."

Struktur: "Lav en liste med fem trin til problemløsning, organiseret med stikord."

Roller: "Forestil dig, at du er en historielærer. Fortæl om renæssancen på en enkel måde."

Format: "Skriv en e-mail til en kollega med forslag til nye initiativer."

Kreativitet: "Skab en kort fortælling om en robot, der lærer at føle."

Forklaringsniveau: "Forklar kvantemekanik for en 12-årig, på en simpel måde."

Problemløsning: "Hvad er de tre største årsager til forsinkede projekter, og hvordan kan de løses?"

Analyser: "Udled fordelene og ulemperne ved fjernarbejde, præsenteret som en tabel."

Sammenligning: "Sammenlign Tesla og SpaceX med fokus på deres teknologi og innovationsniveau."

Opsummering: "Lav en kortfattet opsummering af hovedpunkterne i en artikel om bæredygtighed."

Finjustering og træning af chatbots til specifikke formål

JAN ENGELBRECHT PEDERSEN

Selvom store sprogmodeller som GPT-4 og Claude er trænet på enorme mængder generel tekstdata og besidder en bred vifte af sproglige færdigheder, er de ofte ikke optimalt egnet til specifikke anvendelsesområder eller domæner uden yderligere træning eller tilpasning. Finjustering (fine-tuning) er en proces, hvor en prætrænet LLM trænes yderligere på et mindre, mere specifikt datasæt, der er relevant for den pågældende opgave eller det pågældende domæne.

Formålet med finjustering er at tilpasse LLM'ens eksisterende viden og sproglige færdigheder til de særlige krav og karakteristika ved den ønskede applikation. For eksempel kan en generel LLM finjusteres på et datasæt af kundeservice-samtaler for at forbedre dens evne til at håndtere kundeforespørgsler effektivt og i en bestemt virksomheds tone. Ligeledes kan en LLM finjusteres på medicinske tekster for at forbedre dens forståelse og generering af medicinsk relateret indhold.

Finjustering involverer typisk at træne LLM'en på et labelled datasæt, hvor input (f.eks. en kundeforespørgsel) er parret med det ønskede output (f.eks. et relevant svar). Under finjusteringen justeres vægtene i LLM'en yderligere for at optimere dens præstation på dette specifikke datasæt. Da LLM'en allerede har lært generelle sproglige mønstre under prætræningen, kræver finjustering typisk mindre data og færre computerressourcer end træning fra bunden.

Ud over finjustering kan andre træningsteknikker anvendes til at forbedre chatbotens adfærd og ydeevne. Reinforcement Learning from Human Feedback (RLHF) er en avanceret teknik, der har spillet en vigtig rolle i udviklingen af chatbots som ChatGPT. I RLHF trænes en belønningsmodel baseret på menneskelig feedback om kvaliteten og relevansen af chatbotens svar. Denne belønningsmodel bruges derefter til at finjustere LLM'en ved hjælp af reinforcement learning algoritmer, så den genererer svar, der er mere i overensstemmelse med menneskelige præferencer for hjælpsomhed, sandfærdighed og harmløshed.

Valget af træningsdata, finjusteringsstrategi og eventuel brug af teknikker som RLHF er afgørende for at skabe en chatbot, der er effektiv, pålidelig og i overensstemmelse med de ønskede mål og etiske retningslinjer.

Finjustering (Fine-tuning): En proces, hvor en prætrænet LLM trænes yderligere på et mindre, mere specifikt datasæt for at tilpasse den til en bestemt opgave eller domæne.

Reinforcement Learning from Human Feedback (RLHF): En træningsteknik, hvor en belønningsmodel trænes baseret på menneskelig feedback om kvaliteten og relevansen af chatbotens svar, og denne model bruges til at finjustere LLM'en ved hjælp af reinforcement learning algoritmer.

Finjustering og træning
Af Chatbots

Jan Engelbrecht Pedersen

Finjustering og træning af chatbots til specifikke formål indebærer tilpasning af en generel AI-model til en specifik opgave eller branche. Først indsamles relevante data, såsom kundesamtaler, domænespecifikke tekster eller tekniske dokumenter. Disse data bruges til at træne chatbotten, så den forstår det ønskede sprog og kontekst.

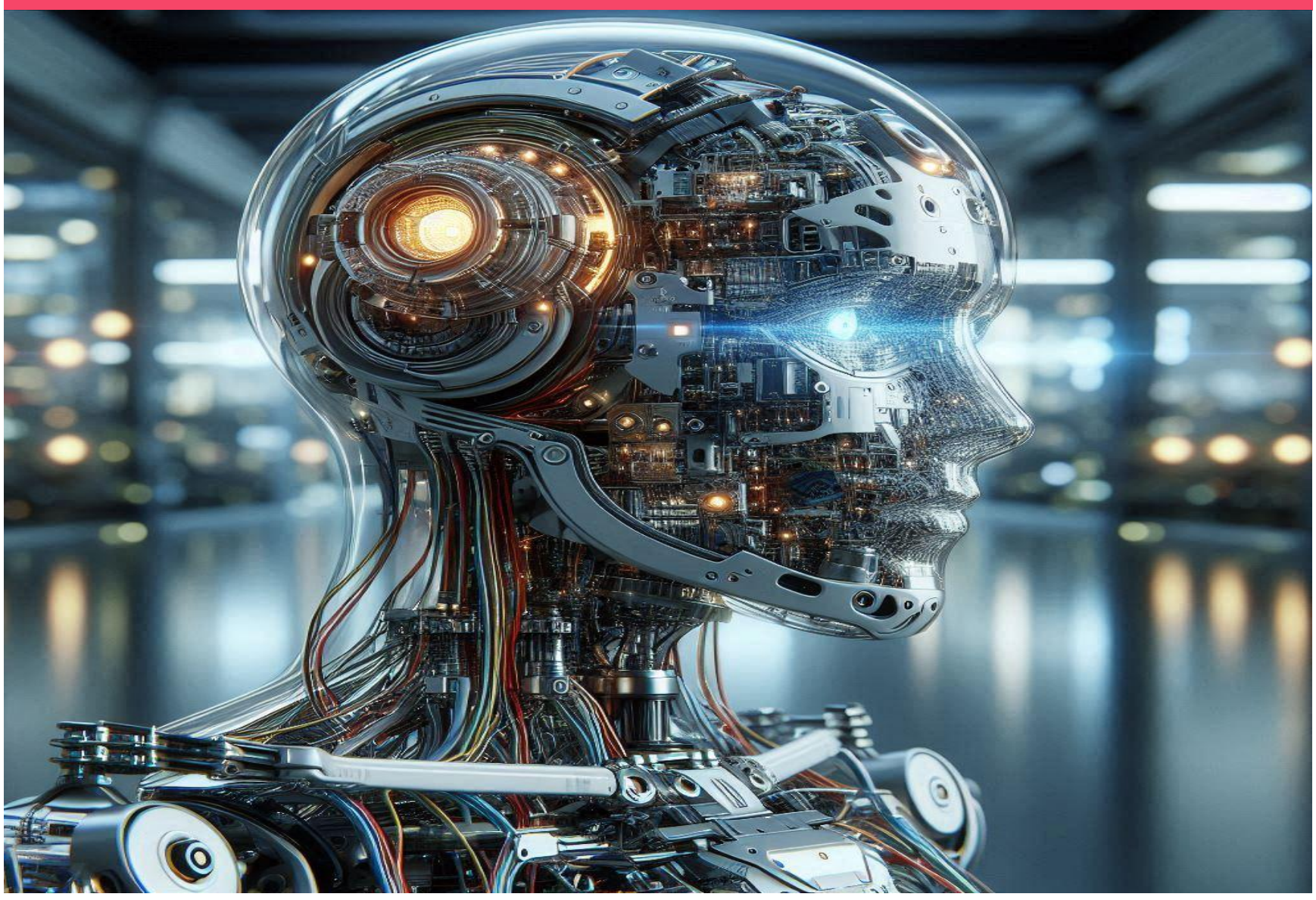
Finjusteringen justerer de eksisterende AI-modeller gennem maskinlæring, hvor små ændringer forbedrer nøjagtighed og præcision. Modellen testes derefter med realistiske scenarier for at sikre, at den fungerer effektivt.

Træning kan også inkludere tilpasning af tonen, stilen og formen på chatbotten, så den matcher målgruppen. Resultatet er en chatbot, der er skræddersyet til specifikke anvendelser såsom kundeservice, teknisk support eller undervisning, med højt niveau af effektivitet og brugertilfredshed.

Selve træningen udføres ofte ved hjælp af superviseret eller usuperviseret maskinlæring. Superviseret læring involverer brug af labeldata, hvor modellen trænes til at forstå input og forventet output. Usuperviseret læring anvender ulabelerede data for at finde mønstre og relationer

Under processen finjusteres modellen med parametre og algoritmer, der sikrer optimal ydeevne. Testning med realistiske scenarier hjælper med at identificere svagheder, og yderligere iterationer kan forbedre dens funktion.

Til sidst anvendes metoder som reinforcement learning for at sikre, at chatbotten lærer kontinuerligt fra brugerfeedback og nye data. Dette resulterer i en robust og skræddersyet chatbot, der leverer effektivt inden for sit domæne



Hukommelse

Kontekstforståelse i samtaler

JAN ENGELBRECHT PEDERSEN

En af de centrale udfordringer i udviklingen af avancerede chatbots er at give dem evnen til at huske og forstå konteksten af en samtale over tid. Menneskelige samtaler er typisk sammenhængende, hvor tidligere udsagn påvirker fortolkningen og relevansen af efterfølgende udsagn. For at en chatbot skal kunne føre naturlige og meningsfulde samtaler, er det derfor afgørende, at den kan opretholde en form for "hukommelse" af samtalens forløb.

De fleste moderne chatbots implementerer en form for samtalekontekst ved at inkludere de tidligere udvekslinger i den prompt, der sendes til LLM'en. Når brugeren stiller et nyt spørgsmål, inkluderes historikken af de tidligere spørgsmål og svar i inputtet til LLM'en. Dette giver LLM'en mulighed for at referere tilbage til tidligere emner og give svar, der er relevante i den aktuelle kontekst. Længden af denne kontekst er dog ofte begrænset af LLM'ens inputvindue (den maksimale mængde tekst, den kan behandle ad gangen).

Udvidet hukommelse i chatbots gør det muligt at gemme tidligere interaktioner, brugerdetaljer og præferencer, hvilket skaber mere personaliserede, kontekstrige og vedvarende samtaler, der forbedrer brugeroplevelsen over tid og interaktion.

For at håndtere længere samtaler og mere kompleks kontekst kan der implementeres mere avancerede hukommelsesmekanismer. Disse kan inkludere:

Samtalesammendrag (Conversation Summarization): Chatbotten kan periodisk opsummere de vigtigste punkter fra den hidtidige samtale og inkludere dette resumé i den efterfølgende prompt for at bevare konteksten uden at overskride inputvinduet.

Det kan være nyttigt til at sikre, at en chatbot forstår essensen af tidligere interaktioner, uden at skulle holde hele samtalen i hukommelsen. Dette hjælper med effektivitet, især i scenarier, hvor brugeren vender tilbage efter længere tid.

Ekstern Hukommelse

Ekstern hukommelse er en mekanisme, hvor chatbots kan gemme data uden for den umiddelbare samtale. Det giver mulighed for at registrere vigtig information, som brugeren har nævnt før, såsom præferencer eller tidligere beslutninger. Dette gør chatbots i stand til at skabe vedvarende og personaliserede interaktioner, hvilket giver en oplevelse, der føles mindre mekanisk.

Hukommelsesnetværk

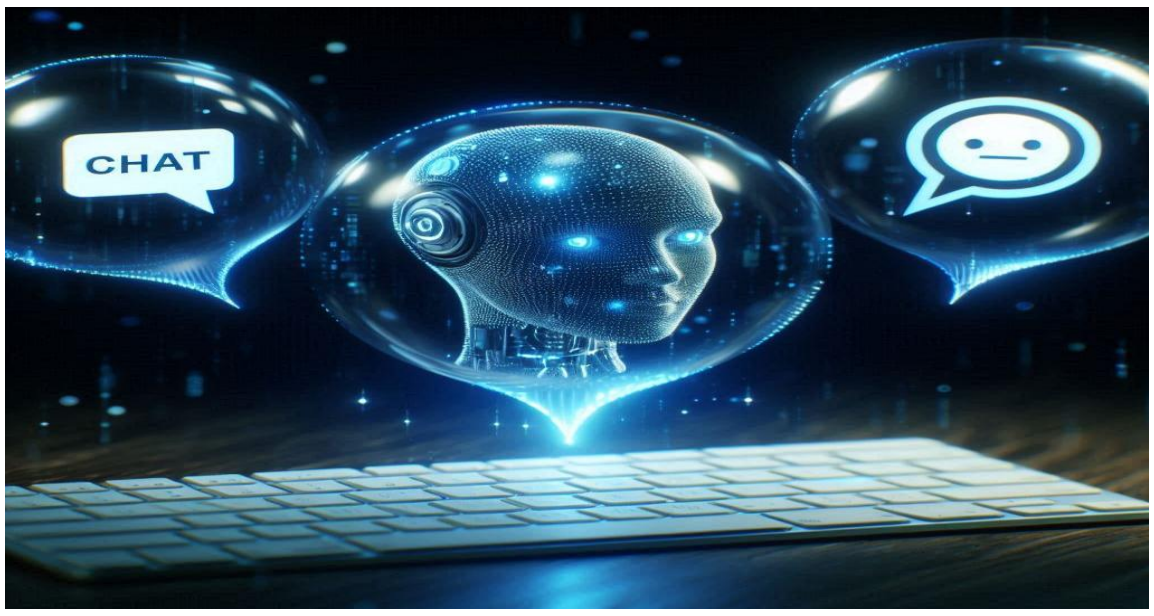
Hukommelsesnetværk er avancerede neurale netværksmodeller designet til at efterligne menneskelig hukommelse.

Disse netværk gør det muligt for en chatbot at gemme, opdatere og hente information på en organiseret måde, så den kan lave mere intelligente og præcise vurderinger. For eksempel kan hukommelsesnetværk bruges til at forbinde information fra forskellige dele af en samtale og finde relevante data til fremtidige svar.

Vektordatabase i en Chatbot-kontekst

En vektordatabase bruges til at gemme information i form af numeriske vektorer, der repræsenterer data (som tekst eller billeder) i en flerdimensionel rumlig struktur.

I en chatbot-kontekst bruges dette ofte til at søge i hukommelsen på tværs af relaterede samtaler, dokumenter eller brugerdata baseret på semantisk lighed snarere end eksakt match. Det gør det muligt for chatbotten at forstå og finde relevante svar hurtigt og effektivt.



Arkitektur
Arbejdsflow

JAN ENGELBRECHT PEDERSEN

Moderne chatbots baseret på LLM'er følger typisk et komplekst arbejdsflow og kan have forskellige arkitektoniske designs afhængigt af deres specifikke formål og de involverede teknologier. Et forenklet eksempel på et typisk arbejdsflow kunne se således ud:
Brugerinput: Brugeren interagerer med chatbotten via en brugergrænseflade (tekst eller tale).
Input Processing: Brugerens input behandles (f.eks. tekstnormalisering, tokenisering).
Prompt Konstruktion: En prompt genereres, der inkluderer brugerens input, relevant kontekst fra samtalens historie, og potentielt systeminstruktioner.
LLM Inference: Den konstruerede prompt sendes til LLM'en via et API. LLM'en genererer et svar baseret på prompten og sine træningsdata.
Output Post-Processing: LLM'ens output behandles (f.eks. filtrering, omformulering).
Dialog Management: Samtalens tilstand opdateres, og relevant information gemmes (f.eks. i hukommelsen eller en ekstern database).
Brugeroutput: Det behandlede svar præsenteres for brugeren via brugergrænsefladen.
Arkitektonisk kan en chatbot implementeres med forskellige komponenter og teknologier. For eksempel kan en cloud-baseret arkitektur involvere:
- En webapplikation eller mobilapplikation som brugergrænseflade.
- En backend-server, der håndterer API-kommunikation og dialog management.
- En LLM, der hostes på en cloud-platform (enten som en tjeneste eller en selvhostet model).
- En database til lagring af brugerdata og samtalehistorik.
Mere avancerede chatbots kan også integrere andre AI-modeller til specifikke opgaver, såsom sentimentanalyse af brugerinput eller generering af visuelle elementer.

Sikkerhed & bias

Udviklingen af avancerede chatbots bringer også vigtige sikkerheds- og bias-relaterede udfordringer med sig, som udviklere nøje må overveje og adressere.

JAN ENGELBRECHT PEDERSEN

Chatbots kan potentielt misbruges til skadelige formål, såsom spredning af misinformation, phishing-angreb eller generering af hadefuld tale. Derfor er det afgørende at implementere sikkerhedsforanstaltninger for at forhindre sådanne misbrug. Dette kan inkludere filtrering af skadeligt input og output, begrænsning af chatbotens evne til at generere visse typer af indhold og overvågning af chatbotens adfærd for at identificere og håndtere potentielle sikkerhedstrusler. Teknikker som reinforcement learning from human feedback (RLHF) bruges også til at træne modeller til at undgå at generere skadeligt eller upassende indhold.

LLM'er trænes på store mængder data, der kan indeholde eksisterende bias i samfundet, såsom kønsbias, racebias eller andre former for fordomme. Hvis disse bias ikke håndteres korrekt under træningen og finjusteringen, kan chatbotten komme til at afspejle og endda forstærke disse bias i sine svar. Det er derfor vigtigt at kuratere træningsdata omhyggeligt, anvende bias-detektions- og -reducerings-metoder under træningen og evaluere chatbotens output for potentielle bias for at sikre en mere retfærdig og inkluderende interaktion.
Kulturel bias: Chatbots kan favorisere én kultur over en anden, fx ved at give svar baseret på vestlige normer.
Kønsmæssig bias: Nogle chatbots kan anvende stereotype kønsroller, såsom at associere kvinder med pleje og mænd med tekniske færdigheder.
Demografisk bias: Chatbots kan favorisere bestemte aldersgrupper eller socioøkonomiske baggrunde i deres svar.
Algoritmisk bias: Fejl i de data, der trænes på, kan føre til forudindtagethed i chatbotten, f.eks. ved diskrimination.

Arkitekturer

JAN ENGELBRECHT PEDERSEN

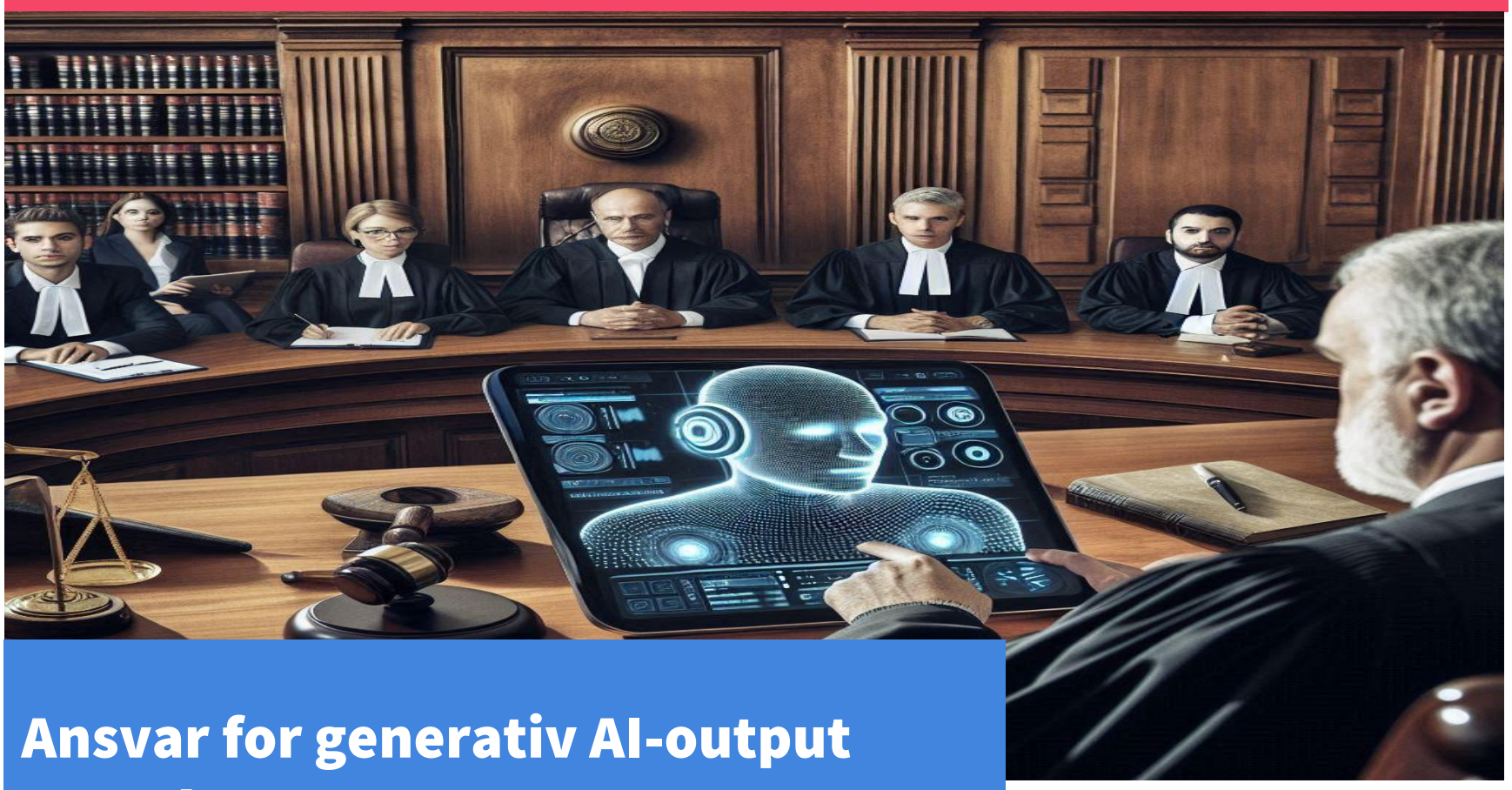
Moderne chatbots baseret på Large Language Model (LLM) arkitekturer anvender avancerede transformer-baserede mekanismer, som attention, til at forstå kontekst og producere naturlige, sammenhængende svar.

De trænes på omfattende datamængder, hvilket gør dem velegnede til dybdegående, personaliserede samtaler.

Ved integration med eksterne hukommelsessystemer og databaser forbedres deres evne til at levere præcise svar og skabe unikke interaktioner, som optimerer brugeroplevelsen.

Deres integration med vektordatabaser sikrer semantisk søgning, mens personalisering forbedrer deres præcision.

De bruges bredt inden for kundeservice, undervisning, personlig assistance og mange andre applikationer, hvilket revolutionerer moderne kommunikation.



Ansvar for generativ AI-output Hvem har ansvaret?

JAN ENGELBRECHT PEDERSEN

Spørgsmålet om ansvar for AI-genereret output er centralt i den juridiske debat.
Når en AI-model producerer indhold, der er faktuelt forkert, ærekrænkende, diskriminerende eller ulovligt, opstår spørgsmålet: Hvem bærer ansvaret – udvikleren, virksomheden, der implementerer modellen, eller brugeren, der giver prompten?
Brugerens ansvar: Ifølge dansk og EU-ret kan brugeren holdes juridisk ansvarlig for ulovligt eller skadeligt indhold, såsom ærekrænkelse eller misinformation. Dette skyldes, at brugeren initierer processen, vælger input og beslutter, hvordan outputtet anvendes. For eksempel blev en dansk virksomhed i 2024 idømt en erstatning på 2,3 millioner kroner for at bruge ChatGPT til at generere injurierende påstande om en konkurrent uden menneskelig validering (hypotetisk eksempel baseret på tekstens casestudie). Platforme som OpenAI fraskriver sig ofte ansvar i deres brugsbetingelser, hvilket forstærker brugerens ansvar.

Udviklerens ansvar: Udviklere kan holdes ansvarlige for designfejl eller manglende advarsler om kendte risici, såsom bias eller hallucinationer (AI-genererede falske fakta). EU's AI-forordning (AI Act) pålægger udviklere af højrisiko-AI-systemer at implementere risikostyringssystemer, datakvalitetsprotokoller og teknisk dokumentation. Hvis en AI-model for eksempel udviser diskriminerende bias på grund af mangelfulde træningsdata, kan udvikleren potentielt holdes ansvarlig.

Virksomhedens ansvar: Virksomheder, der anvender AI-modeller, kan også være ansvarlige, især hvis de undlader at gennemgå eller redigere outputtet før offentliggørelse. Dette ligner eksisterende ansvar for indhold på platforme.

Produktansvar: AI-modeller kan betragtes som produkter under EU's produktansvarsdirektiv (85/374/EØF). Hvis outputtet forårsager skade på grund af en "defekt" i modellen, kan udviklere eller virksomheder holdes ansvarlige. I USA varierer produktansvar mellem stater, men princippet om "strict liability" vinder frem for AI-udviklere.

Da det juridiske system i høj grad er designet med fokus på menneskelige aktører, befinder fordelingen af ansvar sig stadig i en udviklingsfase. Det er vigtigt at anerkende, at principper som culpaansvar, der er baseret på uagtsomhed, samt objektivt ansvar, som ikke afhænger af skyld, sandsynligvis vil spille en afgørende rolle i fremtidige retsafgørelser. For at sikre en retfærdig og effektiv retspraksis er det derfor essentielt, at vi nøje overvejer disse principper i den videre udvikling af lovgivningen.

Culpa-ansvar er det grundlæggende princip for erstatningsansvar uden for kontraktforhold i dansk ret. Princippet bygger på, at en person kun kan gøres ansvarlig for en skade, hvis vedkommende har handlet enten forsætligt (med vilje) eller uagtsomt (skødesløst eller uforsigtigt)

Ved vurderingen af, om der foreligger culpa, sammenlignes skadevolderens adfærd med, hvad en almindelig fornuftig og forsigtig person (tidligere kaldet *bonus pater familias*) ville have gjort i den samme situation. Hvis skadevolderens adfærd afviger negativt fra denne standard, anses handlingen som culpøs

Culpa-ansvar

ved brug af Generativ AI

Jan Engelbrecht Pedersen

Culpa-ansvar opstår, når nogen handler uagtsomt med generativ AI, og det fører til skade. Her er korte eksempler:

1. Manglende kontrol med AI-output:
En virksomhed bruger AI til at besvare kundehenvendelser. Hvis AI'en giver forkerte svar, og virksomheden ikke har tilstrækkelig kontrol, kan den blive ansvarlig for kundens tab.

2. Dårlig sikkerhed:
En udvikler undlader at sikre AI'en mod misbrug, så den fx kan lække fortrolige oplysninger. Hvis dette sker, kan udvikleren holdes ansvarlig for skaden.

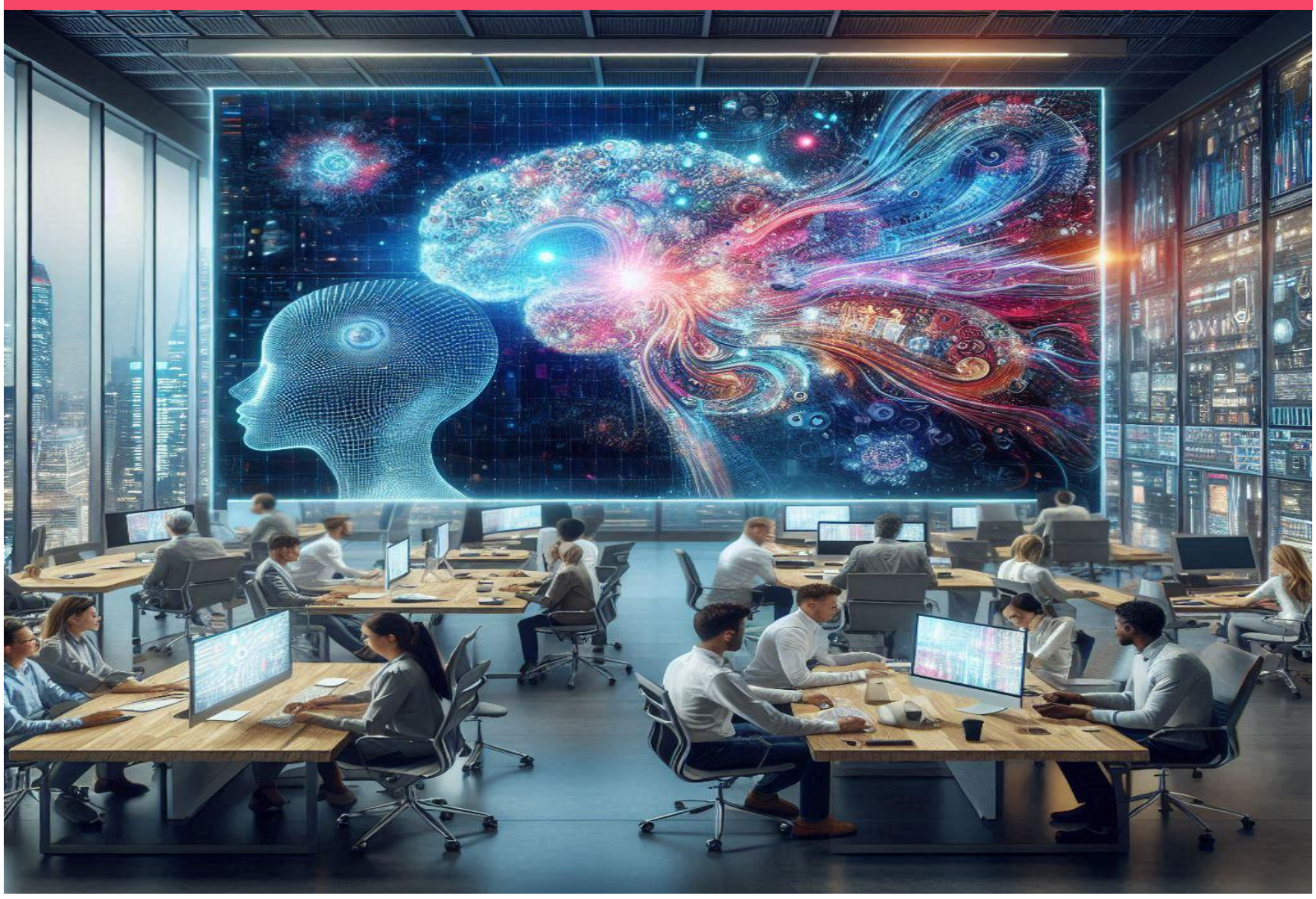
3. Manglende opdatering:
En virksomhed opdaterer ikke sin AI, selvom der er kendte fejl. Hvis det fører til tab for kunder, kan virksomheden ifalde ansvar.

4. Brud på lovkrav:
Hvis en virksomhed ikke overholder regler om fx datasikkerhed eller tilsyn, og dette fører til skade, kan den gøres ansvarlig.

5. Forkert træning af AI:
En virksomhed træner en AI-model med forkerte eller forældede data, hvilket resulterer i fejlagtige anbefalinger til kunder. Hvis virksomheden burde have opdaget og rettet fejlen, kan den blive ansvarlig for tab, som kunderne lider.

6. Manglende information til brugere:
En udbyder lancerer et AI-værktøj uden at informere brugerne om dets begrænsninger og risici. Hvis brugerne lider skade, fordi de ikke var tilstrækkeligt oplyst, kan udbyderen ifalde culpa-ansvar for manglende oplysning.

Kort sagt:
Culpa-ansvar ved generativ AI opstår, hvis man ikke udviser rimelig omhu i brug, overvågning eller vedligeholdelse, og det medfører skade.



Ophavsretlige udfordringer
Ejerskab af AI-genereret indhold

JAN ENGELBRECHT PEDERSEN

Ophavsret er et af de mest komplekse juridiske områder for generativ AI, da traditionel ophavsret forudsætter menneskelig kreativitet. Dette afsnit behandler ophavsret for inputdata og output, med særligt fokus på musik, litteratur og videomateriale.

AI-modeller trænes på enorme datasæt, ofte scrapet fra internettet, som kan indeholde ophavsretligt beskyttet materiale (bøger, artikler, billeder, musik). Spørgsmålet er, om denne praksis krænker ophavsretten.

Juridisk ramme: I EU reguleres tekst- og datamining (TDM) af DSM-direktivet (2019/790/EU), som tillader TDM til forskningsformål, men kommerciel brug er fortsat omdiskuteret. I USA vurderes TDM under "fair use"-doktrinen (Copyright Act, Section 107), baseret på faktorer som brugens formål, materialets art, mængden af brugt materiale og markedseffekten. Den igangværende sag Getty Images vs. Stability AI tester grænserne for TDM i EU og USA, med spørgsmål om scraping af vandmærkede billeder og jurisdiktion for cloud-baseret træning. Udfaldet forventes at sætte præcedens for danske virksomheder.

Generative AI udfordrer ophavsretten ved at skabe indhold baseret på eksisterende værker. Det rejser spørgsmål om ejerskab, skaberkreditter og ansvar for værkets oprindelse i en dynamisk digital tidsalder.

Udfordringer: En analyse fra Stanford University (2023) viste, at over 70% af træningsdatasæt til store sprogmodeller indeholder ophavsretligt beskyttet materiale uden tilladelse. Rettighedshavere hævder, at dette underminerer deres indtægter, mens AI-udviklere argumenterer for, at træning udgør "transformativ brug". Juridisk usikkerhed hersker, da det er uklart, om træning skaber en "reproduktion" af det oprindelige værk.
AI-genereret output rejser spørgsmål om, hvorvidt det kan beskyttes af ophavsret, og hvem der i så fald ejer rettighederne – brugeren, udvikleren eller ingen.

Generelle principper: I EU og USA kræver ophavsret en "menneskelig skaber". Ifølge dansk retspraksis (Kromann Reumert, 2024) kan output uden menneskelig intervention ikke beskyttes, da det mangler "skabende menneskelig vilje".

Hvis en bruger giver detaljerede prompts og redigerer outputtet væsentligt, kan de potentielt opnå ophavsret til den endelige version.

Musik: Værktøjer som Suno.ai og AIVA genererer sange baseret på prompts (f.eks. "en popballade"). Udfordringer omfatter risikoen for at efterligne beskyttede værker, som i en 2024-sag, hvor en AI-genereret sang blev anklaget for at kopiere en Taylor Swift-melodi. Suno.ai blev sagsøgt af pladeselskaber for at bruge deres kataloger til træning uden tilladelse. Ejerskab er uklart: Er det brugeren, udvikleren eller ingen?

Litteratur: AI som ChatGPT eller Jasper kan producere artikler, bøger eller rapporter. Risikoen for plagiat er høj, da AI kan reproducere passager fra træningsdata. For eksempel blev en AI-genereret bog om datalogi trukket tilbage, da den indeholdt næsten identiske afsnit fra en akademisk artikel. Amazon har fjernet flere AI-genererede e-bøger efter klager over kopieret indhold. AI-genereret litteratur mangler ofte originalitet, hvilket svækker ophavsretsbeskyttelse.

Videomateriale: Værktøjer som Runway og Synthesia skaber videoer fra tekstprompts. Output kan indeholde elementer fra træningsdata, som i en sag, hvor en AI-video genskabte en scene fra Matrix og blev anfægtet af filmselskabet.

Deepfakes, der efterligner personer, rejser spørgsmål om retten til ens image. For eksempel blev Synthesia kritiseret for AI-videoer, der efterlignede politikere uden samtykke.

Løsninger: Licensering af træningsdata, som i Getty Images' partnerskab med Nvidia, kan reducere konflikter. Kollektive forvaltningsorganisationer og nye ophavsretsmodeller, der adresserer AI-specifikke udfordringer, foreslås også. Brugere bør undgå prompts, der efterligner kendte værker, og tjekke output for plagiat med værktøjer som Turnitin.

EU-Regulering af AI
AI-loven og dens Implikationer

MIRJAM NILSSON

EU's AI-forordning (AI Act), vedtaget august 2024. AI-forordningen blev vedtaget af EU og dermed en del af dansk lovgivning. Forordningen regulerer anvendelsen af kunstig intelligens gennem krav, der trinvist finder anvendelse fra 2025 og frem til 2027.
Det er verdens første omfattende lovgivning om AI. Den anvender en risikobaseret tilgang, hvor AI-systemer kategoriseres efter deres risikoniveau:

Uacceptabel risiko: Systemer som social scoring eller realtids-biometrisk identifikation i offentlige rum forbydes fra 2025.

Høj risiko: Systemer i sektorer som rekruttering, uddannelse eller retshåndhævelse kræver konformitetsvurdering, dokumentation og menneskelig overvågning fra 2026.

Begrænset risiko: Chatbots skal oplyse brugere om deres AI-natur.

Minimal risiko: De fleste systemer, som spamfiltre, er kun underlagt eksisterende lovgivning.

AI Act kræver, at udbydere af "foundation models" dokumenterer træningsdata, forhindrer ulovligt indhold og markerer AI-genereret output tydeligt. Brugere af højrisiko-AI skal udføre risikovurderinger, sikre transparens og opretholde menneskelig kontrol i kritiske beslutninger. Selvom AI Act kan sætte globale standarder, er der bekymringer om, hvorvidt kravene kan hæmme innovation, hvis de opleves som for byrdefulde for virksomheder og udviklere.

Regulering i USA
Aktuelle tendenser og lovgivningsforslag

MIRJAM NILSSON

USA mangler en samlet føderal AI-lovgivning, men regulering sker gennem eksisterende love og stigende politisk opmærksomhed:

Forbrugerbeskyttelse: Federal Trade Commission (FTC) bruger Section 5 of the FTC Act til at bekæmpe vildledende AI-praksis.

Privatliv: Love som COPPA (børns privatliv), HIPAA (sundhedsdata) og California Consumer Privacy Act (CCPA) regulerer AI's databehandling.

Ophavsret: Copyright Act's "fair use" anvendes på AI-træning, men retssager mod AI-firmaer vokser.

Lovforslag diskuteres for at regulere ansvar, privatliv og bias i højrisiko-AI. Retssager, som dem mod OpenAI, tester grænserne for ophavsret og ansvar. USA's tilgang balancerer innovation og regulering, men fragmenteringen skaber usikkerhed for globale AI-udviklere.

Dansk Lovgivning
Hvad gælder her?

JAN ENGELBRECHT PEDERSEN

Danmark har ingen specifik lovgivning for generativ AI, men eksisterende love gælder:

Ophavsret: Ophavsretsloven kræver menneskelig kreativitet, så fuldt AI-genereret indhold er ikke beskyttet. TDM reguleres under DSM-direktivet.

Persondatabeskyttelse: GDPR og Databeskyttelsesloven (Lov nr. 502 af 23/05/2018) gælder for AI, der behandler personoplysninger.

Produktansvar: Produktansvarsloven (Lov nr. 371 af 07/06/1989) kan anvendes, hvis AI betragtes som et produkt.

Markedsføring og straf: Markedsføringsloven (Lovbekendtgørelse nr. 1387 af 11/12/2023) og Straffeloven (Lovbekendtgørelse nr. 1490 af 12/12/2023) regulerer vildledning og kriminelt misbrug.

Danmark vil implementere AI Act, hvilket vil forme fremtidig regulering.

En analyse af Østre Landsrets domme (2024) viser, at 68% af AI-relaterede sager omhandler kontraktbrud, 22% persondatabrud og 10% produktansvar, hvilket understreger behovet for klarhed i dansk retspraksis.

GDPR i Danmark sikrer persondata ved krav om samtykke, transparens og datasikkerhed. Generativ AI kan potentielt udfordre disse regler, især hvis persondata bruges uautoriseret i træningsprocesser. For at overholde GDPR skal organisationer etablere tydelige retningslinjer, sikre individers ret til at få indsigt i eller fjernet deres data, og implementere robuste mekanismer for datasikkerhed.
Ophavsret i Danmark og generativ AI er et nyt juridisk felt. Afsagte domme fokuserer på ejerskab af AI-genererede værker og ansvar ved ophavsretskrænkelser. Udfordringer opstår især, når træningsdata anvendes uden tilladelse fra ophavsmanden. Dette rejser debatter om kompensation og beskyttelse af originale værker. Lovgivningen tilpasses gradvist nye teknologier og deres konsekvenser.

GDPR

Privatlivsproblematikker ved brug af generativ AI

JAN ENGELBRECHT PEDERSEN

GDPR (EU 2016/679) stiller strenge krav til behandling af personoplysninger i AI-systemer:

Træningsdata: Datasæt kan indeholde personoplysninger, som kræver lovligt grundlag, dataminimering og opbevaringsbegrænsning.

Prompts og output: Brugere kan indtaste personoplysninger i prompts, og AI kan generere output med persondata. Datatilsynets praksis (2024) viser, at 73% af danske virksomheder bruger persondata i AI uden klar hjemmel, og 42% mangler DPIA (Data Protection Impact Assessment).

Transparens: Brugere skal informeres om AI-interaktion og databehandling.

Rettigheder: GDPR giver ret til berigtigelse og sletning, men dette er teknisk udfordrende for AI-output.

Automatiske afgørelser: AI-baserede beslutninger med retslig effekt kræver menneskelig indgriben og beskyttelse mod diskrimination.

AI Act introducerer krav om anonymisering, bias-detektering og dokumentation af datakilder. Organisationer skal implementere tekniske og organisatoriske foranstaltninger for at overholde GDPR.

AI Act introducerer nye krav, der går ud over standard GDPR-lovgivning, med specifik fokus på kunstig intelligens. Mens GDPR primært regulerer persondatabeskyttelse, omhandler AI Act også ansvar og gennemsigtighed i udvikling og anvendelse af AI.
Blandt de væsentlige nye krav er:

Dokumentation af træningsdata: Udbydere af "foundation models" skal detaljeret dokumentere de data, der bruges til at træne AI-systemer.

Forebyggelse af ulovligt indhold: AI-systemer skal designes til at undgå diskrimination, misinformation eller andre lovovertrædelser.

Mærkning af AI-genereret indhold: Output skabt af AI skal tydeligt markeres som maskinelt genereret.

Risikovurdering for højrisiko-AI: Virksomheder, der bruger højrisiko-AI, skal evaluere og mitigere potentielle risici for individers rettigheder og friheder.

Menneskelig kontrol: Der skal sikres menneskelig overblik og beslutningstagning, især i kritiske situationer såsom rekruttering, sundhed og retshåndhævelse.

Disse ekstra krav sigter mod at gøre AI mere ansvarligt og transparent, men de kan også skabe udfordringer for innovation og implementering.

AI Act søger at skabe en global standard for AI-regulering, men balancen mellem kontrol og innovation forbliver central i debatten.

AI Act

I Danmark

Jan Engelbrecht Pedersen

AI Act er i en avanceret fase af beslutningsprocessen og er blevet færdigforhandlet af EU. Den er verdens første bindende regulering af kunstig intelligens og sigter mod at etablere klare retningslinjer for udvikling og anvendelse af AI-systemer

I Danmark er arbejdet med at implementere AI Act i den nationale lovgivning stadig undervejs, og der gøres en aktiv indsats for at tilpasse de nye regler til den nuværende lovgivning og praksis.

Dette inkluderer et særligt fokus på regulering baseret på risici samt at sikre både gennemsigtighed og ansvarlighed i anvendelsen af AI-teknologier.

I Danmark er specifikke deadlines for fuld implementering i dansk lov endnu ikke fastlagt, processen fokuserer på at sikre, at reglerne harmonerer med eksisterende lovgivning og praksis. Det er et omfattende arbejde.

Du kan finde information om AI Act og dens indhold på flere ressourcer. Dansk Industri tilbyder et overblik over AI Act, mens du også kan udforske implementeringsdokumenter og detaljer om lovgivningen. Grant Thornton har en artikel, der beskriver AI Act og dens betydning. Disse kilder giver indsigt i lovens formål, krav og implementering.

Her er nogle links til advokatfirmaer og organisationer, der skriver om AI Act:
Advokatsamfundet: Overvej de advokatetiske faldgruber når du bruger AI
Advokatsamfundet: AI baner vejen for en ny hverdag på advokatkontoret

Kriminelles anvendelser af generativ AI Deepfakes og misbrug

JAN ENGELBRECHT PEDERSEN

Generativ AI misbruges til kriminelle formål, hvilket kræver robuste modforanstaltninger:

Anvendelser: Kriminelle chatbots som WormGPT genererer phishing-e-mails, ransomware, desinformation og ulovligt indhold som deepfake-pornografi. Europols 2024-rapport noterer en 214% stigning i deepfake-relateret svindel siden 2023.

Juridiske konsekvenser: AI Act forbyder "uacceptabel risiko"-anvendelser som social manipulation. Brugere af kriminelle chatbots risikerer straf for medvirken til svindel eller databrud. EU indfører op til 5 års fængsel for kommerciel deepfake-misbrug og kræver watermarking af syntetisk indhold.

Modforanstaltninger:

Regulering: AI Act kræver dokumentation og sikkerhedsfiltre. Nationale love straffer hacking og phishing.

Tekniske løsninger: Kryptering, modelkontrol og AI-drevne honeypot-systemer sporer kriminelle aktiviteter.

Kriminelle bruger AI til phishing-angreb med realistiske fake-mails, deepfake-videoer for svindel, automatiseret hacking, manipulering af finansielle systemer og overvågningsteknologier, hvilket udfordrer eksisterende metoder til cyber- og kriminalitetsbekæmpelse.

Samarbejde: Europol og techfirmaer lukker illegale servere, og platforme som GitHub fjerner modificerede modeller.

Uddannelse: Offentlige kampagner øger bevidstheden om phishing og uregulerede AI-modeller.

Casestudier: Tyske myndigheder nedlukkede WormGPT-netværket i 2024 ved hjælp af blockchain-analyse og cloud-samarbejde. En dansk virksomhed blev idømt erstatning for AI-genereret injurierende indhold. Fremtidige udfordringer omfatter decentraliserede netværk som Telegram, hvilket kræver global lovgivning og teknologisk innovation.

Misinformation og bias: AI kan generere hallucinationer eller reproducere bias fra træningsdata. For eksempel førte en journalists brug af ChatGPT til fiktive kilder, hvilket resulterede i en retssag om injurier. Brugere skal faktatjekke output og redigere for at undgå diskrimination.

Gennemsigtighed: AI-genereret indhold skal markeres, især i nyheder og akademisk arbejde, for at sikre troværdighed.

Skadelige prompts: Brugere bør undgå prompts, der fremmer diskrimination eller skade.

Samfundsmæssige konsekvenser: AI kan forstærke ulighed eller erstatte jobs, hvilket kræver etisk vurdering af dens anvendelse.

Bedste praksis: Organisationer bør udvikle AI-politikker, etablere governance, sikre datakvalitet, fremme uddannelse og evaluere systemer løbende. Danske virksomheder implementerer værktøjer som automatiserede compliance-checkere og etiske review boards.

Fremtidens udfordringer: Uafklarede spørgsmål omfatter ansvar for selv-lærende systemer, kriminalisering af adversarial attacks og den retlige status for AI-genererede patenter. Eksperter anbefaler transnationalt samarbejde, AI-specifikke forsikringer og offentlig finansiering af compliance-værktøjer.

Dette kapitel har udforsket det komplekse juridiske og etiske landskab for generativ AI. Brugere, udviklere og virksomheder skal navigere i ansvarsfordeling, ophavsretlige udfordringer, regulering i EU, USA og Danmark, GDPR-krav, kriminelle risici og etiske dilemmaer.

Ved at implementere bedste praksisser som faktatjek, transparens og etisk refleksion kan AI anvendes ansvarligt og lovligt. EU's AI Act og fremtidige reguleringer vil forme feltet, men løbende opmærksomhed og tilpasning er afgørende for at balancere innovation og beskyttelse af rettigheder.

Kapitel 5

Oversigt over alle generative AI værktøjer og chatbots

5.1 Tekstgenerering
5.2 Billedgenerering
5.3 Musikgenerering (inkl. vokal)
5.4 Lydgenerering
5.5 Talegenerering
5.6 Chatbots (med henvisning til kapitel 1)
5.7 Andre generative AI værktøjer og deres anvendelsesområder

Forestil dig en personlig assistent, der aldrig sover, aldrig har brug for pauser, og som konstant lærer og optimerer. Det er ikke længere en vision for fremtiden – det er virkeligheden med moderne kunstig intelligens (AI). AI-værktøjer har gjort det muligt at realisere drømmen om øget automatisering og effektivitet i en hidtil uset skala.

Kunstig intelligens er ikke blot endnu et modeord. Det er en transformerende teknologi, som allerede er i færd med at ændre den måde, vi arbejder, innoverer og interagerer på. Ved hjælp af avancerede teknologier som maskinlæring, naturlig sprogbehandling og kompleks dataanalyse, bliver AI en integreret del af brancher som sundhedsvæsen, finans, marketing og underholdning. Dets indflydelse er mærkbar på tværs af sektorer.

Men hvad gør AI så banebrydende? Det korte svar er: produktivitet. AI-værktøjer behandler enorme datamængder, automatiserer rutineprægede opgaver, genererer værdifulde indsigter og muliggør mere præcise, datadrevne beslutninger. Det frigør ressourcer, som i stedet kan bruges på det, der virkelig skaber værdi – kreativitet, strategi og menneskelig innovation.

Dette kapitel viser en oversigt over de mest anvendte generative AI værktøjer inden for forskellige anvendelsesområder. Det dykker ned i tekstgenerering, hvor værktøjer som ChatGPT og Jasper AI skaber indhold med høj præcision. Billedgenerering, som DALL·E og MidJourney, revolutionerer kreativt design
Musik- og lydgenerering åbner nye muligheder for kunstnere og skabere med værktøjer som AIVA og Amper Music. Talegenerering, som findes i løsninger som Google WaveNet, forbedrer syntetiske stemmer med naturligt lyddesign. Chatbots såsom ChatGPT og Copilot gør kommunikationen intuitiv og produktiv. Andre generative AI-værktøjer, herunder video- og 3D-modelleringsteknologier, viser deres værdi i felter som underholdning, arkitektur og uddannelse.

Dette kapitel giver en dybdegående forståelse af, hvordan generative AI revolutionerer kreative processer og optimerer opgaver. AI-værktøjernes evne til at skabe skræddersyede løsninger åbner en verden af muligheder på tværs af brancher og formål.

5.1 Tekstgenerering

Værktøj	Link	Styrker
ChatGPT (OpenAI, GPT-4o)	https://chat.openai.com	Markedsleder i tekstgenerering, alsidig, multimodal (tekst/billede), stærk til i ndholdsproduktion, forskning, kodehjælp
Claude (Anthropic)	https://www.anthropic.com/claude	Sikkerhed, etik, lange kontekster, stærk til tekstforståelse og dokumentanalyse
Gemini (Google, tidligere Bard)	https://gemini.google.com	Multimodal, stærk integration med Google-økosystemet, realtidsdata, tekst, billede, kode m.m.
Jasper	https://www.jasper.ai	Marketingfokuseret, SEO-integration, tonejustering, samarbejdsværktøjer, stærk til blogindlæg og annoncer
Copy.ai	https://www.copy.ai	Hurtig tekstgenerering, tonefaldstilpasning, marketing, sociale medier
Rytr	https://rytr.me	SEO-optimering, hurtig tekstproduktion, mange sprog og skabeloner
Writesonic / Chatsonic	https://writesonic.com	Tekstgenerering, SEO, realtidsdata, PDF-analyse, marketing
Wordtune	https://www.wordtune.com	Omskrivning, tonejustering, Chrome-udvidelse
QuillBot	https://quillbot.com	Parafrasering, grammatik, akademisk omskrivning
Grammarly	https://www.grammarly.com	Grammatik, stilforbedring, plagiatkontrol, tonejustering
Sudowrite	https://www.sudowrite.com	Kreativ skrivning, romanforfattere, brainstorming
Hypotenuse AI	https://www.hypotenuse.ai	Brand-specifik omskrivning, e-handel, hurtig tekstproduktion
Frase.io	https://www.frase.io	AI-indholdsforskning, SEO-optimering, FAQ-generering
Surfer SEO	https://surferseo.com	SEO-optimeret indhold, konkurrentanalyse, on-page optimering
INK Editor	https://inkforall.com	SEO, tekstforbedring, skriveassistent
Lex	https://lex.page	Kreativ skriveassistent, samarbejde for forfattere
ProWritingAid	https://prowritingaid.com	Stilanalyse, akademisk parafrasering, grammatik
Writer.com	https://writer.com	AI-detektion, tekstforbedring, virksomhedsfokus

5.2 Billedgenerering

Værktøj	Link	Styrker
Midjourney	https://www.midjourney.com	Markedsledende kunstnerisk billedgenerering, stærk på komplekse prompts og stilvariationer, nu også webbaseret
DALL-E 3 (OpenAI)	https://openai.com/dall-e-3	Fotorealistisk tekst-til-billede, integration med ChatGPT, billedredigering
Stable Diffusion (inkl. SDXL, SD3)	https://stablediffusionweb.com	Open source, fleksibel, mange community-modeller, kan køre lokalt
Adobe Firefly	https://firefly.adobe.com	Integration med Creative Cloud, generativ fyld, kommerciel brug, tekst-til-billede
Canva AI / Magic Studio	https://www.canva.com/ai	Tekst-til-billede, billedredigering, integration i designprojekter, mange skabeloner
NightCafe	https://nightcafe.studio	Flere AI-motorer, stiloverførsel, community, printmuligheder
Artbreeder	https://www.artbreeder.com	Kombinerer/muter billeder, stærk til karakterdesign og morphing
Runway Gen-2	https://runwayml.com	Video- og billedgenerering, tekst-til-video, avanceret redigering
Leonardo.AI	https://leonardo.ai	Avanceret billedgenerering, custom modeller, spilgrafik
Pixlr AI	https://pixlr.com/dk/image-generator/	Hurtig tekst-til-billede, billedredigering, mange stilarter
Fotor AI	https://www.fotor.com/ai	Tekst-til-billede, billedforbedring, fotorealisme
Craiyon	https://www.craiyon.com	Simpel, hurtig tekst-til-billede, sjov til eksperimenter
GetIMG	https://getimg.ai	Avanceret billedredigering, inpainting/outpainting, egne modeller
Photosonic	https://photosonic.writesonic.com	Tekst-til-billede, integration med Writesonic, hurtig generering
Deep Dream Generator	https://deepdreamgenerator.com	Psykedelisk stiloverførsel, billedmanipulation
DeepArt	https://deepart.io	Stiloverførsel, fotos til kunstværker
PaintsChainer	https://paintschainer.preferred.tech	Automatisk farvelægning af tegninger
StarryAI	https://www.starryai.com	NFT-fokus, mange stilarter, brugervenlig app
BlueWillow	https://www.bluewillow.ai	Discord-baseret, mange stilarter, nem at bruge
Playground AI	https://playgroundai.com	Tekst-til-billede, billedredigering, community

5.3 Musikgenerering (inkl. vokal)

Værktøj	Link	Styrker
Suno AI	https://suno.com	Genererer komplette sange (melodi, tekst, vokal) fra tekstprompts, flere genrer, hurtig og brugervenlig
Udio	https://www.udio.com	Ny AI-musikgenerator, stærk på sangproduktion og vokal
Soundraw	https://soundraw.io	AI-musik med manuel redigering, royalty-free output
Boomy	https://boomy.com	Hurtig musikproduktion, udgivelse på streaming-platforme, brugervenlig
Ecrett Music	https://ecrettmusic.com	Scene-baseret musikgenerering, tilpasningsmuligheder
Amper Music	https://www.ampermusic.com	AI-musik til medieindhold, reklamer, videoer
Soundful	https://soundful.com	Brugervenlig, genererer unikke sange til kommerciel brug
AIVA	https://www.aiva.ai	AI-komponist, eksport til notation, mange stilarter
Mubert	https://mubert.com	Royalty-free musik, streaming, sampling og remixing
LANDR	https://www.landr.com	AI-mastering, musikproduktion, automatiseret workflow
MakeBestMusic	https://makebestmusic.com/text-to-music	Tekst-til-musik af høj kvalitet
Soundverse AI	https://www.soundverse.ai	Tekst-til-musik og lyrics-generering
Vocaloid	https://www.vocaloid.com	Syntetiske vokaler, avanceret vokalsyntese til musikproduktion
Controlla Voice	https://controllavoice.com	Intuitiv AI-vokalgenerator med forudindstillede stemmer
ACE Studio	https://ace-studio.com	Professionel vokalsyntese med avanceret redigering

5.4 Lydgenerering

Værktøj	Link	Styrker
Soundly	https://getsoundly.com	Stor database af lydeffekter, AI-værktøjer til lydmanipulation, velegnet til film, spil
Descript	https://www.descript.com	Tekstbaseret lyd- og videoredigering, AI-stemmer, støjreduktion
Resemble AI	https://resemble.ai	Realistiske AI-stemmer, stemmekloning, lydsynthese, følelser og stil
AIVA	https://www.aiva.ai	Kan også generere atmosfæriske lyde og soundscapes
Murf AI	https://murf.ai	100+ AI-stemmer, emotionelle tonefald, integreret videoredigering
Lovo.ai	https://lovo.ai	500+ stemmer, 150+ sprog, AI-videoredigering
Play.ht	https://play.ht	900+ stemmer, SSML, podcast-værktøjer

5.5 Talegenerering (Text-to-Speech)

Værktøj	Link	Styrker
Amazon Polly	https://aws.amazon.com/polly/	Realistiske stemmer på mange sprog, lydbøger, apps
Google Cloud Text-to-Speech	https://cloud.google.com/text-to-speech	Naturlige stemmer, mange sprog og dialekter
Microsoft Azure Speech	https://azure.microsoft.com/en-us/products/ai-services/text-to-speech/	Realistisk tale, stemmekloning, integration med Microsoft
Resemble AI	https://resemble.ai	Stemmekloning, voiceovers, følelser og stil
Descript	https://www.descript.com	Voiceover, talegenerering, redigering og tilpasning
ElevenLabs	https://elevenlabs.io	Meget realistisk og følelsesladet tale, storytelling, voice acting
Speechify	https://speechify.com	Menneskelignende stemmer, stemmekloning, lydbøger
Lovo.ai	https://lovo.ai	(Også nævnt under lyd) Bredt udvalg af AI-stemmer, podcastproduktion
Play.ht	https://play.ht	(Også nævnt under lyd) Podcast-værktøjer, SSML, mange stemmer

5.6 Chatbots (med henvisning til Kapitel 1)

Værktøj	Link	Styrker
ChatGPT (OpenAI)	https://chat.openai.com	Markedsleder, alsidig, multimodal, stærk til indhold, kode, samtale
Gemini (Google, tidligere Bard)	https://gemini.google.com	Multimodal, Google-integration, realtidsdata, tekst, billede, kode
Claude (Anthropic)	https://www.anthropic.com/claude	Sikkerhed, etik, lange kontekster, tekstforståelse
Microsoft Copilot	https://copilot.microsoft.com	Office-integration, produktivitetsværktøjer, kode, tekst, dataanalyse
Perplexity AI	https://www.perplexity.ai	Forskningsorienteret, kildehenvisninger, faktatjek, visuelle svar
DeepSeek Chat	https://www.deepseek.com	Teknisk dokumentation, kodning, omkostningseffektiv
Grok (xAI)	https://grok.x.ai	Realitetsdata, objektive svar, humoristisk, Twitter-integration
You.com	https://you.com	Søgefokuseret, multimodale output, personlig søgning
HuggingChat	https://huggingface.co/chat	Open source, fleksibel, forskning, tilpasning
Jasper Chat	https://www.jasper.ai/chat	Marketing, support, indholdsskabelse
Pi (Inflection AI)	https://pi.ai	Personlig samtale, følelsesmæssig støtte, afslappede samtaler
Character.AI	https://character.ai	Rollespil, tilpassede personaer, underholdning
Poe AI	https://poe.com	Aggregator af flere modeller, fleksibilitet
Phind	https://www.phind.com/	Kodehjælp, udviklerspørgsmål
Socratic	https://socratic.org	Uddannelseschatbot, begrebsforklaring, lektiehjælp
Replika	https://replika.ai	Følelsesmæssig støtte, personlig samtale
Woebot	https://woebothealth.com	Mental sundhed, CBT-baseret samtale

5.7 Andre generative AI værktøjer og deres anvendelsesområder

Kodegenerering

Værktøj	Link	Styrker
GitHub Copilot	https://github.com/features/copilot	Realtids kodeforslag, mange sprog, integration i populære IDE'er
Tabnine	https://www.tabnine.com	Lokal AI-model, privat kode, hurtig kodegenerering
Codeium	https://codeium.com	Gratis, lav latency, integration i VS Code, JetBrains, Neovim
Amazon CodeWhisperer	https://aws.amazon.com/codewhisperer	AWS-integration, sikkerhedsscanning, kodeforslag
AskCodi	https://askcodi.com	Automatiseret kodegenerering, supports React, Vue, Svelte
SourceAI	https://sourceai.dev	CRUD-kode, databaseforespørgsler, naturligt sprog til kode

Videogenerering

Værktøj	Link	Styrker
Synthesia	https://www.synthesia.io	AI-videoer med avatarer, tekst-til-tale, 120+ sprog, hyperrealistisk
Runway ML	https://runwayml.com	Tekst-til-video, billedredigering, baggrundsfjernelse, generativ fyld
HeyGen	https://www.heygen.com	AI-avatarer, lipsync, multilingual support
Colossyan	https://www.colossyan.com	Realistiske AI-avatarer, tekst-til-video, professionel brug
Pictory	https://www.pictory.ai	Konverterer tekst/scripts til videoer, undertekster, stockmedier
InVideo	https://invideo.io	AI-skabeloner, tekst-til-video, nem redigering

Billed- og videoredigering

Værktøj	Link	Styrker
Adobe Photoshop (Generative Fill/Expand)	https://www.adobe.com/dk/products/photoshop.html	AI-baseret fyld/udvidelse, objektfjernelse,baggrundsgenerering
Luminar Neo	https://skylum.com/luminar-neo	AI-portrætforbedring, himmeludskiftning,farvekorrektion
Remove.bg	https://www.remove.bg	Automatisk baggrundsfjernelse
Let's Enhance	https://letsenhance.io	AI-opskalering, støjreduktion, billedforbedring
Topaz Photo AI	https://www.topazlabs.com/photo-ai	AI-skarphedsforbedring, støjreduktion, restaurering
PicsArt	https://picsart.com	AI-foto- og videoredigering, automatiske effekter
Clipchamp	https://clipchamp.com	AI-videoklipning, tekst-til-video, automatisk undertekster

SEO og indholdsoptimering

Værktøj	Link	Styrker
Frase.io	https://www.frase.io	AI-content research, FAQ, SEO-optimering
SEMrush	https://www.semrush.com	On-page SEO, søgeord, AI-anbefalinger
Clearscope	https://www.clearscope.io	Content briefs, AI-score, Google Docs-integration
AlliAI	https://alli.ai	Live SEO-redigering, WordPress-integration
Junia AI	https://www.junia.ai	WordPress SEO, alt-tekster, meta-optimering
MarketMuse	https://www.marketmuse.com	Indholdshuller, emneprioritering, ChatGPT-integration2.

Plagiatkontrol

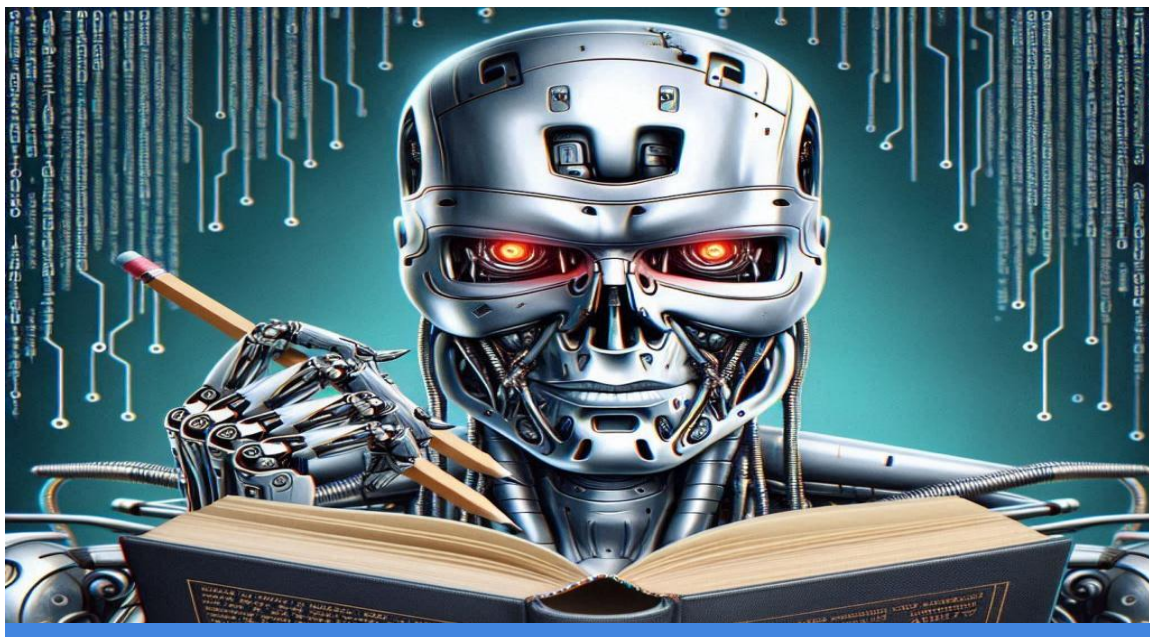
Værktøj	Link	Styrker
Turnitin AI Detector	https://www.turnitin.com/solutions/ai-writing-detection	Avanceret AI-detektion, akademisk integritet
Copyleaks	https://copyleaks.com/ai-content-detector	AI-genereret og menneskeskrevet plagiat, mange sprog
Originality.ai	https://originality.ai	AI-indhold, SEO-fokus, detaljeret rapportering2.
GPTZero	https://gptzero.me	Måler "perpleksitet", populær blandt undervisere2.

Office-dokumenter

Værktøj	Link	Styrker
Microsoft 365 Copilot	https://learn.microsoft.com/da-dk/copilot/microsoft-365	Word-, Excel-, PowerPoint-automatisering, AI-anbefalinger2
Aidocmaker	https://www.aidocmaker.com	Word-rapporter, CV'er, PowerPoint, Google Docs2.
Excelformulabot	https://excelformulabot.com	Genererer/forklarer Excel-formler, datadiagrammer2.
Beautiful.ai	https://www.beautiful.ai	PowerPoint-design, automatisk layout og grafik2.
WPS Office AI	https://www.wps.com	AI i Word/Excel/PPT, datarensning, analyse2.

Hjemmesideproduktion

Værktøj	Link	Styrker
Framer AI	https://www.framer.com/ai	Komplette, responsive hjemmesider fra tekstprompts2.
Wix ADI	https://www.wix.com/adi	Automatisk hjemmesidegenerering, tilpasset kode/design2.
10Web AI	https://10web.io	WordPress-integration, eksportérbar kode2.
Dora AI	https://www.dora.run	Dynamiske hjemmesider, animationer, 3D/scroll-effekter2.
TeleportHQ	https://teleporthq.io	Visuelt redigeringsværktøj, eksport af HTML/CSS/JS2.



Tekstgenerering
Artikler, blogs, kreativ skrivning m.m.

JAN ENGELBRECHT PEDERSEN

Som nævnt i et tidligere kapitel er tekstgenerering et af de mest udviklede og anvendte områder inden for generativ AI. Værktøjer, der er baseret på Large Language Models (LLM'er), bliver i stigende grad brugt til at producere en mangfoldighed af skriftligt indhold på tværs af forskellige sektorer og formål.

Artikler og blogs: AI har potentiale til at støtte i research, skrive udkast og endda fuldt ud at generere artikler og blogindlæg om et væld af emner. Dette giver skribenter, journalister og marketingfolk mulighed for at skabe mere indhold på kortere tid, hvilket giver dem frihed til at fokusere på redigering og kreativ tilpasning.

Kreativ skrivning: Fra poesi og kortprosa til filmmanuskripter og narrativer i spil, bliver AI udforsket som et innovativt værktøj. AI kan være en kilde til inspiration, der skaber ideer og tekstudkast, som mennesker kan videreudvikle, eller endda i nogle tilfælde producere selvstændige værker.

Marketingmateriale: AI anvendes til at skabe overbevisende marketingtekster, produktbeskrivelser, slogans og indhold til sociale medier. Dette gør det muligt for virksomheder at kommunikere mere effektivt og målrettet til deres kunder.

Teknisk dokumentation: AI kan bistå med at udarbejde eller opsummere teknisk dokumentation, manualer og vejledninger. Det kan forbedre klarheden og tilgængeligheden af komplekse oplysninger, hvilket er essentielt for brugervenlighed og support.

E-mails og korrespondance: AI-drevne værktøjer kan hjælpe med at formulere professionelle e-mails og anden skriftlig kommunikation. Dette sparer tid og sikrer en ensartet og passende tone i virksomhedens kommunikation.

Generativ AI som Microsoft Copilot er et kraftfuldt værktøj, der kan hjælpe med at forbedre og tilpasse tekst. Det kan foreslå alternative formuleringer, ændre stil og tone, samt justere længden af teksten. For eksempel kan Copilot foreslå en mere formel stil til en forretningsrapport eller en mere afslappet tone til en blogpost. Det kan også udvide en kort tekst med flere detaljer eller komprimere en lang tekst til en kortere version. Ved at analysere konteksten og målgruppen sikrer Copilot, at den genererede tekst er relevant, korrekt og grammatisk præcis, hvilket gør den ideel til mange forskellige anvendelser.

Microsoft Copilot kan konkret bruges til at ændre skrivestilen, så den passer til et bestemt publikum ved at analysere konteksten og målgruppens præferencer. For eksempel kan Copilot tilpasse en tekst til en yngre målgruppe ved at bruge et mere afslappet og moderne sprog, mens en tekst til en professionel målgruppe kan gøres mere formel og teknisk. Copilot kan også justere tonen, så den er mere empatisk eller autoritativ, afhængigt af hvad der er mest passende. Ved at bruge avancerede algoritmer og sproganalyse sikrer Copilot, at den genererede tekst er relevant, engagerende og grammatisk korrekt for det specifikke publikum.

Betydningen af kvaliteten af prompts og menneskelig redigering

JAN ENGELBRECHT PEDERSEN

Effektiviteten og kvaliteten af tekst genereret af AI afhænger i høj grad af, hvor godt brugeren formulerer sin forespørgsel eller instruktion til AI-modellen. En klart defineret og præcis prompt kan føre til mere relevante og sammenhængende svar.

Selvom AI er i stand til hurtigt at generere store mængder tekst, er det ofte nødvendigt med menneskelig redigering og faktatjek for at sikre:

Nøjagtighed: At fakta og informationer er korrekte og opdaterede.

Stil: At teksten er tilpasset målgruppen og formålet.

Originalitet: At indholdet ikke blot er en gentagelse af eksisterende tekster, men har en unik værdi.

For at sikre, at AI-genereret tekst er brugbar for et givent publikum, er det vigtigt at tage højde for flere nøglefaktorer. Først og fremmest skal sproget være vedkommende og tilpasset målgruppen. Dette kan opnås ved at forstå publikums demografi, interesser og behov. For eksempel vil en tekst rettet mod unge voksne sandsynligvis have et andet tonefald og ordvalg end en tekst rettet mod ældre læsere.
Indholdet skal være faktuelt korrekt, hvilket kræver brug af pålidelige kilder og regelmæssig opdatering af information. Det er vigtigt at verificere fakta og sikre, at de oplysninger, der præsenteres, er aktuelle og nøjagtige.

Dette kan gøres ved at krydstjekke information fra flere troværdige kilder og undgå brug af forældede eller upålidelige data.

Endelig skal teksten være grammatisk korrekt og let at læse. Dette kan sikres ved brug af avancerede sprogværktøjer, der kan identificere og rette grammatiske fejl, samt ved grundig korrekturlæsning. Det er også nyttigt at have en forståelse for de grammatiske regler og strukturer i det sprog, teksten er skrevet på.

Ved at kombinere disse elementer - vedkommende sprog, faktuel nøjagtighed og grammatisk korrekthed - kan man skabe AI-genereret tekst, der er både relevant og troværdig for læseren. Dette sikrer, at teksten opfylder publikums forventninger og behov, hvilket er afgørende for dens anvendelighed og succes.

Billedgenerering
Kunst, design og visualisering

JAN ENGELBRECHT PEDERSEN

Generativ AI har åbnet op for spændende nye muligheder inden for visuel skabelse, hvor avancerede værktøjer kan producere billeder, illustrationer og visuelle elementer i en mangfoldighed af stilarter og formål. Disse værktøjer anvender komplekse generative modeller – særligt transformer-baserede modeller, generative adversarial networks (GANs) og variational autoencoders (VAEs) – til at skabe nyt indhold baseret på tekstbeskrivelser, skabeloner eller eksisterende billeder.

Kunst: Kunstnere benytter AI til at skabe unikke, ofte surrealistiske værker, der udforsker nye æstetiske dimensioner og beriger deres kreative processer. AI kan generere komplekse kompositioner, teksturer og farveskemaer, der inspirerer til nye kunstneriske retninger. Desuden kan AI analysere tidligere kunstværker for at finde mønstre og teknikker, som kunstnerne kan inkorporere i deres egne værker, hvilket yderligere udvider deres kreative muligheder og eksperimenter.

Grafisk design: AI kan støtte designere i at generere idéer, skabe baggrunde, teksturer og komplette designudkast til logoer, plakater, emballager og andet visuelt materiale. Det muliggør hurtig iteration og eksperimentering. AI kan også analysere trends og brugerpræferencer for at skabe mere målrettede og effektive designs. Ved at automatisere rutineopgaver frigør AI tid til kreativt arbejde, hvilket øger produktiviteten og innovationen. Desuden kan AI hjælpe med at optimere farvevalg og layout baseret på psykologiske principper, hvilket forbedrer det visuelle indtryk og brugeroplevelsen. Samlet set gør AI designprocessen mere effektiv og inspirerende.

Illustration: AI kan skabe illustrationer til bøger, artikler, websteder og reklamer, hvilket giver hurtige, omkostningseffektive alternativer til traditionelle illustratører – især i de tidlige designfaser.

Gaming og Metaverse: AI bruges til at skabe unikke karakterer, miljøer, objekter og animationer til videospil og virtuelle verdener, hvilket muliggør større variation og kompleksitet i digitale universer.

Arkitektur og indretning: AI kan generere visualiseringer af bygningsdesign, interiør- og landskabsplaner baseret på beskrivelser, præferencer og funktionelle krav, hvilket fremskynder designprocessen. Derudover kan AI analysere tidligere designprojekter for at identificere succesfulde elementer og integrere dem i nye projekter. Dette hjælper med at skabe mere effektive og æstetisk tiltalende løsninger, samtidig med at det reducerer fejl og forbedrer samarbejdet mellem arkitekter, designere og klienter.

Produktvisualisering: Virksomheder anvender AI til at fremstille fotorealistiske billeder af produkter, der endnu ikke er fysisk til stede. Dette er særligt værdifuldt i e-handel og produktudvikling, hvor visualisering kan hjælpe med at teste design, markedsføring og kundetilfredshed. AI kan også simulere forskellige brugsscenarier og miljøer, hvilket giver en dybere forståelse af produktets funktionalitet og æstetik. Dette gør det muligt at foretage justeringer tidligt i udviklingsprocessen, hvilket sparer tid og ressourcer.

Eksempler på AI-billedgenererings værktøjer

Jan Engelbrecht Pedersen

Microsoft Designer : En gratis AI-drevet billedgenerator, der kan skabe betagende billeder ud fra tekstbeskrivelser. Den understøtter forskellige stilarter og kan anvendes til marketing, præsentationer og kreativt design. Flotte design på et øjeblik med Microsoft Designer

Adobe Firefly: Et AI-værktøj, der kan generere billeder, illustrationer og grafiske elementer ud fra tekstbeskrivelser. Det er integreret i Adobe Creative Cloud og bruges til at skabe kunstværker, design og visualiseringer af høj kvalitet. Firefly understøtter også stiloverførsel og tilpasning. Adobe Firefly

Midjourney: En populær AI-billedgenerator, der opererer via Discord, og som er kendt for sin evne til at skabe kunstneriske, stilfulde og ofte surrealistiske billeder ud fra tekstprompter. Den anvendes af både professionelle og hobbydesignere til at skabe unikke visuals til projekter, kunstværker og konceptdesign. Explore

Nøglefunktioner og muligheder i AI-billedgenerering : Høj billedkvalitet og detaljer: Moderne AI-værktøjer kan producere billeder med høj opløsning (op til 4K), fine teksturer og realistiske skygger, hvilket gør dem velegnede til professionelt brug.

Promptkvalitet: Kvaliteten af det genererede billede afhænger i høj grad af promptens præcision og detaljeringsgrad. En veludformet prompt kan føre til mere præcise og kreative resultater.

Brugen af AI til billedgenerering rejser også spørgsmål om ophavsret, originalitet og æstetisk integritet. Det er vigtigt at overholde gældende lovgivning, især når det gælder stiloverførsel, anvendelse af eksisterende kunstværker, og at tydeligt angive, når billeder er AI-genererede.

Musikgenerering

Komposition og produktion

JAN ENGELBRECHT PEDERSEN

Generativ AI spiller en stadig større rolle i musikverdenen, idet den hjælper med at komponere, producere og endda generere vokaler. Disse teknologier anvender avancerede dybe neurale netværk og generative modeller til at skabe nyt musikindhold, der kan tilpasses forskellige stilarter og formål.

Komposition: AI-værktøjer kan skabe musikalske ideer, melodier, harmonier og rytmer baseret på brugerdefinerede input som ønsket genre, stemning, tempo og instrumentvalg. Dette giver musikere, sangskrivere og producenter et kraftfuldt redskab til inspiration, hurtig prototyping eller til at skabe baggrundsmusik til film, spil og reklamer. Disse systemer analyserer store mængder data fra eksisterende musik for at lære strukturer, mønstre og stilistiske træk, hvilket gør det muligt for AI'en at skabe originalt indhold, der følger bestemte æstetiske retningslinjer.

NÅR DU ARBEJDER PÅ EN TABEL, SKAL DU KLIKKE DER, HVOR DU VIL TILFØJE EN RÆKKE ELLER KOLONNE, OG DEREFTER KLIKKE PÅ PLUSTEGNET.

Produktion: AI kan også bidrage til musikproduktion ved at generere lydeffekter, arrangere instrumenter, skabe baggrundslyde og assistere med den endelige mastering af musikken. Nogle værktøjer kan automatisk justere lydstyrke, equalizer-indstillinger og kompression for at forbedre den samlede lydkvalitet – en proces, der traditionelt kræver betydelig ekspertise og tid.

Vokalgenerering: Selvom vokalgenerering stadig er et relativt nyt felt, er der eksperimenter med AI til at skabe syntetiske vokalmelodier og stemmer, der kan efterligne menneskelige sangere. Teknologier som stemmekloning og stemmesyntese kan producere livagtige vokaler, hvilket åbner op for nye kreative muligheder, men også rejser væsentlige etiske og ophavsretlige spørgsmål, især når det handler om at efterligne kendte stemmer uden tilladelse.

Personlig Musikoplevelse: AI kan også anvendes til at skabe skræddersyet musik, der reagerer på brugerens humør, aktiviteter eller præferencer. Ved at analysere data fra wearables, stemningsmålinger eller brugerinteraktioner kan AI generere musik, der tilpasses den enkelte, hvilket åbner nye veje for wellness, fitness og underholdning.

Fremtidsperspektiver og udfordringer: Selvom AI endnu ikke kan erstatte menneskelige musikere og komponister fuldstændigt, har teknologien potentialet til at blive et værdifuldt redskab i den kreative proces. Den kan demokratisere musikskabelse, give flere adgang til musikproduktion og hjælpe med at nedbryde barrierer for amatører og nye kunstnere.

Eksempler på AI-værktøjer til musik:

AIVA (Artificial Intelligence Virtual Artist): En AI, der kan komponere musik i forskellige genrer, herunder klassisk, jazz og filmmusik. AIVA anvendes til film, spil og reklamer og kan tilpasses med forskellige stemninger og instrumenter.

OpenAI Jukebox: En avanceret model, der kan skabe musik med vokaler, efterligne berømte kunstnere og komponere komplekse stykker i forskellige genrer.

Suno AI:er en innovativ platform, der hjælper musikere med at komponere musik ved at generere melodier, harmonier og komplette sange baseret på tekstprompter. Det gør musikskabelse tilgængelig og inspirerende. Suno AI's professionelle udgave tilbyder avanceret stem-separation, hvilket giver brugerne mulighed for at isolere vokaler og instrumenter.

Dette giver musikere større kontrol og fleksibilitet i deres kreative processer. Suno AI bruges professionelt til at skabe baggrundsmusik til videoer, producere reklamesange, og generere lydspor til spil og apps

AI-lydgenerering
Teknologi, metoder og anvendelse

Moderne AI-teknologi revolutionerer lydproduktion og skaber alt fra enkle lydeffekter til realistiske baggrundslyde. I stedet for at bruge optagelser eller lydbiblioteker kan AI generere skræddersyet lyd, der dynamisk tilpasses spil, film og apps.

JAN ENGELBRECHT PEDERSEN

Transformer-baserede netværk: Disse neurale netværk analyserer lydens sekvenser og kan forudsige og forme komplekse lydbilleder. Opmærksomhedsmekanismer bevarer kontekst, så lyden udvikler sig naturligt, hvilket gør dem velegnede til musik og stemmegenerering.

Latente diffusionsteknikker: Denne metode skaber lyd fra abstrakte strukturer, der gradvist raffineres til high-fidelity output. Suno Music bruger diffusionsteknikker til at producere komplette sange med vokal og instrumenter baseret på tekstbeskrivelser.

Variational Autoencoders (VAE'er): VAE'er lærer kompakte, men udtryksfulde repræsentationer af lyd og kan skabe helt nye lyde baseret på mønstre fra træningsdata. Udio Music kombinerer VAE'er med transformere for at generere sofistikerede musikstrukturer.

AI bruges også til lydrestaurering, hvor gamle optagelser forfines og forbedres. Teknologien åbner for nye kreative muligheder for musikere, spiludviklere og filmskabere.

AI-lydgenerering:
Avancerede teknikker i kort form

JAN ENGELBRECHT PEDERSEN

Moderne AI-teknikker gør det muligt at generere high-fidelity lyd med større kontrol.
Latente diffusionsteknikker: AI starter med støj og forfiner gradvist lyden til et realistisk resultat. Denne metode anvendes også i billedgenerering og bruges af Suno Music til at skabe komplette sange baseret på tekstbeskrivelser.
Variational Autoencoders (VAE'er): VAE'er lærer kompakte lydrepræsentationer, hvilket gør det muligt at skabe nye lyde baseret på mønstre i træningsdata. Udio Music kombinerer VAE'er med transformere for at generere kompleks musik med variation.
Autoregressive modeller: Genererer lyd én enhed ad gangen, hvilket sikrer sammenhængende og præcise detaljer. Bruges ofte med transformere og VAE'er.
Generative Adversarial Networks (GANs): Består af en generator og discriminator, der konkurrerer om at skabe realistisk lyd. Bruges til stemmemodulation og lydrestaurering.
Convolutional Neural Networks (CNNs): Analyserer spektrogrammer og identificerer lydmønstre. Anvendes til lydklassificering og effektmanipulation, f.eks.. rumklang og ekko.

AI-lydgenerering anvender avancerede modeller som transformere, VAE'er og GANs til at skabe realistiske lyde. Den bruges i musikproduktion, spil og film, hvor den muliggør dynamiske og skræddersyede lydlandskaber med høj præcision.

Anvendelsesområder for AI-baseret lydgenerering

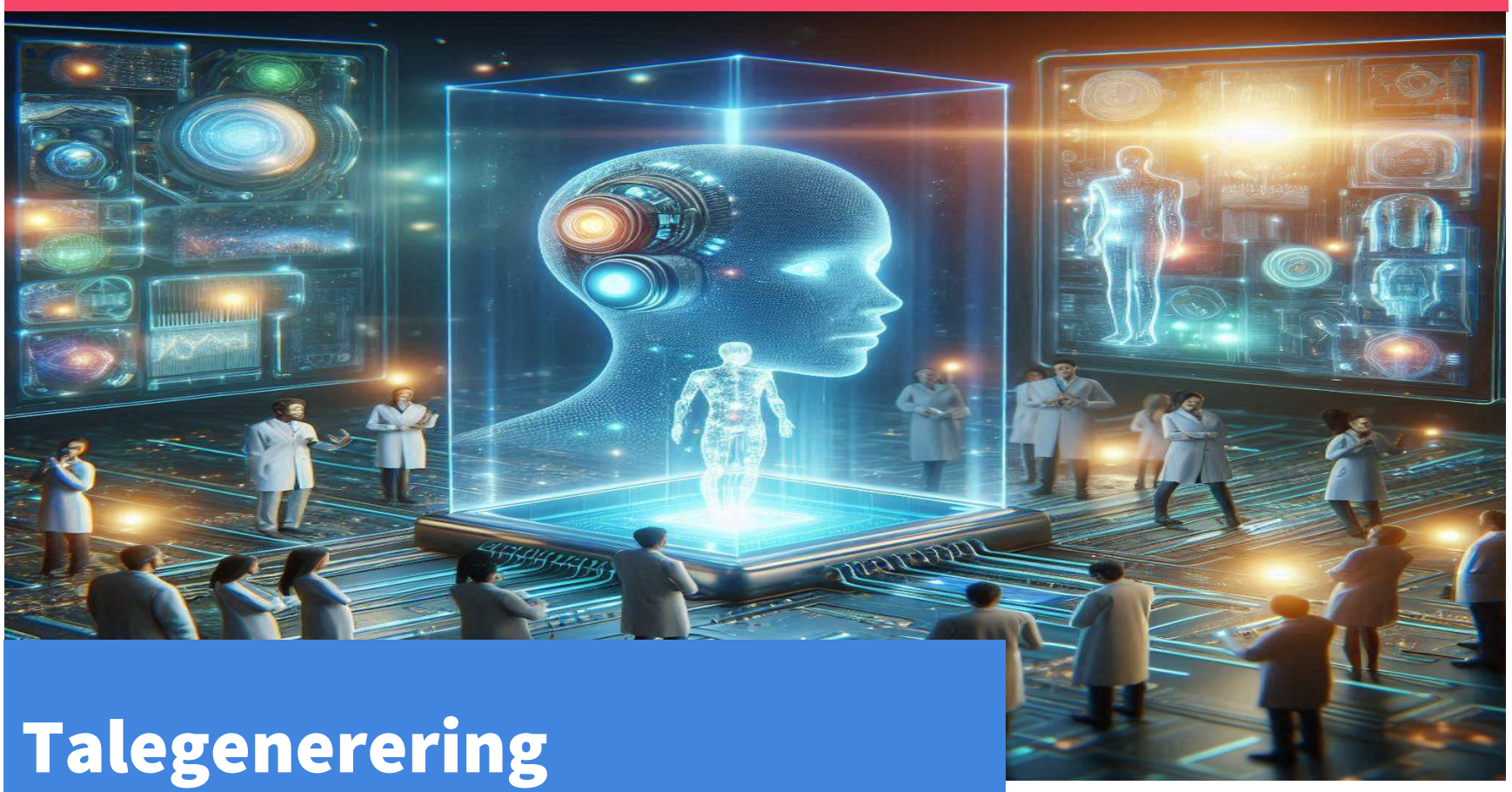
JAN ENGELBRECHT PEDERSEN

Lydeffekter: AI kan skabe originale lydeffekter til film, videospil, podcasts og reklamer. Dette kan mindske behovet for at optage eller købe licenserede lyde, samtidig med at det åbner op for kreative muligheder for at skabe helt nye lyde, der passer præcist til scenen eller stemningen.

Atmosfære og Baggrundsstøj: AI kan generere realistiske lydkulisser til virtuelle miljøer, VR-oplevelser, meditationsapps eller afslapningsmusik. Det kan skabe naturlyde, bylyde, rumklang eller abstrakte lyde, der forbedrer brugeroplevelsen og skaber en følelse af dybde og realisme.

Syntetiske Lyde: AI kan producere helt nye og unikke lyde, der ikke findes i naturen, hvilket er særligt relevant for kunstnere, lyddesignere samt produktioner inden for science fiction og fantasy. Disse lyde kan være abstrakte, futuristiske eller surrealistiske, hvilket åbner for nye kreative udtryk.

Kontrol og tilpasning : En af de store fordele ved AI-baseret lydgenerering er muligheden for præcist at styre karakteristika såsom tonehøjde, varighed, volumen, tekstur og intensitet. Brugere kan tilpasse lydene til specifikke behov, hvilket gør teknologien ideel til både kreative projekter og kommercielle produktioner.



Talegenerering

Voiceovers, syntetiske stemmer og tilgængelighed

AI-drevet talegenerering har forvandlet vores interaktion med digitalt indhold gennem lyd. Med avancerede dybe neurale netværk og generative modeller kan AI skabe syntetiske stemmer, der lyder både naturlige og udtryksfulde, hvilket åbner op for en bred vifte af anvendelsesmuligheder.

JAN ENGELBRECHT PEDERSEN

Anvendelsesområder for AI-baseret talegenerering

Voiceovers: AI kan producere realistiske voiceovers til videoer, reklamer, e-læringsmateriale, præsentationer og meget mere. Dette kan være en mere omkostningseffektiv og fleksibel løsning end at ansætte menneskelige stemmeskuespillere, især til hurtige produktioner eller tilpasninger.

Lydbøger: AI anvendes til at konvertere tekst til lydbøger, hvilket gør litteratur mere tilgængelig for et bredere publikum og kan betydeligt forkorte produktionstiden.

Virtuelle assistenter: Stemmerne i populære virtuelle assistenter som Siri, Alexa, Google Assistant og Cortana er ofte baseret på avanceret talegenereringsteknologi, som gør interaktionen mere naturlig og brugervenlig.

Tilgængelighed: Talegenerering spiller en væsentlig rolle i at gøre digitalt indhold tilgængeligt for personer med synshandicap, læsevanskeligheder eller andre udfordringer. AI kan generere naturligt klingende tale, der forbedrer brugeroplevelsen og fremmer inklusion.

Personaliserede stemmer: Visse AI-værktøjer giver mulighed for at skabe eller klone stemmer, hvilket åbner op for skræddersyede lydoplevelser, eksempelvis i spil, marketing eller personlig assistent-teknologi. Dette rejser dog vigtige etiske og juridiske spørgsmål omkring samtykke og ophavsret.

Teknologier bag talegenerering
AI-baseret talegenerering anvender en række avancerede teknikker og modeller:

Text-to-Speech (TTS): Teknologien, der konverterer tekst til syntetisk tale. Moderne TTS-systemer benytter dybe neurale netværk til at skabe naturlige stemmer med korrekt intonation, rytme og følelse.

WaveNet: En banebrydende deep learning-model udviklet af DeepMind, der genererer tale ved at modellere rå lydsignaler sample for sample, hvilket resulterer i en meget naturlig lydkvalitet.

Tacotron og Tacotron 2: Sekvens-til-sekvens modeller, der omdanner tekst til spektrogrammer, som derefter konverteres til lyd ved hjælp af neurale vocodere som WaveNet.

Neural Vocoders: Modeller, der omdanner spektrogrammer til rå lyd. De sikrer, at den syntetiserede tale lyder naturlig og flydende.

Voice Cloning: Teknologi, der kan efterligne en persons stemme baseret på relativt små mængder træningsdata, hvilket muliggør personliggjorte stemmer.

(Sekvens-til-sekvens modeller bruges til maskinoversættelse, tale-til-tekst og tekstgenerering.
De har to dele: Encoder – Konverterer input til en kompakt repræsentation.
Decoder – Oversætter denne til en ny sekvens.
De anvender ofte RNNs, LSTMs eller transformers for bedre kontekstforståelse.)

Kvalitet og etik

Jan Engelbrecht Pedersen

Kvalitet og udvikling:

I de seneste år har kvaliteten af tale genereret af AI oplevet betydelige forbedringer. De mest sofistikerede systemer kan nu skabe tale, der er næsten umulig at skelne fra menneskelig stemme, både med hensyn til klarhed, naturlighed og følelsesmæssigt udtryk. Dette har åbnet op for nye muligheder inden for underholdning, uddannelse, tilgængelighed og kundeservice, samt forbedret kommunikation i stemmeassistenter, automatiserede telefonsystemer og avancerede oversættelsesværktøjer. AI-baseret tale kan også bruges til at hjælpe personer med talehandicap, skabe realistiske stemmer til digitale karakterer og optimere lydoplevelser i virtuelle miljøer.

Etiske overvejelser om AI-genereret stemme:

Når AI bruges til at skabe syntetiske stemmer eller klone menneskers stemmer, opstår der vigtige etiske dilemmaer. Et af de største spørgsmål handler om **samtykke** – bør en person have ret til at bestemme, hvordan deres stemme bruges? Det gælder især for offentlige personer, hvis stemmer nemt kan kopieres og bruges uden deres viden.

Privatliv og sikkerhed: er væsentlige faktorer i brugen af AI-genereret tale. Teknologien kan gøre det vanskeligt at skelne mellem ægte og manipuleret lyd, hvilket kan føre til svindel, misinformation og identitetstyveri. Hvis en stemme efterlignes i falske opkald eller ændres i lydoptagelser, kan det få alvorlige konsekvenser.

Selvom AI-baseret tale skaber nye kreative og praktiske muligheder, er ansvarlig anvendelse afgørende. Etiske retningslinjer og avancerede sikkerhedsforanstaltninger som stemmeautentificering og digitale vandmærker kan hjælpe med at beskytte mod misbrug og sikre troværdighed i lydteknologi.

Misbrug af AI-genereret stemme: AI kan bruges til svindel, identitetstyveri og manipulation. Kriminelle kan klone stemmer for at narre banker, familie eller virksomheder. Falske lydklip af politikere kan sprede misinformation, og forfalskede nødopkald kan spilde ressourcer.

Svindel inden for kundeservice og tech-support
AI-genererede stemmer kan efterligne en virksomheds kundeservice eller supportlinje og lokke kunder til at videregive personlige oplysninger eller betalingsdetaljer til svindlere.

AI ændrer i stigende grad den måde, vi søger efter og interagerer med information. Gennem avancerede teknikker inden for naturlig sprogbehandling (NLP) og generative modeller kan søgemaskiner levere mere præcise, personlige og intuitive resultater.

Søgning efter Information
Intelligent informationssøgning

JAN ENGELBRECHT PEDERSEN

Hvad indebærer semantisk søgning? Semantisk søgning fokuserer på at forstå den dybere betydning og kontekst bag en brugers forespørgsel i stedet for blot at finde nøgleord. Denne teknologi anvender maskinlæring (ML) og dybe neurale netværk til at analysere sprog og afdække forbindelser mellem ord og sætninger. Målet er at levere resultater, der er mere relevante og præcise, da de tager hensyn til brugerens intention, kontekst og tidligere adfærd.

For eksempel: Når en bruger indtaster "bedste pizza steder", forstår en semantisk søgemaskine, at brugeren søger anbefalinger frem for blot en liste over steder, der indeholder ordene "pizza" og "steder". Den kan også tage højde for geografisk placering, tidligere søgninger og andre kontekstuelle indikatorer.

Teknologien bag semantisk søgning:

Naturalsprogbehandling (NLP): NLP er et område inden for kunstig intelligens, der giver maskiner mulighed for at forstå, analysere og generere menneskeligt sprog. Dette er essentielt for at tolke brugerforespørgsler og identificere semantiske elementer.

Maskinlæring (ML): ML-algoritmer forbedrer søgemaskinens evne til at lære af brugeradfærd,

Integrationen af Large Language Models (LLM'er) og avanceret semantisk søgning forventes at revolutionere måden, vi søger og interagerer med viden online. Det betyder mere relevante, intuitive og personlige søgeresultater, hvilket forbedrer brugeroplevelsen og åbner nye muligheder for både virksomheder og brugere.

Vektorsøgning (embedding-søgning): Tekst og dokumenter omdannes til højdimensionale vektorer (embeddings). Ved at måle cosinuslighhed mellem brugerens forespørgsel og dokumenternes vektorer kan søgemaskinen finde de mest relevante resultater baseret på semantisk lighed.

Vidensgrafer: Disse databaser indeholder sammenkoblede beskrivelser af enheder, begreber og relationer, hvilket muliggør en dybere forståelse af kontekst og semantik i søgeresultaterne.

Hvordan fungerer semantisk søgning?

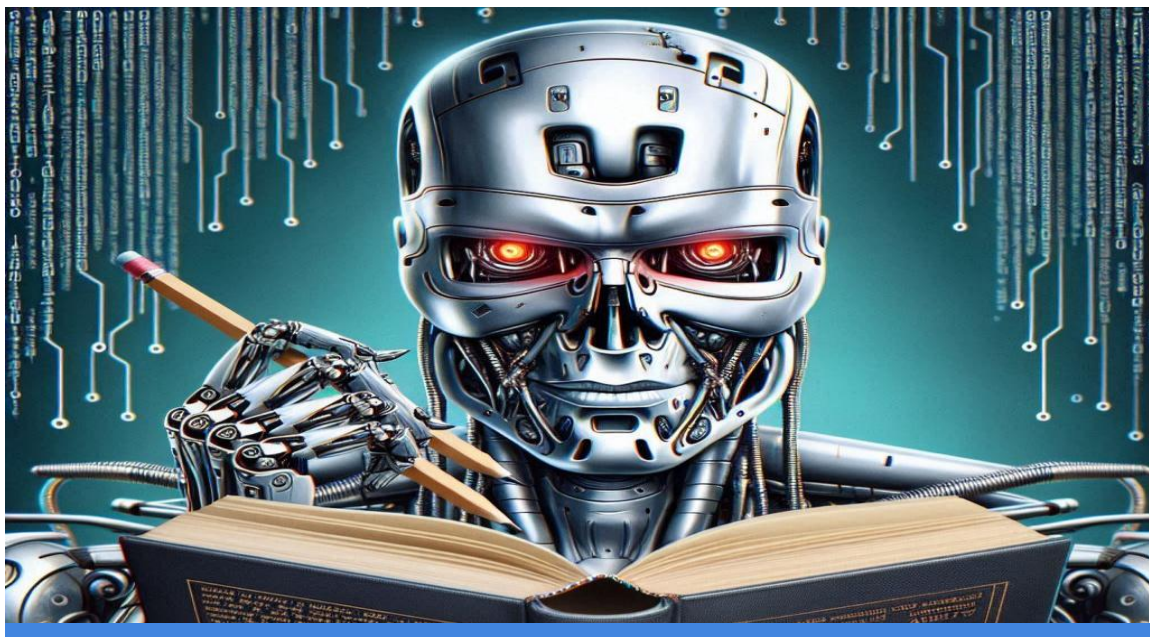
Processen består af flere trin:
Brugerforespørgsel: Brugeren indtaster en søgning i naturligt sprog eller nøgleord.
Fortolkning af forespørgslen: NLP-teknikker analyserer teksten for at identificere nøgleord, synonymer og den overordnede intention.
Kontekstuel analyse: Yderligere data, såsom brugerens geografiske placering, tidligere søgehistorik og adfærd, vurderes for at forbedre relevansen.

Vektorrepræsentation: Både forespørgslen og dokumenterne konverteres til numeriske vektorer i et højdimensionelt rum.

Lighedsmåling: Ved hjælp af metoder som cosinuslighhed vurderes, hvor tæt brugerens forespørgsel matcher de indekserede dokumenter.

Resultatlevering: De mest relevante resultater præsenteres for brugeren, prioriteret efter relevans og kontekst.

Denne tilgang muliggør mere præcise, personlige og naturlige søgeresultater, hvilket markant forbedrer brugeroplevelsen.



Generering af dokumenter
Rapporter, præsentationer

Generativ AI er begyndt at finde vej ind i produktivitetsværktøjer, hvilket muliggør skabelse, manipulation og forbedring af dokumenter på tværs af forskellige arbejdsprocesser. Disse værktøjer anvender avancerede dybe neurale netværk og generative modeller for at forbedre både effektiviteten og kvaliteten af dokumentproduktionen.

JAN ENGELBRECHT PEDERSEN

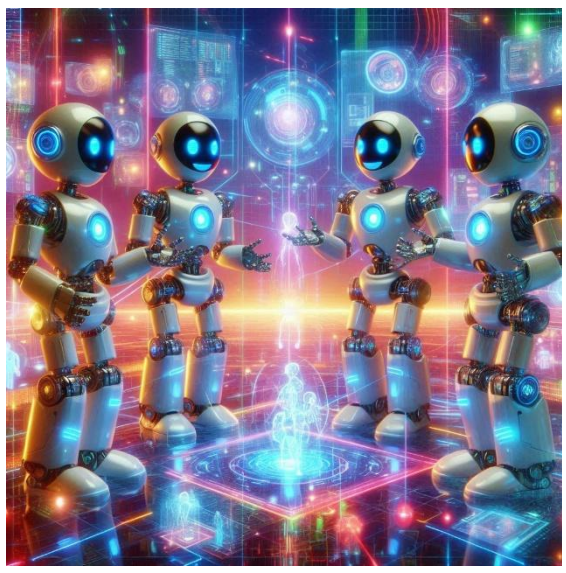
Rapportgenerering : AI kan hjælpe med at skabe rapporter ved automatisk at udtrække, analysere og opsummere data fra forskellige kilder. Ved at bruge teknikker inden for naturlig sprogbehandling (NLP) kan AI generere strukturerede rapporter, der præsenterer komplekse data på en forståelig måde. Dette sparer tid og reducerer fejl i manuel dataopsætning.

Præsentationsgenerering: Værktøjer som Microsoft CoPilot, der er integreret i PowerPoint, kan generere udkast til præsentationer baseret på tekstinput, foreslå designlayouts, indsætte billeder og ikoner samt hjælpe med at organisere slides. CoPilot anvender GPT-4-baseret generativ AI til at skabe indhold, der er skræddersyet til brugerens behov, og kan automatisk omorganisere slides, tilføje noter og visualisere data.

Dokumentskabelse: AI kan assistere med at udarbejde udkast til forskellige dokumenttyper, såsom forretningsplaner, juridiske kontrakter, rapporter og breve, baseret på brugerens input og eksisterende skabeloner. Ved at analysere eksisterende dokumenter kan AI foreslå strukturer og formuleringer, hvilket øger både produktiviteten og kvaliteten.

Gemini (Google) : Google Gemini er en avanceret AI-model, der er integreret direkte i Google Workspace (Docs, Sheets, Slides). Den tilbyder funktioner som tekstgenerering, opsummering, brainstorming og dataanalyse, hvilket muliggør en mere effektiv arbejdsflow. Gemini kan generere indhold, forbedre tekst og hjælpe med at strukturere komplekse dokumenter uden at forlade Google-økosystemet.

Microsoft CoPilot: Microsoft CoPilot, som beskrevet i Kapitel 1, er dybt integreret i Microsoft 365-applikationerne (Word, Excel, PowerPoint, Outlook). Det anvender GPT-4-baseret generativ AI til at assistere med skrivning, dataanalyse, præsentationsskabelse og automatisering af rutineopgaver. CoPilot kan generere tekst, foreslå layouts, organisere slides, indsætte billeder og endda hjælpe med at forbedre indholdets design og struktur. Det kan også importere data fra Excel, generere diagrammer og samarbejde i realtid via Teams.



Eksempler på funktioner og anvendelser
Fremtidsperspektiver

JAN ENGELBRECHT PEDERSEN

Rapportgenerering: AI kan udtrække og opsummere data fra flere kilder og præsentere dem i strukturerede rapporter, hvilket sparer tid og forbedrer nøjagtigheden.

PowerPoint-præsentationer: AI kan skabe udkast, organisere slides, foreslå design, indsætte billeder og generere talernoter, hvilket gør det lettere at lave professionelle præsentationer.

Samarbejde: AI kan hjælpe med at integrere feedback, foreslå forbedringer og automatisere opdateringer i dokumenter, især når man arbejder i teams via Microsoft Teams eller Google Workspace.

Support for forskellige platforme :

Både Gemini og CoPilot tilbyder kompatibilitet på tværs af platforme. Gemini er fuldt integreret i Google Workspace og kan tilgås via browser, mobil og desktop. CoPilot fungerer problemfrit på tværs af Microsoft 365-apps, både online og offline, med integration til OneDrive og SharePoint for fælles adgang og samarbejde.

Fremtidsperspektiver :

Disse AI-integrationer har potentiale til betydeligt at øge produktiviteten, forbedre dokumentkvaliteten og reducere manuelt arbejde i kreative og administrative processer. De vil fortsætte med at udvikle sig med flere funktioner, bedre forståelse af brugerbehov og en mere intuitiv brugeroplevelse.

Andre innovative anvendelsesområder for AI

JAN ENGELBRECHT PEDERSEN

Personlig medicin : AI spiller en vigtig rolle i udviklingen af personlig medicin ved at generere skræddersyede behandlingsplaner baseret på individets genetiske profil, medicinske historie og livsstil. AI kan også designe nye lægemidler ved at simulere molekylære interaktioner, hvilket fremskynder udviklingen af effektive og målrettede behandlinger. Maskinlæring (ML) og dynamiske modeller bruges til at analysere store medicinske datasæt med høj præcision, hvilket forbedrer både diagnose og prognose.

Klimaforskning: AI bidrager væsentligt til klimaforskning ved at modellere komplekse klimaforandringer, analysere store mængder klimadata og udvikle strategier for bæredygtighed. Prediktive modeller og simuleringsværktøjer kan forudsige vejrmønstre, optimere energiforbrug og støtte beslutningstagning i klimaindsatser. AI hjælper med at forstå dynamikken i klimaændringer, forbedre vedvarende energikilder og vurdere risici for økosystemer og biodiversitet.

Finansiel modellering: I finanssektoren anvendes AI til at forudsige markedsudviklinger, opdage svindel og automatisere handelsstrategier. Dataanalyse og generative modeller som transformer-arkitekturer bruges til at analysere store datamængder, identificere skjulte mønstre og træffe mere præcise beslutninger. AI forbedrer risikovurdering, porteføljestyring og kundeservice gennem automatiserede, intelligente systemer.

Robotik: AI giver robotter evnen til at lære, tilpasse sig og udføre komplekse opgaver i dynamiske og uforudsigelige miljøer. Autonom navigation og mobilitet er centrale områder, hvor AI-algoritmer som SLAM (Simultaneous Localization and Mapping) gør det muligt for robotter at navigere i ukendte omgivelser uden menneskelig indgriben. AI forbedrer også robotters evne til at håndtere uventede situationer, hvilket er afgørende i logistik, produktion og serviceindustrien.

Udover de tidligere nævnte anvendelsesområder udforskes generativ AI i en række banebrydende og innovative sektorer, hvor teknologien hjælper med at løse komplekse udfordringer, skabe nye muligheder og fremme bæredygtig udvikling.

Materialevidenskab: AI assisterer i designet af nye materialer med specifikke egenskaber til industrielle anvendelser, såsom letvægtsmaterialer, højtemperaturlegeringer eller biokompatible stoffer. Generative modeller og simuleringsværktøjer kan forudsige materialers strukturer og egenskaber, hvilket forkorter udviklingscyklussen og åbner nye muligheder for innovation.

Udvikling af bæredygtige løsninger gennem kunstig intelligens.

AI spiller en afgørende rolle i at skabe bæredygtige løsninger ved at optimere energiforbrug, reducere CO2-udledning og forbedre ressourceudnyttelse. Eksempler inkluderer:

Energieffektivisering: AI kan styre energiforbruget i industri og hjem gennem prediktiv vedligeholdelse og automatiserede styringssystemer.
Vedvarende energikilder: AI optimerer driften af sol- og vindkraftanlæg gennem vejrudsigter og produktionsovervågning.

Generativ AI

Nye roller

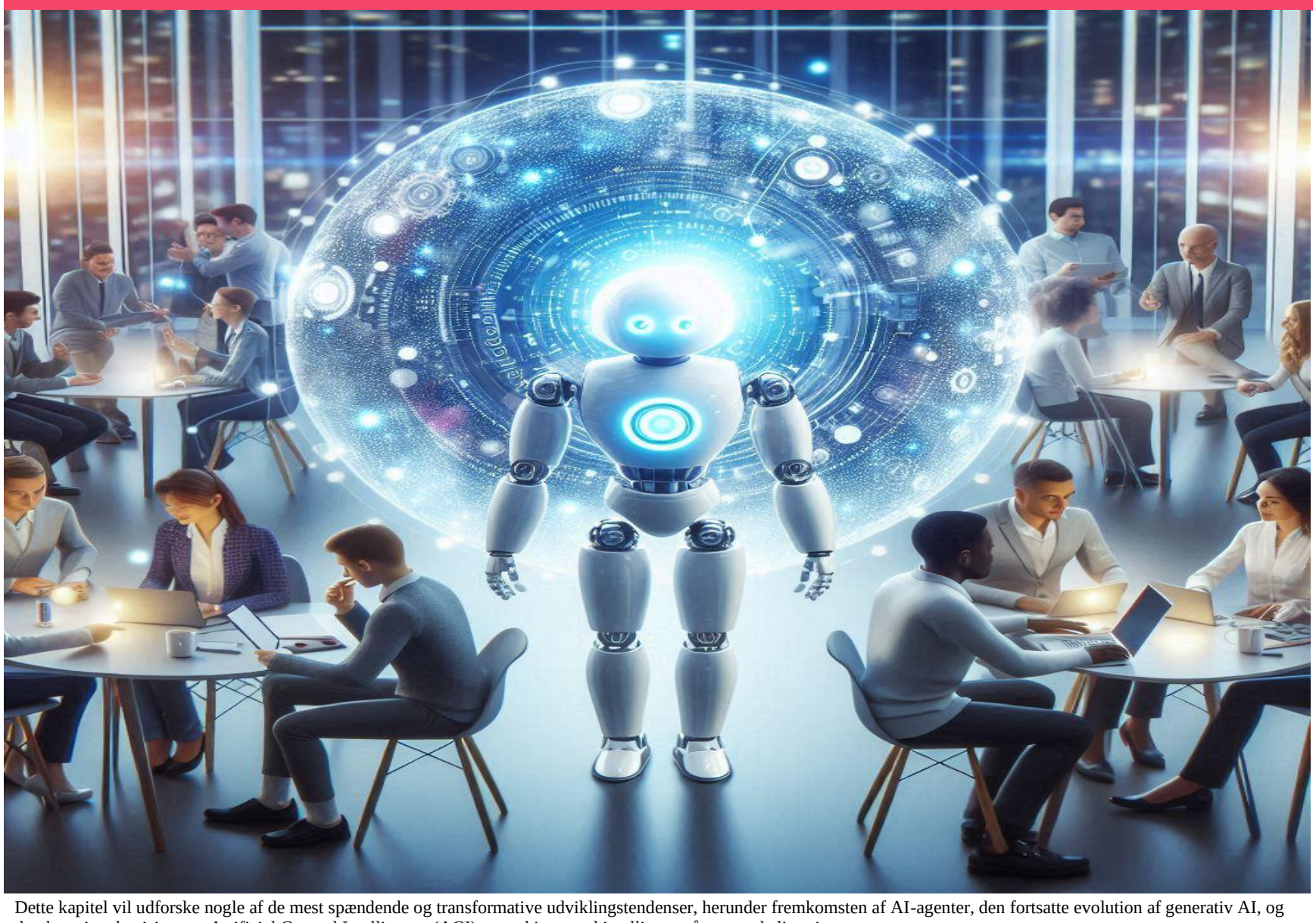
Jan Engelbrecht Pedersen

Intelligente transportsystemer: AI kan reducere CO2-udslip ved at planlægge mere effektive ruter, optimere trafikflow og støtte elektrificering. Ved at analysere trafikmønstre i realtid kan AI identificere de mest brændstofeffektive ruter, hvilket mindsker emissioner. Desuden kan AI styre trafiksignaler dynamisk for at reducere propper og forbedre flowet. Elektrificering understøttes gennem optimeret opladningsinfrastruktur.

Overvågning af biodiversitet: AI analyserer sensor- og satellitdata for at overvåge økosystemer, identificere truede arter og støtte biodiversitetsbeskyttelse. Ved hjælp af avanceret dataanalyse kan AI opdage ændringer i økosystemer i realtid, hvilket gør det muligt at reagere hurtigt og effektivt. Dette bidrager til at beskytte biodiversiteten og bevare truede arter.

Klimamodeller: AI forbedrer præcisionen i klimaforudsigelser og hjælper beslutningstagere med at udvikle effektive tiltag mod klimaforandringer. Ved at analysere omfattende datasæt kan AI identificere mønstre og tendenser, hvilket gør det muligt at forudsige klimaforandringer med større nøjagtighed. Dette understøtter udviklingen af strategier, der kan afbøde de negative effekter af klimaforandringer.

AI's bidrag til klimaforståelse og bæredygtighed:
AI har transformeret klimaforskningen ved at muliggøre mere præcise modeller, hurtigere dataanalyse og bedre forståelse af komplekse systemer.
Gennem praktiske case-studier har AI vist sig at kunne forudsige vejrmønstre, simulere klimascenarier og støtte beslutningstagning for at håndtere klimaudfordringer. På trods af etiske og sociale overvejelser er AI's potentiale til at bidrage til en mere bæredygtig fremtid uomgængeligt.



Dette kapitel vil udforske nogle af de mest spændende og transformative udviklingstendenser, herunder fremkomsten af AI-agenter, den fortsatte evolution af generativ AI, og den langsigtede vision om Artificial General Intelligence (AGI) – maskiner med intelligens på menneskeligt niveau.

AI-agenter

Den næste evolution inden for AI

JAN ENGELBRECHT PEDERSEN

En af de mest fascinerende fremskridt inden for kunstig intelligens (AI) er udviklingen af AI-agenter. Mens mange nuværende AI-systemer er designet til at udføre specifikke opgaver baseret på direkte input fra brugeren, er AI-agenter konstrueret som mere autonome enheder, der har evnen til at opfatte deres omgivelser, træffe beslutninger og handle for at opnå bestemte mål uden behov for konstant menneskelig indgriben.

Autonome opgaver og beslutningstagning: AI-agenter er designet til at være proaktive, selvstyrende og i stand til at håndtere komplekse opgaver uden vedvarende menneskelig indblanding. For at opnå dette kræves det, at de kan:

Sanse eget miljø: Modtage og fortolke information: Agenter anvender sensorer, datafeeds eller digitale input til at opfatte deres omgivelser. Dette kan inkludere visuelle data, lyd, tekst, sensorinformation eller andre datatyper, afhængigt af den specifikke anvendelse.

Anvende naturlig sprogforståelse (NLU) og computer vision: Dette er nødvendigt for at kunne analysere og forstå komplekse miljøer.

Definere og prioritere: Agenter kan fastlægge mål baseret på deres design, brugerinput eller kontekstuelle oplysninger.

Målstyring: De er i stand til at håndtere flere mål samtidigt og justere prioriteringer efter behov.

Udvikle strategier: Agenter benytter planlægningsalgoritmer og beslutningstræer til at formulere handlingsplaner, der skal føre til opnåelse af de fastsatte mål.

Forudse konsekvenser: Ved at anvende simuleringer og forudsigelsesmodeller kan de vurdere, hvilke handlinger der har størst sandsynlighed for at føre til succes.

Vælge de mest hensigtsmæssige handlinger: Baseret på perception, planlægning og målsætning anvender agenten beslutningsteorier som markov-beslutningsprocesser (MDP'er), reinforcement learning (RL) eller heuristikker til at træffe valg om handlinger.

Vægtning af usikkerhed: Agenter kan håndtere usikkerhed og risiko ved hjælp af probabilistiske modeller.

Evaluere resultater: Agenter vurderer konsekvenserne af deres handlinger ved hjælp af feedback og belønningssystemer.

Justere adfærd: Gennem maskinlæring og forstærkningslæring (RL) kan de forbedre deres præstation over tid og tilpasse sig ændringer i deres miljø.

Fremtiden inden for AI er fyldt med både enormt potentiale og betydelige usikkerheder, og en forståelse af de mulige veje forude er afgørende for at kunne navigere i denne dynamiske udvikling.

Forskellen mellem autonome AI-agenter og traditionelle AI-systemer:

Traditionelle AI-systemer kræver ofte eksplicitte instruktioner for hver enkelt handling og er begrænset til specifikke, veldefinerede opgaver.

AI-agenter er i stand til at håndtere mere komplekse, åbne og langsigtede opgaver, hvor de selv skal planlægge, træffe beslutninger og lære undervejs. Dette gør dem velegnede til anvendelser som robotik, autonome køretøjer, intelligente assistenter og selvstyrende systemer.

Denne grad af autonomi muliggør, at AI-agenter kan operere i uforudsigelige og dynamiske miljøer, hvor de skal træffe beslutninger uden konstant menneskelig overvågning, og hvor de skal tilpasse sig nye situationer for effektivt at opnå deres mål.

Eksempler på AI Agenter

Du kan lave AI-agenter med simple midler ved at bruge programmeringssprog som Python eller JavaScript. Kombinér open-source biblioteker som TensorFlow eller Hugging Face med API'er som OpenAI's GPT. Brug HTML og CSS til brugergrænsefladen.

JAN ENGELBRECHT PEDERSEN

Implementering af en AI-agent afhænger af dens formål og anvendelsesområde. Generelt følger processen disse trin:

1. **Definér formålet**: Bestem, hvad AI-agenten skal gøre, f.eks. kundeservice, dataanalyse eller automatisering.
2. **Vælg en AI-model**: Dette kan være en generativ AI som ChatGPT, hvis agenten skal forstå og generere naturligt sprog.
3. **Træning og finjustering**: Brug datasæt til at træne modellen, så den kan udføre specifikke opgaver.
4. **Integration**: Implementér AI-agenten i en applikation eller platform, f.eks. en chatbot eller et softwareværktøj.
5. **Test og optimering**: Test agenten for at sikre, at den fungerer korrekt, og justér efter behov.

Generative AI som ChatGPT kan bruges som backend, især hvis agenten skal interagere med brugere via tekst eller tale. Modellen kan integreres via API'er, hvilket gør det muligt at sende forespørgsler og modtage svar i realtid.

Her er nogle eksempler på simple AI-agenter og hvordan de kan implementeres:

Chatbot til kundeservice:
Implementering: Kan bygges med JavaScript ved hjælp af OpenAI's API. Brug fetch-funktion til at sende forespørgsler til API'en og modtage svar.
Generativ AI: OpenAI GPT-modeller via API-adgang.
Anvendelse: Besvarer kundespørgsmål og giver automatiserede løsninger.

Automatisk tekstoversætter:

Implementering: Kan udvikles i Python med Google Translate API.
Generativ AI: Google Translate AI.
Anvendelse: Oversætter tekst mellem forskellige sprog.

Enkel stemmeassistent:

Implementering: Kan bygges med C++ og integreres med IBM Watson API.
Generativ AI: IBM Watson Assistant.
Anvendelse: Udfører simple stemmekommandoer som at starte en timer eller afspille musik.

Disse agenter bruger API'er til at interagere med generative AI-modeller, hvilket gør det muligt at integrere avancerede funktioner uden at skulle bygge modellerne fra bunden.

Praktiske eksempler
AI agenter

JAN ENGELBRECHT PEDERSEN

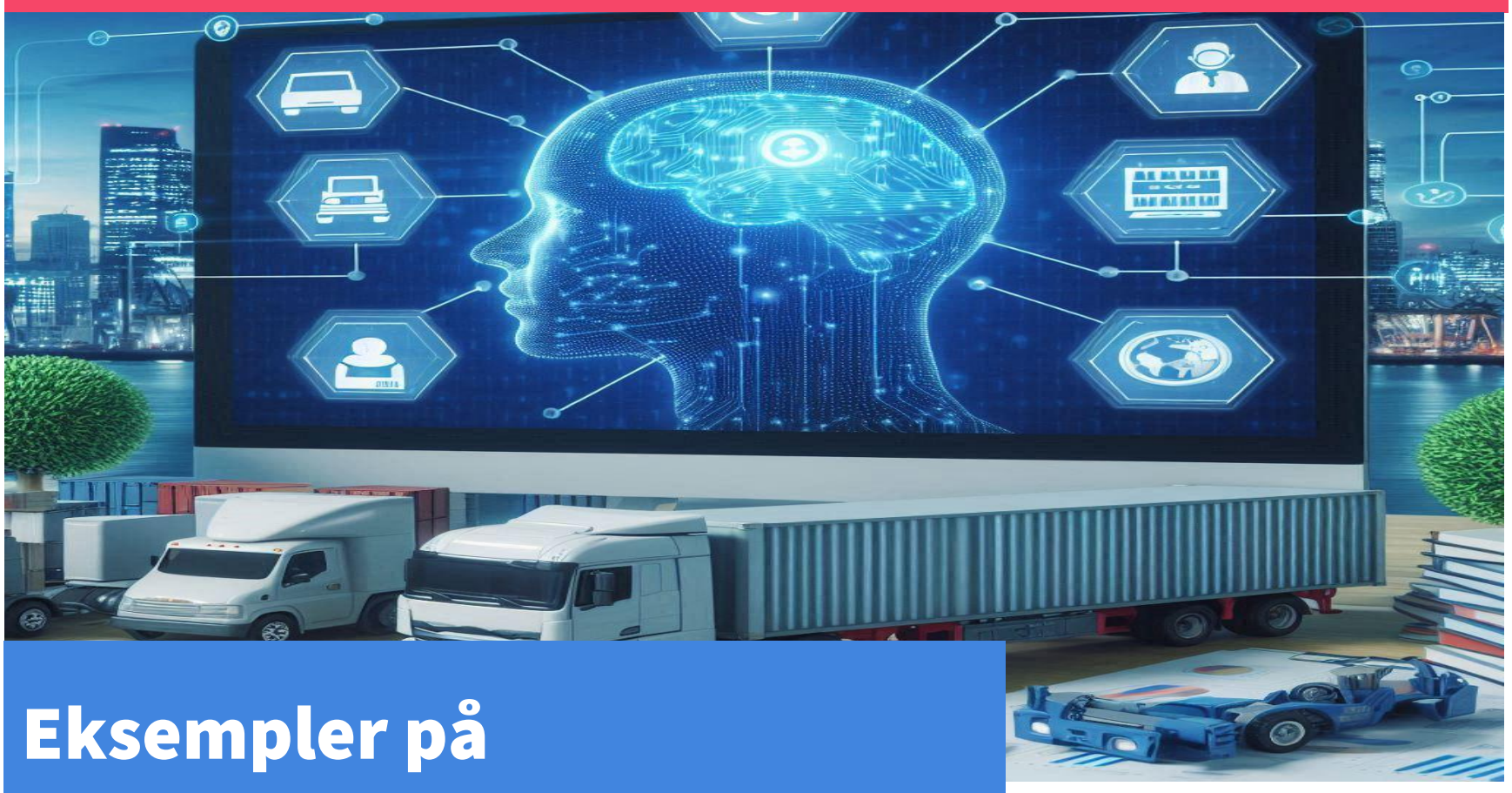
Et godt eksempel på en AI-agent bygget med JavaScript, HTML5 og CSS, der bruger ChatGPT som backend, er en chatbot-applikation. Denne type applikation kan integrere OpenAI's API for at generere intelligente og kontekstuelle svar baseret på brugerinput.

Du kan finde en detaljeret vejledning til at bygge en sådan chatbot her og her. Disse projekter demonstrerer, hvordan HTML og CSS bruges til at skabe en brugervenlig grænseflade, mens JavaScript håndterer interaktivitet og API-forespørgsler. Chatbotten kan give realtidsrespons og tilpasses til forskellige anvendelser, såsom kundeservice eller personlig assistance.

AI-agenter forenkler opgaver ved at automatisere gentagne processer, analysere data og levere intelligente løsninger. Svar fra generative AI'er via API'er kan efterbearbejdes ved at tilpasse dem til specifikke behov, filtrere irrelevante oplysninger og integrere dem med andre systemer for at optimere præcision og anvendelighed. Dette skaber effektive og skræddersyede løsninger.

AI-agenter er nemme at lave med grundlæggende programmeringsfærdigheder og adgang til API'er som ChatGPT. Ved hjælp af værktøjer som JavaScript, HTML og CSS kan man bygge funktionelle agenter. Generative AI'er via API'er gør det muligt at integrere avancerede funktioner uden kompleks udvikling, hvilket gør processen tilgængelig for de fleste.

AI agenter - Fremtidens AI



Eksempler på AI-agenter

JAN ENGELBRECHT PEDERSEN

Personlige assistenter: Nutidens virtuelle assistenter, såsom Siri, Alexa og Google Assistant, kan betragtes som grundlæggende AI-agenter. De er i stand til at udføre opgaver som at sætte påmindelser, afspille musik, besvare spørgsmål og styre smart home-enheder. Forventningerne til fremtidige versioner inkluderer en øget proaktivitet samt evnen til autonomt at håndtere mere komplekse opgaver, eksempelvis at planlægge møder, foretage indkøb eller koordinere flere tjenester.

Autonome køretøjer: Selvkørende biler repræsenterer avancerede AI-agenter, der opfatter deres omgivelser gennem en række sensorer, herunder kameraer, radar og lidar. Disse køretøjer er i stand til at planlægge ruter, træffe beslutninger om kørsel i realtid og handle autonomt for at nå deres destination på en sikker måde. Google Clouds Automotive AI-agent, som er baseret på Gemini-modellen, er et eksempel på en intelligent platform, der muliggør naturlig sprogforståelse og multimodal interaktion i biler, hvilket forbedrer både brugeroplevelsen og sikkerheden.

Robotter i industri og logistik: AI-drevne robotter anvendes i produktions- og lagerstyring til at udføre komplekse opgaver som samlebåndsarbejde, pakkehåndtering og levering. Disse agenter har evnen til at tilpasse sig ændringer i deres miljø og optimere processer autonomt, hvilket resulterer i øget effektivitet og reducerede fejl.

Softwareagenter: Softwarebaserede AI-agenter kan automatisere IT-relaterede opgaver som netværksovervågning, cybersikkerhed, systemoptimering og endda kodegenerering. Eksempler på sådanne agenter inkluderer AI-værktøjer, der kan identificere fejl i software, optimere ydeevne eller autonomt generere kode, hvilket bidrager til øget produktivitet i udviklingsmiljøer.

Finansielle agenter: Inden for finanssektoren anvendes AI-agenter til at analysere markedsdata, forudsige prisbevægelser, træffe investeringsbeslutninger og administrere porteføljer. Disse agenter arbejder med realtidsdata og komplekse algoritmer for at optimere afkast og minimere risiko.

Videnskabelige agenter: AI-agenter kan designe eksperimenter, analysere store mængder forskningsdata og endda formulere nye hypoteser. Denne kapacitet kan accelerere videnskabelige opdagelser og forbedre præcisionen i forskning inden for områder som medicin, fysik og bioteknologi.

Typer af AI-agenter og deres funktioner :

Kundeagenter: Disse agenter automatiserer kundeservice og support ved at kommunikere med brugere i naturligt sprog.
Hierarkiske agenter: Disse agenter organiserer opgaver på flere niveauer, hvor lavniveau-agenter håndterer specifikke opgaver, mens højere niveauer koordinerer bredere processer.
Dataagenter: Disse agenter udfører kompleks databehandling, analyse og informationstilgængelighed.
Ansatte agenter: Disse agenter automatiserer HR- og administrative opgaver, såsom onboarding, skemalægning og dokumenthåndtering, hvilket bidrager til at øge medarbejdernes produktivitet.
Lærende agenter: Disse agenter tilpasser deres adfærd baseret på tidligere erfaringer ved hjælp af maskinlæring. De anvendes eksempelvis i e-handel til at anbefale produkter og tilpasse annoncer.

Fremtidens AI-agenter

Transport & logistik

Jan Engelbrecht Pedersen

Fremtidens AI-agenter i transport og logistik revolutionerer sektoren ved at integrere avancerede teknologier som generativ AI og maskinlæring. De optimerer ruteplanlægning ved at analysere trafikdata i realtid og foreslå brændstofeffektive ruter, hvilket reducerer CO2-udledning.

Laststyring forbedres gennem automatiserede systemer, der balancerer vægt og volumen for maksimal effektivitet.

Lageradministration bliver smartere med AI-drevne algoritmer, der forudsiger efterspørgsel og optimerer vareplacering, hvilket mindsker spild og øger produktiviteten.

AI-agenter kan også minimere tomgangskørsel ved at koordinere køretøjer og reducere unødvendige ture.

Automatiseret overvågning og vedligeholdelse sikrer, at udstyr fungerer optimalt, hvilket reducerer driftsomkostninger og miljøpåvirkning.

Desuden kan AI forudsige markedstendenser og hjælpe med strategisk planlægning, hvilket gør transport og logistik mere bæredygtig og effektiv.

Samlet set er AI-agenter en nøglefaktor i fremtidens intelligente og grønne logistiksystemer.

Artificial General Intelligence (AGI)

Drømmen om menneskelignende intelligens

JAN ENGELBRECHT PEDERSEN

AGI's karakteristika og potentiale:
Alsidighed: Kunstig generel intelligens (AGI) har evnen til at udføre enhver intellektuel opgave, som et menneske er i stand til, hvilket omfatter komplekse problemløsninger, abstrakt tænkning samt en dyb forståelse af følelser. Denne alsidighed gør AGI til en potentielt revolutionerende teknologi, der kan anvendes i mange forskellige sammenhænge og industrier.

Adaptiv læring: Kunstig generel intelligens (AGI) har evnen til at tilpasse sig nye situationer og håndtere ukendte opgaver uden at have modtaget forudgående træning inden for specifikke domæner. Denne fleksibilitet gør det muligt for AGI at navigere komplekse og varierende udfordringer i forskellige kontekster.

Kreativitet og intuition: Udover logisk ræsonnering har AGI potentiale til at demonstrere både kreativitet og intuition, hvilket giver den mulighed for at udvikle originale ideer og innovative løsninger.

Artificial General Intelligence (AGI), også kendt som stærk AI, er det langsigtede mål for mange AI-forskere.

Mulige konsekvenser af AGI: Realisationen af AGI vil være en af de mest transformative begivenheder i menneskehedens historie med potentielt enorme konsekvenser:

Positive aspekter ved kunstig generel intelligens (AGI) inkluderer dens evne til at tackle komplekse globale udfordringer, fremme videnskabelige gennembrud, øge produktiviteten og generere hidtil usete muligheder for økonomisk velstand og innovation. AGI har potentialet til at revolutionere måden, hvorpå vi løser presserende problemer, og kan dermed bidrage til en mere bæredygtig og innovativ fremtid.

Risici og bekymringer: Udviklingen af Artificial General Intelligence (AGI) medfører betydelige eksistentielle risici, som omfatter tab af kontrol, uforudsete konsekvenser samt komplekse etiske dilemmaer. Det er af største vigtighed at sikre, at AGI udvikles med fokus på ansvarlighed, sikkerhed og transparens. En sådan tilgang vil bidrage til at minimere potentielle trusler og fremme en etisk anvendelse af teknologien i fremtiden.

Status og udfordringer : Selvom der er gjort betydelige fremskridt inden for AI, er vi stadig langt fra at realisere ægte AGI.

De nuværende AI-systemer, herunder avancerede generative modeller som GPT-4, er specialiserede inden for snævre domæner og mangler den brede, fleksible og dybtgående forståelse, der kendetegner AGI.

De vigtigste udfordringer inkluderer:
Teknologiske barrierer: Udvikling af modeller, der kan generalisere på tværs af opgaver og lære effektivt fra begrænsede data.
Teoretiske spørgsmål: Forståelse af bevidsthed, intelligens og kognition i maskiner.
Etiske og samfundsmæssige aspekter: Sikring af sikkerhed, kontrol og ansvarlig anvendelse.

Etiske og samfundsmæssige konsekvenser
Fremtidens AI

JAN ENGELBRECHT PEDERSEN

Udviklingen af avancerede AI-systemer som selvkørende agenter og generativ AI vil få store etiske og samfundsmæssige konsekvenser. Samfundet skal aktivt håndtere udfordringer som:

- **Autonomi og kontrol**: Sikring af menneskelig kontrol via mekanismer som "human-in-the-loop".

- **Arbejdsløshed**: Automatisering kan skabe ulighed; uddannelse og opkvalificering er afgørende.

- **Bias**: AI kan forstærke fordomme; metoder til at minimere bias er nødvendige.

- **Privatliv**: Regulering og teknologier som differentieret privatliv beskytter data.

- **Sikkerhed**: Forebyggelse af misbrug kræver robuste foranstaltninger og aftaler.

- **Eksistentielle risici**: AGI kan udgøre trusler; forskning i sikkerhed og governance er kritisk.

Når AI-agenter bliver mere selvstændige, opstår spørgsmålet om, hvordan vi bevarer menneskelig kontrol. Hvordan sikrer vi, at AI handler i overensstemmelse med menneskelige værdier og normer? Det kræver udvikling af mekanismer som "human-in-the-loop", hvor mennesker kan overvåge, styre og om nødvendigt afbryde AI-systemers handlinger.

AI trænes på store datasæt, som kan indeholde eksisterende fordomme. Dette kan føre til diskrimination i eksempelvis ansættelser, kreditvurdering eller retspraksis. Det er vigtigt at udvikle metoder til at identificere og minimere bias i AI-modeller for at sikre retfærdighed.

Autonome AI-agenter, der indsamler og analyserer store datamængder, rejser spørgsmål om privatlivets fred. Risikoen for masseovervågning og datamisbrug kræver regulering, transparens og teknologiske løsninger som differentieret privatliv (en metode, der beskytter individuelle data i store datasæt ved at tilføje kontrolleret støj, så det bliver svært at identificere enkeltpersoner, mens de samlede data stadig kan analyseres effektivt) for at beskytte individets rettigheder.

Etiske og samfundsmæssige konsekvenser
Samfundsmæssige udfordringer

JAN ENGELBRECHT PEDERSEN

Sikkerhed og misbrug: Kraftfuld AI kan misbruges til skadelige formål som autonome våben, cyberkriminalitet eller misinformation. Det er nødvendigt med robuste sikkerhedsforanstaltninger, overvågning og internationale aftaler for at forhindre misbrug.

Eksistentielle risici (ved AGI): Hvis AGI realiseres, kan en superintelligent AI, der ikke er i overensstemmelse med menneskelige værdier, udgøre en eksistentiel trussel. Forskning i AI-sikkerhed og governance bliver derfor afgørende.

Håndtering af disse udfordringer kræver en tværfaglig og international indsats, hvor forskere, politikere, virksomheder og borgere samarbejder. Nøgleinitiativer inkluderer:

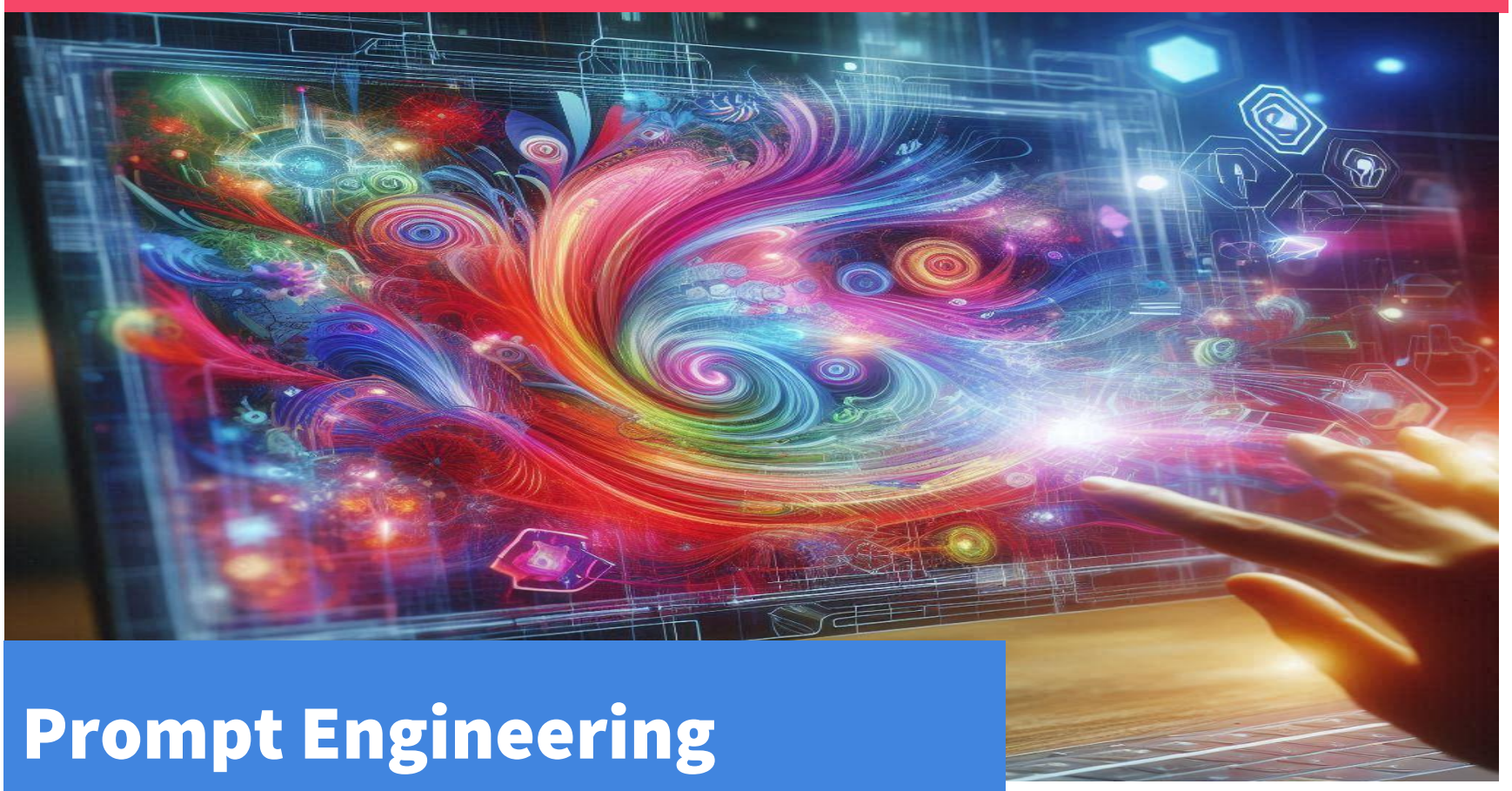
- **Etiske retningslinjer:** Udvikling af internationale og nationale standarder for ansvarlig AI-udvikling.

- **Regulering og lovgivning:** Skabelse af love, der beskytter privatliv, sikrer transparens og forhindrer diskrimination.

- **Tekniske sikkerhedsforanstaltninger:** Indbygning af kontrolmekanismer som fail-safe og audit trails i AI-systemer.

- **Offentlig dialog:** Åben debat om AI's rolle, så borgerne kan bidrage til teknologiens udvikling.

- **Uddannelse og opkvalificering**: Investering i uddannelse, så befolkningen kan tilpasse sig AI's muligheder og krav.



Prompt Engineering
Kunsten at kommunikere med AI

JAN ENGELBRECHT PEDERSEN

Kapitel 3 introducerede vigtigheden af prompt engineering, som udgør fundamentet for effektiv styring af generativ AI.

I dette kapitel vil vi gå i dybden med de avancerede metoder, der muliggør en optimal udnyttelse af AI-modeller ved hjælp af nøje udformede prompts.
Avanceret prompt engineering handler om at forstå og anvende de små detaljer, der forvandler en almindelig prompt til en fremragende og præcisionsorienteret kommando.
Det kræver en dybdegående indsigt i, hvordan forskellige teknikker kan benyttes til at generere resultater, der er både kreative, relevante og målrettede.

Kapitlet fokuserer på tre centrale aspekter:
- **De grundlæggende principper** for konstruktion af prompts: Herunder en systematisk tilgang til at skabe prompts, der er klart definerede og målrettede.
- **Ordvalg og sprogbrug**: Identifikation af nøgleord og formuleringer, som effektivt kan udløse specifikke handlinger i AI-systemer.
- **Avancerede teknikker**: Strategier til at finjustere prompts for at opnå større kontrol over den genererede output og sikre nøjagtighed i komplekse scenarier.

Formålet med denne gennemgang er at udstyre dig med en praktisk og fleksibel værktøjskasse af metoder og tilgange, der forbedrer din evne til at kommunikere med AI-modeller. Ved at anvende disse strategier kan du frigøre det fulde potentiale, som generativ AI tilbyder, og benytte det som et innovativt værktøj til problemløsning, beslutningstagning og kreativ udfoldelse. Kapitlet bygger på praktiske eksempler, der gør det let at forstå og anvende stoffet.

Prompt engineering er en central disciplin i arbejdet med avancerede AI-modeller, især de store sprogmodeller (LLM'er) som GPT-4, Bard/Gemini, Claude og LLaMA. Disciplinen handler om at formulere effektive prompts – de instruktioner og spørgsmål, vi stiller til AI – for at opnå de mest præcise, relevante og anvendelige resultater.

En velkonstrueret prompt er nøglen til at optimere AI's potentiale og sikre kvaliteten af dens output.

Dette kapitel giver et dybdegående indblik i følgende områder:

- **Grundlæggende principper:** Vi starter med fundamentet for, hvordan man strukturerer en klar og målrettet prompt, der guider AI til at levere de ønskede svar.

- **Avancerede teknikker:** Her undersøger vi metoder til at finjustere og forbedre prompts, der kan håndtere komplekse og specifikke opgaver, samt hvordan eksperimentering med prompts kan åbne op for nye kreative muligheder.

- **Etiske overvejelser**: Hvordan kan vi som brugere sikre, at prompts anvendes ansvarligt? Dette omfatter diskussioner om at undgå misbrug og om at maksimere fordelene ved AI, mens etiske retningslinjer overholdes.

Hvad er en prompt?
Det grundlæggende

Jan Engelbrecht Pedersen

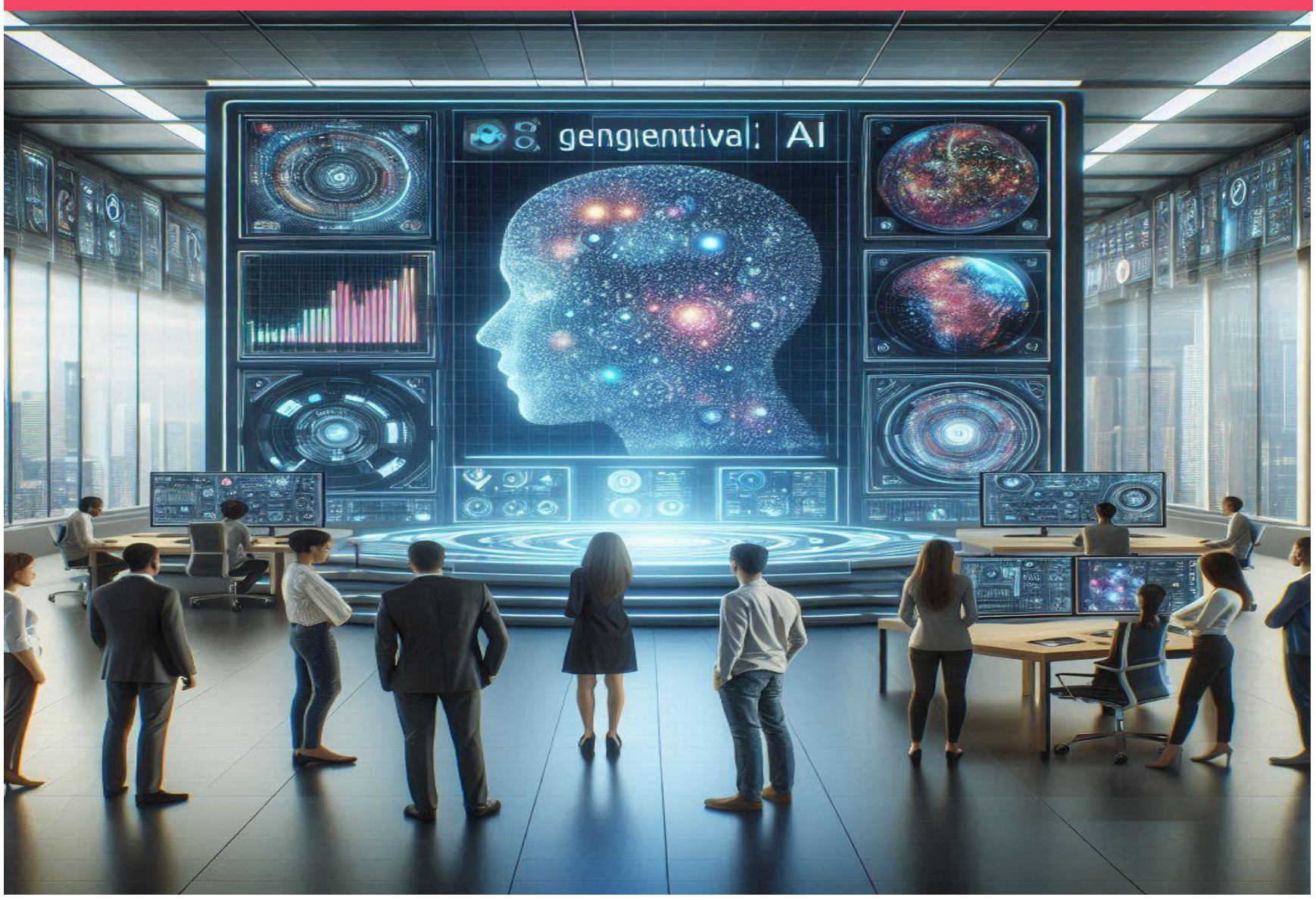
Formålet med dette kapitel er at udstyre dig med viden og værktøjer, der gør det muligt at arbejde målrettet med AI-modeller og deres unikke kapaciteter. Ved at mestre kunsten at formulere effektive prompts kan du ikke blot kommunikere med AI på en mere struktureret og produktiv måde, men også frigøre dens fulde potentiale som et redskab til innovation, problemløsning og kreativitet. Kapitlet er designet med en praktisk tilgang, så teorien bliver letforståelig og anvendelig.

Prompt engineering er afgørende for at opnå brugbare resultater fra AI. Effektive prompts giver klare instruktioner, kontekst og detaljer, hvilket sikrer relevante og præcise svar. Dårligt formulerede prompts kan føre til misforståelser og irrelevante output. Ekspertise i prompt design optimerer AI's funktionalitet og gør komplekse opgaver lettere at løse.

Prompt engineering er en tværfaglig disciplin, der kombinerer elementer fra datalogi, lingvistik, logik og filosofi. Det kræver teknisk forståelse fra datalogi, sproglig indsigt fra lingvistik og kritisk tænkning fra logik og filosofi. Denne kombination sikrer effektive prompts, der optimerer AI-modellers output og tilpasser dem til menneskelige behov.

En prompt er en klar tekstbesked, der bruges til at instruere eller stille spørgsmål til en AI-model. Den kan være enkel, som en sætning, eller kompleks med specifikke detaljer. Programmeringssprog som Python eller JavaScript kan også anvendes til at strukturere prompts, især når de skal integreres med API'er eller automatiseres.

En prompt laves ved at formulere en klar og præcis tekst, der beskriver opgaven eller spørgsmålet, AI'en skal besvare. Den kan være enkel, som en sætning, eller detaljeret med specifikke krav.



For at skabe en effektiv interaktion med AI-modeller er det afgørende at forstå og anvende tre fundamentale principper: klarhed, koncished og kontekst. Disse principper danner grundlaget for vellykkede prompts, der sikrer, at AI leverer nøjagtige og relevante resultater.

Grundprincipper for opbygning af prompts

Klarhed, koncished og kontekst

JAN ENGELBRECHT PEDERSEN

Klarhed: En tydelig og velstruktureret prompt er essentiel for at undgå misforståelser i kommunikationen med AI-modellen. Det betyder, at instruktionerne skal være entydige og lette at forstå. Undgå at bruge vagt sprog, komplekse formuleringer eller tvetydige udtryk, der kan forvirre AI. Klare prompts kræver specifikke detaljer omkring, hvad du ønsker, at AI skal gøre. For eksempel, hvis målet er at få et resumé, bør du præcist angive teksten, der skal opsummeres, samt hvor detaljeret eller omfattende resuméet skal være.

Koncished : Koncished i prompt konstruktion handler om at formulere en besked, der er kortfattet og præcis. Det betyder, at unødvendige detaljer fjernes, så AI'en kan fokusere på det væsentlige. En koncis prompt sikrer effektiv kommunikation og reducerer risikoen for misforståelser, hvilket gør det lettere for AI'en at levere relevante og brugbare svar.
En effektiv prompt indeholder kun de relevante oplysninger og undgår overflødige eller irrelevante ord.

Selvom det er vigtigt at inkludere kontekst, skal instruktionerne være korte og præcise for at gøre det lettere for AI at fokusere på de væsentlige elementer i anmodningen.

Lange og indviklede prompts kan mindske modellens evne til at levere målrettede svar.

Prompt engineering handler om at skabe klare, koncise og kontekstuelle instruktioner til AI for at optimere dens output og funktionalitet.

Kontekst: Kontekst er afgørende for at sikre, at AI forstår formålet med din anmodning og kan levere et meningsfuldt svar. Det indebærer at give relevant baggrundsinformation, definere formålet med interaktionen og inkludere nødvendige detaljer, som AI skal tage højde for. For eksempel, hvis du beder AI om at skrive en marketingtekst, bør du angive oplysninger om produktet, den tiltænkte målgruppe og ønsket tone for teksten. Disse elementer hjælper AI med at tilpasse sin respons, så den bliver så relevant som muligt.

Ved at følge principperne om klarhed, koncished og kontekst kan du optimere interaktionen med AI-modeller og øge sandsynligheden for at få præcis det output, du ønsker. En struktureret tilgang til promptopbygning gør det lettere for AI at forstå instruktionerne, hvilket skaber mere effektive og produktive resultater. Disse grundprincipper udgør derfor en afgørende del af arbejdet med at kommunikere med AI-modeller på en målrettet og succesfuld måde.

Konklusion: Grundprincipperne for opbygning af prompts er klarhed, koncished og kontekst. Klarhed sikrer, at AI forstår opgaven uden tvetydighed, hvilket fører til præcise svar. Koncished undgår unødvendige detaljer, så prompten bliver effektiv og fokuseret. Kontekst giver AI den nødvendige baggrundsinformation, så svarene bliver relevante og tilpasset opgaven. En velstruktureret prompt kan eksempelvis indeholde specifikke krav, målgruppe og ønsket format. Ved at kombinere disse principper kan man optimere AI's output og sikre, at komplekse opgaver løses korrekt. Effektiv prompt engineering er nøglen til at udnytte AI's fulde potentiale og opnå brugbare resultater.

Dansk fungerer godt som sprog i prompt konstruktion, da det er præcist og struktureret. Det danske sprog har klare grammatiske regler og et rigt ordforråd, hvilket gør det muligt at formulere detaljerede og kontekstuelle prompts. Sprogets evne til at udtrykke nuancer og relationer mellem begreber gør det ideelt til at skabe effektive og målrettede instruktioner for AI-modeller.

En effektiv prompt skal tydeligt afgrænse rækkevidden af det ønskede svar ved at specificere emnet og relevante detaljer. Den skal definere omfanget, f.eks. antal ord eller dybden af analysen. Det faglige niveau skal angives, om det er grundlæggende eller avanceret. Tonen skal være klar, f.eks. formel eller uformel.

Valget af sprogstil afhænger af opgavens krav og den ønskede respons fra AI'en. Akademisk sprog er formelt og præcist, ideelt til videnskabelig analyse. Hverdagssprog er tilgængeligt og letforståeligt, perfekt til brede målgrupper. Casual sprog er afslappet og uformelt, med humor og personlighed, velegnet til kreative interaktioner. Valg af tone afhænger af opgaven og den ønskede respons fra AI'en.

Brug af stil, tone, format og teknikker til optimering af prompts

Visse ord og fraser i dine prompts kan signalere specifikke handlinger eller outputformater til AI-modellen. Ved bevidst at bruge disse, kan du finjustere AI's adfærd og opnå mere præcise resultater.

JAN ENGELBRECHT PEDERSEN

Stil og tone: Når du arbejder med generativ AI, er evnen til at definere stil og tone afgørende for at sikre, at resultatet passer til konteksten. AI kan tilpasse sit sprog og udtryk efter brugerens behov ved hjælp af følgende tilgange:
- **Formel tone:** Brug prompts som "Skriv en professionel rapport om..." eller "Udfør en analyse i en akademisk tone...". Denne tilgang er særligt effektiv til rapporter, analyser og præsentationer, hvor præcision og professionalisme er nødvendige.
Eksempel: "Lav en detaljeret analyse af virksomhedens vækststrategi i en professionel og formel tone."

- **Uformel tone:** For afslappet og hverdagsagtig kommunikation kan du sige "Skriv det på en venlig måde..." eller "Hold det uformelt...". Denne tilgang er god til sociale medier, blogs eller samtaleorienteret indhold.
Eksempel: "Forklar fordelene ved sund kost på en sjov og uformel måde."

- **Humoristisk stil:** Humor kan skabe engagement og lethed i kommunikation. Prøv sætninger som "Gør det underholdende og humoristisk..." eller "Tilføj et sjovt twist til...".
Eksempel: "Beskriv livet som freelancer på en humoristisk måde, der også rummer ironi."

- **Overbevisende:** For at skabe argumenterende eller overtalende indhold kan du bruge "Overbevis læseren om, hvorfor..." eller "Argumenter for fordelene ved...".
Eksempel: "Overbevis en målgruppe om fordelene ved at implementere grøn energi i deres virksomhed."

- **Faktuel og objektiv**: Hvis du har brug for et neutralt og præcist svar, kan du sige "Skriv en objektiv vurdering af..." eller "Præsentér kun fakta uden holdninger."
Eksempel: "Opsummer historiske data om økonomisk vækst fra 2010 til 2020 i et faktuelt format."

- **Kreativ:** For indhold, der kræver fantasi, kan du bede AI om at "Skabe en original fortælling om..." eller "Udforme et kreativt koncept for...".Eksempel: "Skriv et eventyr om en opdagelsesrejsende, der finder en skjult verden."

- **Efterligning af forfatterstil:** Brug specifikke eksempler som "Skriv i Hemingway-stil..." eller "Efterlign Shakespeares poetiske sprog."Eksempel: "Forklar videnskabelige opdagelser i en stil, der minder om Carl Sagans formidling."

Stil refererer til sprogets form og struktur, som kan være formel, uformel, teknisk eller kreativ, afhængigt af opgaven. Tone angiver den følelsesmæssige nuance, såsom venlig, professionel eller neutral. Ved at definere stil og tone præcist i en prompt kan man sikre, at AI leverer svar, der passer til målgruppen og konteksten. Dette gør kommunikationen mere målrettet og relevant, hvilket optimerer AI's output og skaber en bedre brugeroplevelse. Stil og tone er nøglen til succesfuld prompt engineering.

Format og Struktur

JAN ENGELBRECHT PEDERSEN

Format og struktur spiller en vigtig rolle i, hvordan informationen præsenteres og organiseres:

- **Punktopstillinger:** For kortfattet og organiseret indhold kan du sige "Præsentér informationen i punktopstillinger."
Eksempel: "Oplist fordelene ved fjernarbejde i fem punkter."

- **Tabel:** Tabeller er nyttige til at præsentere data systematisk. Brug "Formater oplysningerne i en tabel."
Eksempel: "Lav en tabel, der sammenligner fordele og ulemper ved to forskellige energiformer."

- **Resumé:** Til kortfattet information kan du bede AI om at "Opsummere nøglepunkterne i..."
Eksempel: "Skriv et kort resumé af de vigtigste konklusioner fra en videnskabelig artikel."

- **Afsnitsinddeling:** For velstrukturerede tekster, angiv "Opdel informationen i klart definerede afsnit."
Eksempel: "Skriv en blogopdatering opdelt i tre afsnit – introduktion, diskussion og konklusion."

- **Trin-for-trin forklaringer:** Hvis du ønsker processer forklaret klart, brug "Forklar i trin."
Eksempel: "Beskriv i trin-for-trin-format, hvordan man starter en ny virksomhed."

Format og struktur i prompt engineering er afgørende for at sikre klare og effektive instruktioner til AI. Et godt format inkluderer en præcis beskrivelse af opgaven, specifikke krav og ønsket output. Strukturen bør være logisk og letforståelig, med en tydelig opdeling af information, såsom emne, kontekst og tone. Brug af punktlister eller afsnit kan hjælpe med at organisere komplekse prompts. En velstruktureret prompt reducerer risikoen for misforståelser og optimerer AI's evne til at levere relevante svar. Ved at fokusere på format og struktur kan man maksimere effektiviteten og kvaliteten af AI's output.

Kreativitet og analyse

JAN ENGELBRECHT PEDERSEN

For at fremme kreative løsninger og originale idéer kan du bruge følgende tilgange:

- **Brainstorming:** "Kom med flere ideer til..." eller "Brainstorm alternativer for..."
Eksempel: "Udvikl fem nye produktideer til teknologivirksomheder."

- **Innovative løsninger:** "Foreslå en innovativ tilgang til..." eller "Tænk kreativt omkring..."
Eksempel: "Foreslå innovative løsninger på affaldshåndtering i storbyer."

- **Originalt indhold:** "Skab noget, der aldrig er set før..." eller "Udform en ny tilgang til..."
Eksempel: "Lav et originalt forslag til at designe en bæredygtig skolebygning."

- **Metaforer og analogier:** "Forklar komplekse koncepter med analogier."
Eksempel: "Forklar blockchain-teknologi ved at sammenligne det med en kæde af hængelåse."

- **Udforskning:** "Udforsk alternative metoder til..."
Eksempel: "Undersøg forskellige måder, hvorpå kunst kan bruges i undervisning."

Når der kræves analytiske eller opsummerende svar, er disse prompts effektive:

- **Analyse:** "Identificér de centrale udfordringer i..." eller "Analyser nøgledata fra..."
Eksempel: "Analyser den finansielle performance af et firma baseret på det seneste kvartals data."

- **Sammenligning:** "Sammenlign alternativerne..." eller "Hvad er forskellene mellem to modeller?"
Eksempel: "Sammenlign effekten af to forskellige ledelsesstile på medarbejdermotivation."

- **Evaluering:** "Evaluer styrker og svagheder ved..."
Eksempel: "Vurder fordelene og ulemperne ved at anvende AI i medicinske diagnoser."

- **Konklusion**: "Hvad kan man konkludere ud fra..."
Eksempel: "Opsummer, hvad analyserne viser om tendenserne i klimaforandringer."

Instruktioner og Begrænsninger

Jan Engelbrecht Pedersen

Klare instruktioner og rammer styrker effektiviteten af AI’s output:

- **Specifikke instruktioner:** "Sørg for at inkludere..."
Eksempel: "Forklar med eksempler, hvordan teamarbejde forbedrer effektiviteten."

- **Begrænsninger:** "Undgå brug af komplekst sprog..."
Eksempel: "Skriv en introduktion om global opvarmning, men undgå teknisk jargon."

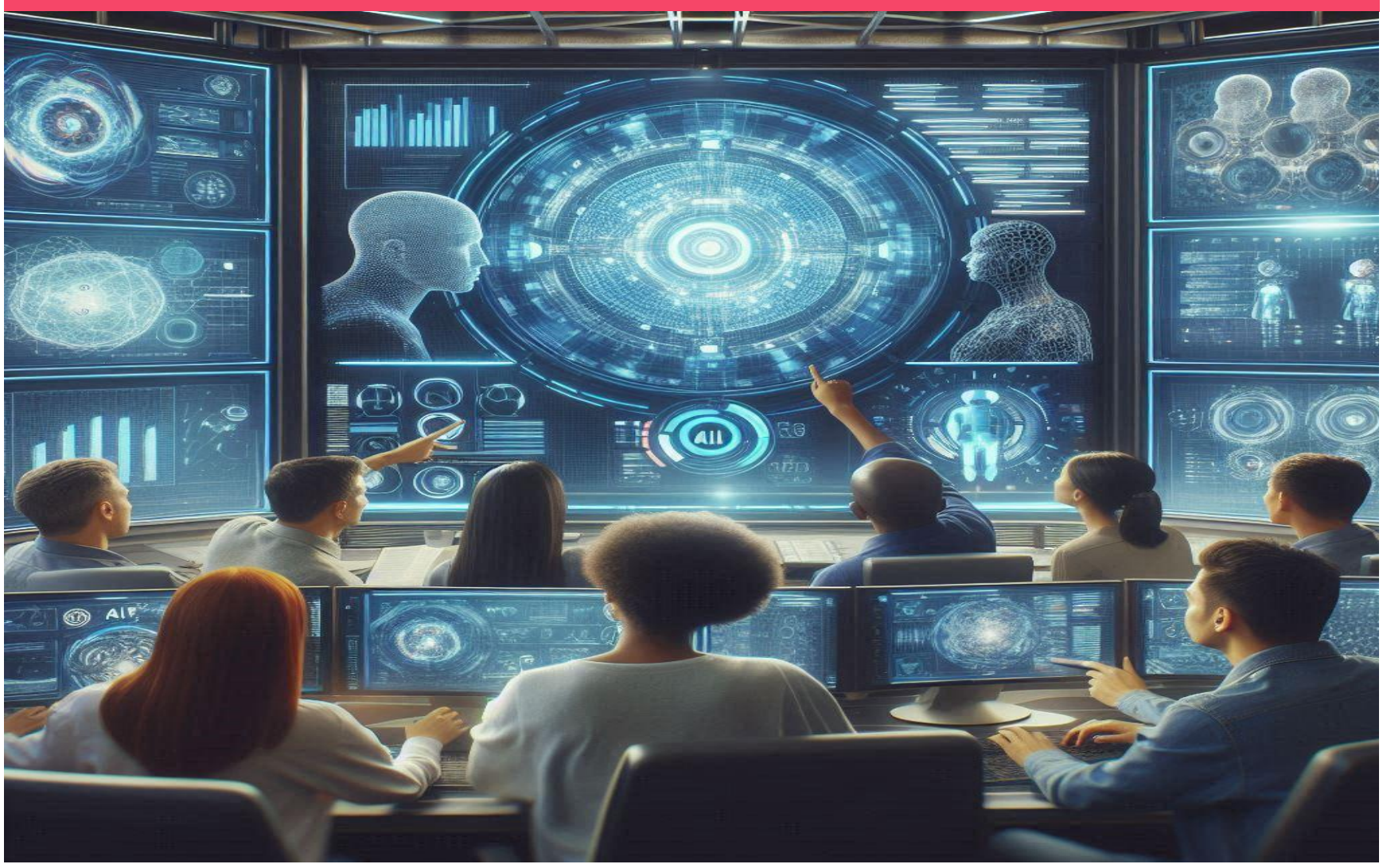
- **Roller:** "Forestil dig, at du er en ekspert i..."
Eksempel: "Agér som en marketingkonsulent og foreslå strategier for et nyt produkt."

Ved at mestre principperne for stil, tone, struktur, kreativitet og instruktioner kan du optimere AI's potentiale og sikre, at den leverer målrettet og effektivt output, der opfylder specifikke krav. Klarhed og præcision i disse aspekter forbedrer kommunikationen med AI, hvilket resulterer i mere relevante og brugbare svar.

Derudover giver det mulighed for at udforske nye kreative løsninger og anvendelser, som kan tilpasses forskellige behov og kontekster.

Denne tilgang fremmer alsidighed og innovation i brugen af AI-teknologi, hvilket gør den til et værdifuldt værktøj i både professionelle og personlige sammenhænge.

Effektiv prompt engineering er nøglen til at udnytte AI's fulde kapacitet.



Avanceret Prompting

JAN ENGELBRECHT PEDERSEN

Few-shot prompting involverer at give AI-modellen et lille antal eksempler på det ønskede input-output-format direkte i prompten. Dette hjælper modellen med at forstå opgaven mere præcist og tilpasse sin respons til forventningerne. I stedet for blot at forklare opgaven, demonstrerer du for AI, hvordan det ønskede output skal se ud. Teknikken fungerer som en form for praktisk "træning", der gør resultatet mere nøjagtigt og relevant.

Eksempel:
Oversæt følgende engelske sætninger til fransk:
"Hello, how are you?" → "Bonjour, comment ça va ?"
"Thank you very much." → "Merci beaucoup."
"What time is it?" →
Ved at inkludere eksempler på oversættelser guider du AI til at levere en korrekt og konsekvent oversættelse af den sidste sætning. Denne metode er særlig effektiv, når opgaven har flere nuancer eller specifikke krav.
Anvendelsesmuligheder:
- Skabelse af dataformater, f.eks. rapporter eller skabeloner.
- Oversættelser, der kræver kontekstspecifikke variationer.
- Generering af indhold med en bestemt stil eller struktur.

Avancerede teknikker inden for prompt engineering giver mulighed for at opnå endnu større kontrol over generativ AI-modeller og forbedre kvaliteten af deres output. Disse teknikker er særligt værdifulde i komplekse og specialiserede opgaver.

Chain-of-Thought Prompting

Med chain-of-thought prompting opfordrer du AI-modellen til at redegøre for sin tankegang trin for trin, før den giver det endelige svar. Denne teknik er særlig nyttig til opgaver, der kræver logisk tænkning, kompleks analyse eller ræsonnement over flere trin.

Eksempel:
Anders har 10 æbler. Han giver 3 til Sofie og spiser derefter 2. Hvor mange æbler har Anders tilbage? Lad os tænke trin for trin. AI vil strukturere sin respons som:
- *Anders starter med 10 æbler.*
- *Han giver 3 til Sofie, hvilket efterlader ham med 10 - 3 = 7 æbler.*
- *Han spiser 2 æbler, hvilket efterlader ham med 7 - 2 = 5 æbler.*

Denne teknik minimerer risikoen for fejltolkninger eller hurtige, upræcise svar og øger modellens evne til at levere velbegrundede resultater.

Few-shot prompting indebærer at give AI-modellen et par eksempler på input-output-par for at guide dens respons. Dette hjælper modellen med at forstå formatet og konteksten, hvilket forbedrer præcisionen i dens svar. Chain-of-Thought Prompting fokuserer på at bryde komplekse opgaver ned i mindre trin, så modellen kan følge en logisk ræsonnementskæde. Denne metode er særligt nyttig til opgaver, der kræver dybdegående analyse eller flere beslutningstrin. Begge teknikker optimerer AI's evne til at levere relevante og kontekstuelle svar, hvilket gør dem værdifulde i mange anvendelser.

Role-Playing , personas og iterativ prompt forbedring

JAN ENGELBRECHT PEDERSEN

Ved hjælp af **role-playing** kan du få AI til at indtage en bestemt rolle eller persona, der styrer den måde, den kommunikerer, vurderer og genererer information. Denne teknik er ideel til opgaver, der kræver kreativitet, simulation eller specifikke perspektiver.
Eksempel:

Du er en hjælpsom og entusiastisk rejseguide, der specialiserer sig i bæredygtig turisme i Thailand. Beskriv tre unikke og miljøvenlige aktiviteter, som turister kan opleve.

AI vil tage rollen som en rejseguide og levere information i en passende tone og fokusere på bæredygtige aktiviteter.

Andre anvendelsesmuligheder:

- Kundeservice simulationer.
- Indhold genereret fra historiske eller fiktive perspektiver.
- Ekspertvurderinger fra specifikke faglige roller.

Iterativ prompt forbedring indebærer en proces, hvor man løbende justerer og finpudser en prompt baseret på feedback fra AI. Dette sikrer, at output bliver mere præcist og relevant med hver iteration. Teknikken er ideel til komplekse opgaver, hvor små ændringer kan gøre en stor forskel.
Eksempel:

Første prompt: "Skriv en kort historie om en kat."

Output: En generisk historie om en kat.

Anden prompt: "Skriv en kort, spændende historie om en sort kat med grønne øjne, der bor i et hjemsøgt hus."

Output: En mere detaljeret og engagerende historie.

Fordele:

Gennem denne iterative proces kan du opnå mere tilfredsstillende og præcise resultater.

Iterativ prompt forbedring, role-playing og personas er centrale teknikker inden for prompt engineering, der optimerer AI's output og gør interaktionen mere målrettet. **Iterativ prompt forbedring** indebærer en proces, hvor man løbende justerer og finpudser en prompt baseret på feedback fra AI. Dette sikrer, at output bliver mere præcist og relevant med hver iteration. **Role-playing** giver AI en specifik rolle, som f.eks. "ekspertanalytiker" eller "venlig mentor," for at styre dens tone og tilgang. **Personas** bruges til at skabe en konsistent stil og tone i AI's svar. Ved at definere en persona, som AI skal efterligne, kan man sikre, at output er skræddersyet til specifikke behov, hvilket gør kommunikationen mere effektiv og målrettet.

Constraint Prompting , prompt- decomposition & Selv-konsistens prompting

JAN ENGELBRECHT PEDERSEN

Med **constraint prompting** sætter du klare regler eller begrænsninger, der styrer, hvad AI må inkludere eller udelade i sit output. Dette sikrer, at svarene forbliver inden for de ønskede rammer.

Eksempel: Skriv en artikel om fordelene ved elbiler, men undgå at nævne Tesla.
Anvendelsesmuligheder:
- Undgå bestemte emner eller stilarter.
- Fokusere på specifikke aspekter af en diskussion.
- Sikkerhed og etisk brug, såsom udeladelse af følsomme data.

Prompt Decomposition: Prompt decomposition indebærer at bryde en kompleks opgave ned i mindre, håndterbare delopgaver, som AI kan bearbejde individuelt. Dette forbedrer strukturering af output og sikrer, at alle aspekter af opgaven bliver behandlet.

Eksempel: I stedet for at bede om en komplet marketingplan, opdeler du opgaven i:
- Identificér målgruppen.
- Beskriv produktets vigtigste fordele.
- Brainstorm marketingkanaler.
- Udarbejd en reklamekampagne.

Selv-konsistens Prompting: Selv-konsistens prompting reducerer variationer og usikkerheder i AI's respons ved at generere flere svar på samme prompt og vælge det mest konsistente output.

Eksempel: Bed AI om at svare fem gange på spørgsmålet: "Hvad er hovedstaden i Australien?"

Hvis fire ud af fem svar er "Canberra," kan du med større sikkerhed acceptere det som korrekt.
Fordele:
- Filtrerer fejl og uregelmæssigheder.
- Øger nøjagtigheden i opgaver, der kræver klare svar.

Relationer og sammenligninger

Kommandoer til identifikation og beskrivelse af relationer

JAN ENGELBRECHT PEDERSEN

AI har evnen til at analysere, beskrive og udforske relationer mellem forskellige emner. Ved at bruge specifikke kommandoer kan du styre denne proces og sikre, at AI leverer organiserede og analytiske svar, der fremhæver sammenhænge, forskelle eller forbindelser. Dette afsnit gennemgår effektive kommandoer til at finde og beskrive relationer mellem emner, og inkluderer praktiske eksempler for at illustrere deres anvendelse.

Kommandoer og deres funktion:

- Sammenlign
Brug "sammenlign" til at anmode AI om at beskrive ligheder og forskelle mellem to eller flere emner. Denne kommando er ideel til analyser, der kræver en struktureret sammenstilling af data eller argumenter.
Eksempel:
"Sammenlign fordele og ulemper ved de to teknologier."
AI vil her fremhæve både positive og negative aspekter ved hver teknologi og præsentere dem i et overskueligt format.

- Kontrastér
Brug "kontrastér" for at fokusere specifikt på forskelle mellem to emner. Kommandoen er nyttig til opgaver, der kræver en distinkt adskillelse af karakteristika.
Eksempel:
"Kontrastér effektiviteten."
AI vil tydeliggøre de områder, hvor de to emner adskiller sig i deres resultater eller anvendelighed.

- Relatér
Når du vil opdage forbindelser eller sammenhænge mellem forskellige områder, er "relatér" den rette kommando. Denne tilgang hjælper med at fremhæve, hvordan emner eller begreber påvirker hinanden.
Eksempel:
"Relatér klimaforandringer til økonomisk udvikling."
AI vil levere en analyse af, hvordan klimaforandringer kan influere økonomiske faktorer såsom investeringer, vækst og globale politikker.

- Forklar sammenhængen
Brug denne kommando til at få AI til at beskrive de specifikke mekanismer, der binder to ting sammen. Det er ideelt til at forstå komplekse relationer eller interaktioner.
Eksempel:
"Forklar sammenhængen mellem søvn og læring."
AI vil forklare, hvordan søvn påvirker hukommelse, koncentration og kognitive evner, og levere en detaljeret respons.

- Associer
"Associer" bruges til at identificere og beskrive forbindelser eller karakteristika, der knytter forskellige emner sammen. Denne kommando er særlig nyttig inden for kreative eller konceptuelle opgaver.
Eksempel:
"Associer de forskellige kunstretninger med de historiske perioder, de opstod i."
AI vil relatere kunstretninger som renæssancen eller impressionismen til deres historiske kontekst og beskrive indflydelse og karakteristika.

Konklusion:

Jan Engelbrecht Pedersen

Fordele ved disse kommandoer:
- **Struktureret analyse:** AI leverer klart organiserede svar, der fremhæver både ligheder og forskelle.

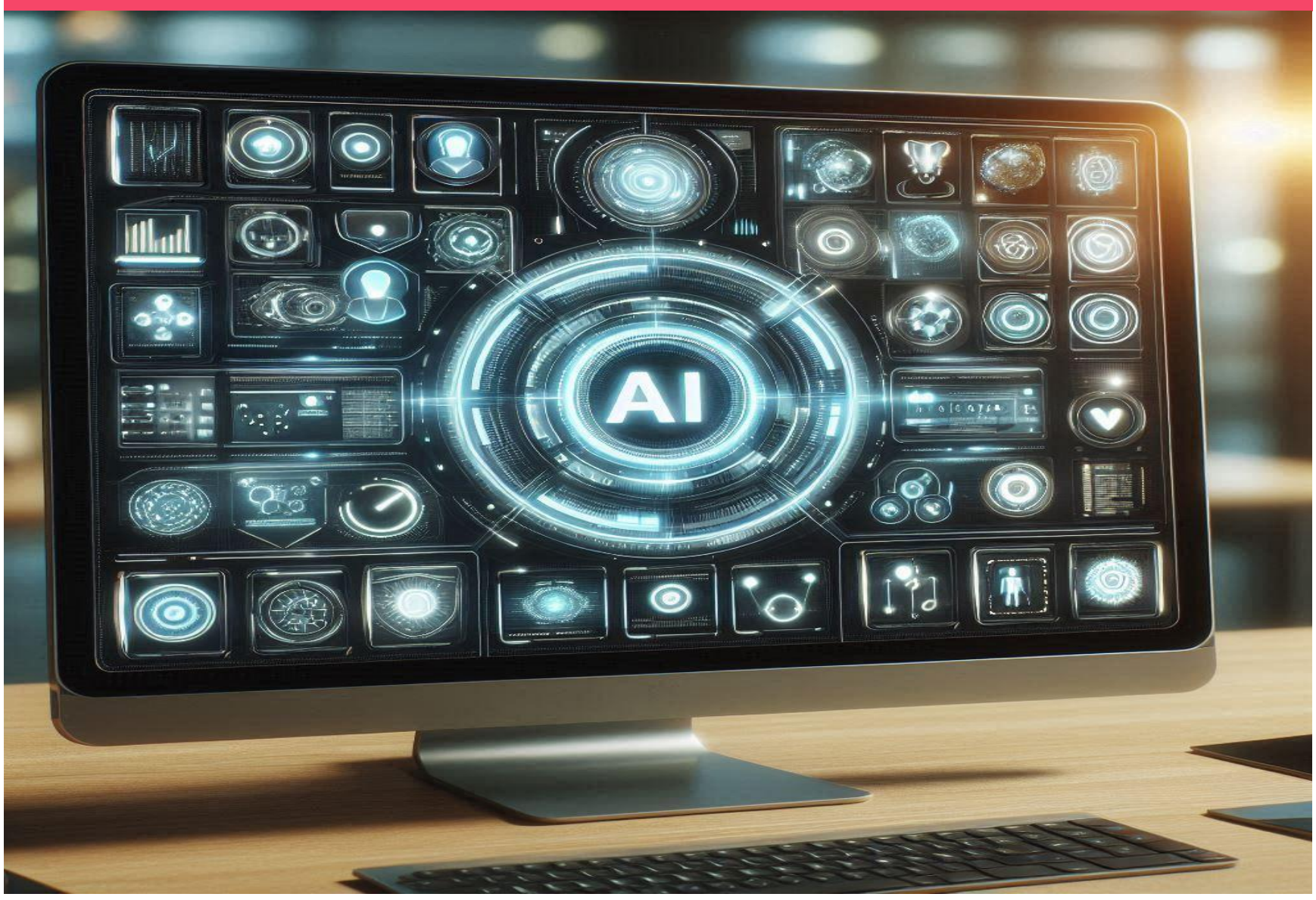
- **Fokuseret output:** Kommandoerne leder AI's opmærksomhed mod præcise aspekter af relationerne.

- **Styrket forståelse**: Forbindelser og interaktioner bliver forklaret på en letforståelig og detaljeret måde.

Anvendelsesmuligheder:
- **Akademisk**: Brug "sammenlign" og "forklar sammenhængen" til forskning og essays.
- **Professionelt**: Analyser markedsdata med "kontrastér" eller "relatér" til forretningsstrategier.
- **Kreativt**: Associer emner som kunst, litteratur eller historiske begivenheder for at skabe inspirerende og dybdegående beskrivelser.

Ved at anvende kommandoer som "sammenlign," "kontrastér," "relatér," "forklar sammenhængen" og "associer" kan du opnå målrettet og struktureret output, der klarlægger relationer mellem emner. Disse redskaber giver dig mulighed for at dykke dybere ned i analyser og sikre, at AI leverer svar, der er både præcise og relevante.

Brug dem strategisk for at maksimere AI's analytiske kapacitet og forståelse.



Logisk tænkning og ræsonnement er afgørende for at sikre, at AI leverer præcise, velbegrundede og analytiske svar. Gennem brug af specifikke kommandoer kan du styre AI til at undersøge, evaluere, analysere og argumentere om et givent emne. Denne tilgang sikrer, at outputtet er både relevant og velovervejet.

Logik og ræsonnement

Kommandoer til styring af AI's analytiske tænkning

JAN ENGELBRECHT PEDERSEN

Kommandoer og deres funktion:

- Argumentér
Brug "argumentér" til at bede AI om at fremlægge en struktureret argumentation for eller imod et specifikt synspunkt. Det er nyttigt i diskussioner, essays eller debatforberedelser.

Eksempel:
"Argumentér for fordelene ved en fire-dages arbejdsuge."
AI vil her levere velbegrundede argumenter, der fokuserer på aspekter som medarbejderproduktivitet, trivsel og økonomiske konsekvenser.

- Begrund
Når du ønsker en forklaring, der understøtter, hvorfor noget er på en bestemt måde, er "begrund" den ideelle kommando. Dette er anvendeligt i sammenhænge, hvor rationelle forklaringer er nødvendige.
Eksempel:
"Begrund nødvendigheden af øget investering i vedvarende energi."
AI vil give en detaljeret begrundelse, der inddrager faktorer som miljøfordele, ressourceknaphed og langsigtede økonomiske gevinster.

- Konkludér
Brug "konkludér" til at få AI til at drage en logisk og sammenhængende konklusion baseret på givne data eller oplysninger. Denne kommando sikrer et klart og afsluttende svar.
Eksempel:
"Konkludér på baggrund af data."
AI vil sammenfatte nøglepunkterne og præsentere en logisk konklusion, der passer til det undersøgte emne.

- Analysér
Med "analysér" kan du bede AI om at udføre en dybdegående undersøgelse af et emne. Det er særligt velegnet til komplekse problemstillinger, der kræver en grundig forståelse.
Eksempel:
"Analysér årsagerne til inflation."
AI vil præsentere en struktureret analyse, der inkluderer økonomiske faktorer som udbud, efterspørgsel, pengepolitik og globale påvirkninger.

Få analyser og argumenter til essays eller projektrapporter. Evaluer strategier, begrund investeringer eller konkludér på komplekse datasæt. Brug logiske analyser til at vurdere dagligdags valg eller større beslutninger.

- Evaluér
Brug "evaluér" til at få AI til at vurdere effektiviteten, styrkerne eller svaghederne ved et emne baseret på specifikke kriterier. Dette er nyttigt for beslutningstagning og vurderinger.
Eksempel:
"Evaluer effektiviteten af de forskellige strategier."
AI vil vurdere strategierne ved at sammenligne deres resultater, omkostninger og implementerbarhed.

Fordele ved disse kommandoer:
- Strukturerede svar: Kommandoerne guider AI til at levere klart organiserede og logisk opbyggede responser.
- Dybde og præcision: Hver kommando fremmer en analytisk tilgang, der sikrer detaljeret og velovervejet output.
- Tilpasningsevne: Disse kommandoer kan anvendes på tværs af fagområder, fra akademisk til professionel kontekst.

Ved at bruge kommandoer som "argumentér," "begrund," "konkludér," "analysér" og "evaluér" kan du udnytte AI's logiske kapacitet fuldt ud. Disse redskaber hjælper med at producere strukturerede, præcise og rationelle svar, der understøtter informeret beslutningstagning og dybere indsigt. Brugen af logik og ræsonnement styrker kvaliteten af AI's output og tilføjer værdi til komplekse opgaver.

Begrænsninger og udelukkelser

Kommandoer til kontrolleret output fra AI

Når du arbejder med AI, er det ofte nødvendigt at sætte klare begrænsninger eller ekskludere visse elementer for at opnå præcist og målrettet output. Ved hjælp af specifikke ord og fraser kan du definere, hvad AI'en ikke må inkludere, hvilket sikrer kontrol over indholdet og tilpasning til dine behov. Dette afsnit gennemgår effektive kommandoer til at styre sådanne begrænsninger.

JAN ENGELBRECHT PEDERSEN

Effektive kommandoer og deres anvendelse:

- Undgå
Brug "undgå" til at instruere AI om ikke at nævne bestemte emner, stilarter eller perspektiver. Det er en bred kommando, der kan anvendes i mange forskellige sammenhænge.
Eksempel:
"Undgå at nævne personlige meninger."
AI vil her producere et neutralt og faktuelt svar, der undgår subjektive holdninger.

- Udeluk
Kommandoen "udeluk" bruges, når du ønsker at ekskludere specifikke detaljer eller synspunkter fra AI's svar. Det sikrer, at unødvendige eller irrelevante oplysninger ikke inkluderes.
Eksempel:
"Udeluk spekulationer og hold dig til fakta."
AI vil fokusere på at levere velunderbyggede og faktuelle oplysninger, mens spekulative elementer bliver udeladt.

- Begræns
Nyttig til at sætte rammer for AI's output, f.eks. ved at indskrænke omfanget til en bestemt tidsperiode, emne eller perspektiv.
Eksempel:
"Begræns svaret til kun at omfatte de seneste fem år."
AI vil levere et svar, der fokuserer udelukkende på de nyeste oplysninger inden for den angivne tidsramme.

- Ignorér
Brug "ignorér," når du ønsker, at AI skal se bort fra visse aspekter eller emner i sit svar. Det er ideelt til at eliminere irrelevante faktorer.
Eksempel:
"Ignorér de politiske aspekter af sagen."
AI vil her koncentrere sig om de ikke-politiske elementer i emnet og udelade politiske analyser.

- Fravælg
"Fravælg" bruges til at ekskludere specifikke muligheder eller valg fra AI's respons. Dette kan tilpasses til stilistiske eller indholdsmæssige behov.
Eksempel:
"Fravælg negative formuleringer i teksten."
AI vil herefter undgå at bruge negative beskrivelser og levere et svar med et positivt fokus.

Fordele ved disse kommandoer:
- **Præcision**: Disse kommandoer giver mulighed for at tilpasse output ved at fjerne irrelevante eller uønskede elementer.
- **Kontrol**: Du styrer indholdet, så det matcher de krav og begrænsninger, du har defineret.
- **Kvalitet**: Ved at ekskludere uvedkommende aspekter forbliver svarene fokuserede og relevante.

Ved at anvende kommandoer som "undgå," "udeluk," "begræns," "ignorér" og "fravælg" kan du finjustere AI's respons og opnå et mere kontrolleret og fokuseret output. Disse redskaber sikrer, at AI leverer præcise svar, der udelukkende indeholder de elementer, du finder relevante og passende for din opgave. Anvend dem strategisk for at få fuld kontrol over resultatet.

Personligt: Brug "justér" eller "forbedr" til blogindlæg eller sociale medieopdateringer.

¤

Rollebaseret promptning

Skabelse af AI-personaer

JAN ENGELBRECHT PEDERSEN

Rollebaseret promptning handler om at formulere forespørgsler på en måde, der instruerer AI til at optræde som en bestemt person, ekspert eller figur. Denne tilgang giver dig mulighed for at kontrollere AI's adfærd og sikre, at svarene er skræddersyet til den valgte rolle, hvilket øger relevans og kreativitet.

Effektive rollebaserede formuleringer:
- "**Agér som**…"
Brug "Agér som…" for at instruere AI til at påtage sig en bestemt rolle, der matcher konteksten af din opgave.
Eksempel:
"Agér som en professionel advokat og rådgiv om juridiske aspekter af ejendomsoverdragelse."

AI vil herefter levere svar, der reflekterer en advokats viden og tone.

- "**Du er en ekspert i**…"
Når du ønsker detaljeret og præcis information, kan du bruge "Du er en ekspert i…" for at sætte AI i en position som fagperson.

Eksempel:
"Du er en ekspert i bæredygtig arkitektur. Forklar principperne bag grønne bygningsteknologier."

AI vil strukturere svaret med dybdegående analyser og teknisk viden.

- "**Forestil dig du er**…"
Brug denne kommando til at fremme AI's kreativitet ved at lade den forestille sig en bestemt rolle eller situation.

Eksempel:
"Forestil dig du er en astronaut, der beskriver sit første rumvandring."

AI vil producere et narrativ, der reflekterer perspektivet fra en astronaut og inkluderer detaljer om rumoplevelsen.

A/B-Testning

Evaluering af prompt performance

Ved at udrulle flere prompt-versioner parallelt kan performance måles kvantitativt.

JAN ENGELBRECHT PEDERSEN

A/B-testning er en systematisk metode til at evaluere effektiviteten af forskellige prompts. Ved at sammenligne to versioner af en prompt – ofte benævnt A og B – kan du analysere, hvilken der bedst opfylder de opstillede succeskriterier. Denne fremgangsmåde er ideel til at optimere prompt engineering, da den giver indsigt i, hvordan små variationer kan påvirke AI's output.

Fremgangsmåde i A/B-testning:

- Opret to versioner af prompts:
Start med at formulere to prompts, der repræsenterer forskellige tilgange til samme opgave. For eksempel:
- **Prompt A:** "Beskriv fordelene ved solenergi på en kortfattet måde."
- **Prompt B**: "Lav en detaljeret beskrivelse af fordelene ved solenergi med fokus på økonomiske og miljømæssige aspekter."
- **Generér output:**
Send begge prompts til AI og indhent de genererede svar. Sørg for, at testen udføres under ensartede betingelser for at sikre fair sammenligning.
- **Sammenlign resultater:**
Evaluér outputtet fra hver prompt baseret på succeskriterier som relevans, nøjagtighed, tone, struktur og kreativitet.

Eksempel: Hvis succeskriteriet er klarhed, kan Prompt A levere en mere overskuelig tekst, mens Prompt B muligvis skaber en omfattende analyse, der er bedre egnet til dybdegående opgaver.

- **Vælg den mest effektive prompt:**
Baseret på analysen kan du identificere, hvilken prompt der performer bedst. Eventuelt kan du justere prompts yderligere for at kombinere deres styrker.

Fordele ved A/B-Testning
- **Data-drevet indsigt:** A/B-testning giver objektive resultater, der gør det muligt at træffe velbegrundede beslutninger.
- **Hurtig iteration:** Ved at teste flere prompts kan du hurtigt tilpasse og forbedre din tilgang.
- **Effektivisering af prompt engineering**: Testning sikrer, at du opnår optimal kommunikation med AI-modellen.

Praktisk Eksempel

Opgave: Opret en marketingtekst for et produkt.
Prompt A: "Skriv en kort og direkte tekst om fordelene ved produktet."
Prompt B: "Skriv en overbevisende tekst, der også fokuserer på produktets unikke egenskaber."
Efter testning konkluderes det, at Prompt B leverer mere engagerende og dybdegående indhold, der egner sig til målgruppen.

Metrikker inkluderer:
Brugertilfredshed (thumbs up/down)
Gennemsnitlig responslængde
Konverteringsrater

En case fra Literal AI viste 28% forbedring i præcision ved A/B-test.

Brugerfeedback

Inddragelse for forbedring

Jan Engelbrecht Pedersen

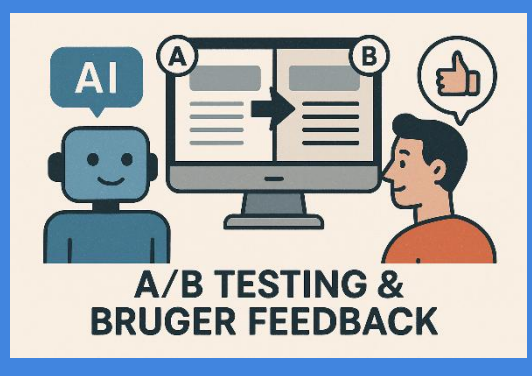
Fremgangsmåde for inddragelse af feedback:
- **Præsenter output:** Vis AI's genererede output til brugerne og bed dem evaluere det med fokus på specifikke kriterier, såsom nøjagtighed, klarhed og relevans.
- **Indsaml feedback:** Brug spørgeskemaer, samtaler eller vurderingsskemaer til at indsamle detaljeret feedback.

Eksempel: Brugerne kan indikere, om teksten fanger formålet og engagerer målgruppen, eller om den har svage formuleringer.

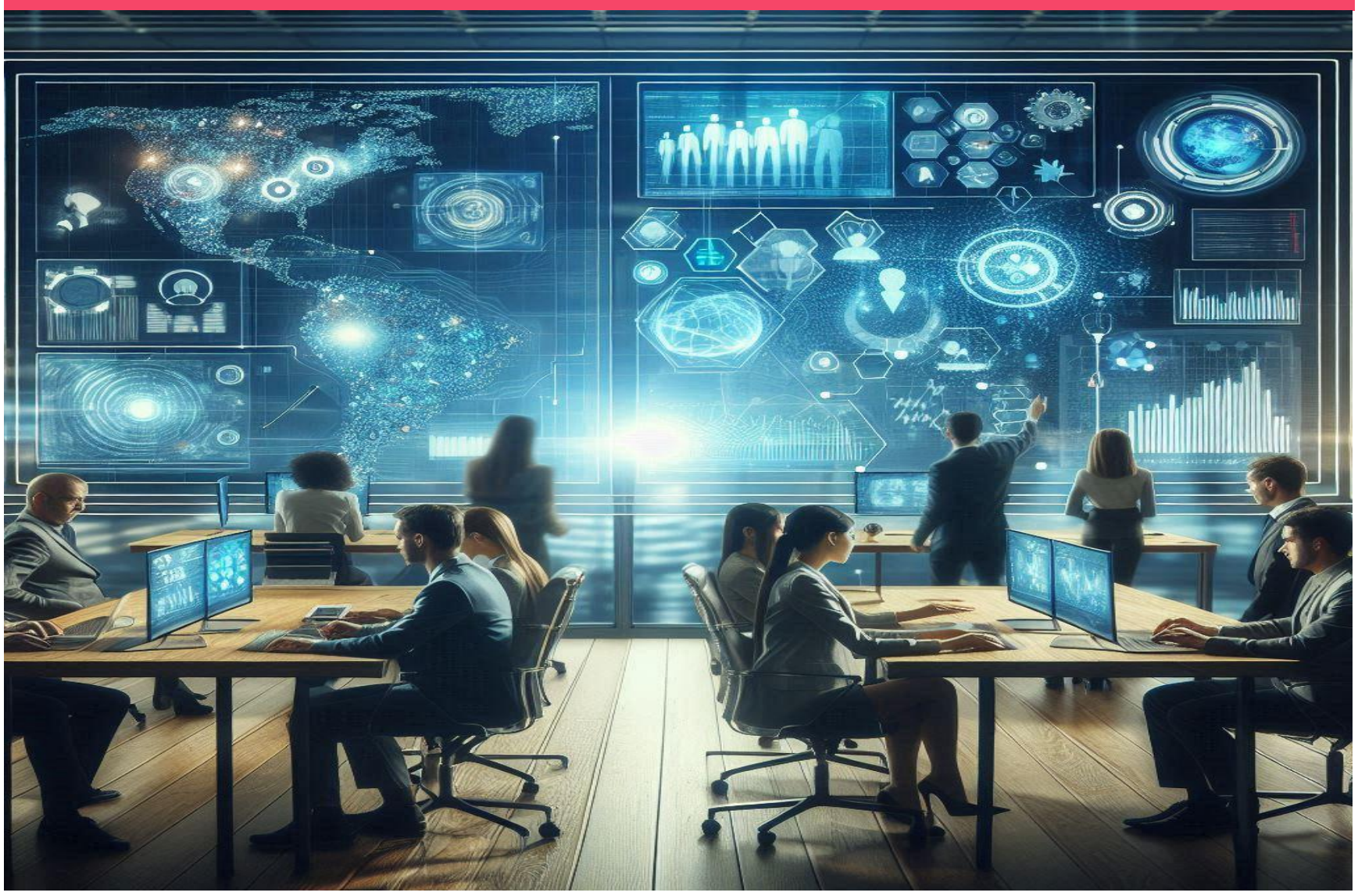
- **Analysér input:** Identificér de områder, hvor outputtet præsterer godt, og de områder, der skal forbedres. Hvis flere brugere bemærker manglende detaljer eller uhensigtsmæssig tone, kan dette være en indikation af nødvendige justeringer.

- **Justér og optimer prompts**: På baggrund af feedback kan du omformulere din prompt for at forbedre outputtet. Dette kan inkludere mere præcise instruktioner, specifik tone eller detaljeret kontekst.

Fordele ved brugerfeedback:
- **Forbedret relevans:** Feedback fra brugerne sikrer, at AI leverer resultater, der opfylder specifikke behov og forventninger.
- **Kvalitetssikring:** Brugerne hjælper med at evaluere og validere, at outputtet lever op til de definerede standarder.
- **Effektiv tilpasning:** Direkte input fra brugerne gør det muligt at justere prompts hurtigt og målrettet.



Kombinationen af A/B-testning og brugerfeedback udgør en effektiv strategi til at finjustere prompt engineering. A/B-testning leverer objektive data, mens brugerfeedback tilfører en kvalitativ dimension, der sikrer, at outputtet er tilpasset målgruppen. Ved at integrere disse metoder kan du markant forbedre både processen og resultatet af din interaktion med AI-modeller.



Social engineering er en metode, hvor kriminelle manipulerer mennesker til at afsløre information eller udføre handlinger. Det udnytter psykologiske reaktioner som tillid, frygt og autoritet. Generativ AI gør angreb mere sofistikerede og svære at opdage, f.eks. ved at automatisere personlig profilering og skabe overbevisende phishing-forsøg.

Introduktion til Social Engineering Generativ AI som redskab

JAN ENGELBRECHT PEDERSEN

Generativ AI har åbnet døren til en ny æra inden for teknologi, hvor kreativitet og automatisering smelter sammen. Men denne kraftfulde teknologi har også skabt nye udfordringer, især når det anvendes som et redskab i social engineering. Social engineering indebærer manipulation af mennesker til at afsløre fortrolige oplysninger eller udføre handlinger, ofte via psykologiske midler snarere end tekniske metoder.

Når generativ AI kombineres med social engineering, kan det bruges til at skabe overbevisende falske e-mails, beskeder og endda samtaler, der efterligner en person eller organisation med skræmmende præcision. Teknologien kan generere personlig skræddersyede angreb, der manipulerer følelser, bygger tillid og udnytter menneskelige svagheder mere effektivt end nogensinde før. Generativ AI har gjort det muligt at skabe overbevisende deepfake-fotos, hvilket udgør en betydelig trussel inden for social engineering. **Deepfake-teknologi** bruger avancerede algoritmer til at manipulere eksisterende billeder eller skabe fuldstændigt falske fotos, der ser autentiske ud. Dette kan bruges i social engineering til at udnytte tillid og manipulere mennesker.

Social engineering er særlig farlig, fordi den ikke nødvendigvis afhænger af teknologiske sårbarheder, men snarere på menneskelige handlinger. Derfor kræver beskyttelse mod sådanne angreb en kombination af tekniske sikkerhedsforanstaltninger og uddannelse af mennesker til at genkende og modstå manipulation.

Ved hjælp af avancerede tekst- og billedgenereringsværktøjer kan AI automatisere processen med at generere falske e-mails, beskeder eller websteder, der efterligner legitime organisationer og personer. For eksempel kan AI analysere offentligt tilgængelige data om en virksomhed eller individ og bruge disse oplysninger til at generere målrettede phishing-angreb, der virker ægte og troværdige. Dette kan indebære alt fra præcist efterlignede firma-logoer og formater til e-mails med en tone, der stemmer overens med organisationens kommunikationsstil. En af de mest skræmmende aspekter ved **AI-drevet phishing** er dens evne til at udnytte socialpsykologiske svagheder som tillid, frygt og travlhed. For at bekæmpe denne trussel er det nødvendigt at kombinere teknologiske modforanstaltninger med menneskelig uddannelse.

Metoder i **social engineering**: Flere teknikker anvendes, f.eks.:
- **Phishing**: AI skaber troværdige e-mails og manipulerer ofre til at dele følsomme data.
- **Vishing:** Stemmekloning bruges til realistiske telefonopkald, hvor angribere udgiver sig for autoriteter.
- **Pretexting:** Falske historier og dokumenter genereres af AI for at udgive sig som troværdige personer.
- **Tailgating**: Ved at kombinere AI med GPS-analyse planlægges fysisk adgang til sikrede områder.

Kriminelle udnytter AI til at forstærke principper som autoritet og social proof. For eksempel kan deepfakes simulere troværdige ledere, og falske likes kan manipulere offentlig opfattelse. AI forstærker også knapphedsfænomenet og skaber realistiske tidsbegrænsede tilbud.
Stemmekloning og deepfakes: Stemmekloning kræver kun få sekunders optagelse og kan bruges til falske samtaler eller omgå sikkerhed. Deepfakes laver realistiske videoer til svindel eller manipulation.

Risici og modforanstaltninger:

Generativ AI skaber nye trusler. For at bekæmpe dem kræves:

- **Bevidsthed:** Uddannelse i at genkende manipulerende AI-angreb.

- **Teknologi:** Biometriske systemer og avancerede spamfiltre.

- **Politikker:** Procedurer som to-faktor-godkendelse og identitetsverifikation.

Social engineering involverer at påvirke personer til at afsløre følsomme oplysninger eller til at udføre handlinger, de normalt ikke ville overveje. Dette opnås ofte ved at udnytte menneskelige psykologiske svagheder som tillid, frygt, skam eller autoritet. Angrebsmetoderne kan variere fra enkle tricks, såsom **phishing-e-mails** og **falske telefonopkald**, til mere sofistikerede teknikker, der udnytter omfattende research om målets adfærd og præferencer. Denne manipulation kan finde sted gennem direkte kontakt, som f.eks. en falsk medarbejder, der vinder en persons tillid, eller gennem teknologibaserede metoder, såsom links til skadelige websteder. Et centralt element i social engineering er at udnytte menneskers tendens til at tage hurtige beslutninger baseret på følelser eller pres fra en tilsyneladende autoritativ kilde.

Phishing
AI & Social Engineering

Phishing er en af de mest udbredte former for cyberangreb, hvor angribere sender falske e-mails eller beskeder, der fremstår som om, de kommer fra troværdige kilder. Formålet er at narre modtageren til at dele følsomme oplysninger, såsom loginoplysninger, kreditkortnumre eller andre personlige data. Angrebet bygger på manipulation af modtagerens tillid og troværdighedsopfattelse.

JAN ENGELBRECHT PEDERSEN

AI's rolle i phishing-angreb:
Generativ AI har drastisk forbedret effektiviteten af phishing-angreb ved at skabe mere overbevisende og skræddersyede beskeder. Ved hjælp af værktøjer som ChatGPT kan angribere generere e-mails med grammatisk korrekt og professionelt indhold, der er næsten umuligt at skelne fra ægte kommunikation. AI kan analysere offentligt tilgængelige data fra sociale medier og andre kilder for at opbygge detaljerede profiler af målgruppen. Disse oplysninger gør det muligt at formulere narrativer, der er nøje tilpasset modtagerens specifikke interesser, adfærd og kommunikationsstil, hvilket øger angrebets troværdighed.

For eksempel kan AI tage hensyn til en medarbejders jobfunktion og arbejdsopgaver, som afsløres gennem LinkedIn eller sociale platforme, og skabe beskeder, der direkte relaterer til medarbejderens daglige virke. Denne personalisering gør det sværere for modtageren at genkende angrebet som falsk.

Eksempler på AI-drevet phishing i praksis

Forestil dig en medarbejder, der modtager en e-mail fra "IT-support". Beskeden indeholder en AI-genereret signatur, der perfekt matcher virksomhedens stil, tone og design. I beskeden står, at modtagerens konto har "sikkerhedsproblemer" og kræver en opdatering via et vedhæftet link. Linket fører til en falsk login-side, der visuelt er identisk med virksomhedens officielle portal. Ved hjælp af AI skabes en besked, som på subtil vis inkluderer data indsamlet fra sociale medier, såsom medarbejderens navn, afdeling eller tidligere interaktion med IT-afdelingen. Resultatet er en phishing-e-mail, der er skræddersyet til modtageren og ekstremt svær at afsløre.

Udvidede scenarier og risici
Phishing

JAN ENGELBRECHT PEDERSEN

AI-drevne phishing-angreb har potentiale til at ramme bredere målgrupper på kort tid, og deres effektivitet truer ikke kun enkeltpersoner, men også organisationer. Eksempler på udvidede phishing-teknikker inkluderer:

- **Falske invitationer til webinarer eller begivenheder:** Angribere kan generere e-mails med personaliserede emner og links, der leder modtagerne til skadelige websites.

- **Fake betalingsanmodninger:** Beskeder, der udgiver sig for at være fra økonomiafdelingen med anmodning om godkendelse af fakturaer.

- **Social engineering via SMS (Smishing):** Brug af AI til at skabe tekstbeskeder, der udgiver sig for at være fra banken eller andre vigtige aktører, med links til falske websites.

Modforanstaltninger og beskyttelse
Phishing

JAN ENGELBRECHT PEDERSEN

For at beskytte sig mod AI-drevet phishing er det afgørende at implementere robuste sikkerhedsstrategier, som kan reducere risikoen og styrke beskyttelsen. Her er en udvidet gennemgang med eksempler på effektive metoder:

- **Avancerede spamfiltre:**
Disse værktøjer anvender maskinlæring til at analysere og identificere phishing-e-mails baseret på mønstre, ordvalg og strukturelle karakteristika. For eksempel kan spamfiltre opdage e-mails, der indeholder links til falske websteder, eller beskeder med unormale sprogvariationer, som ofte bruges i phishing-angreb. Et praktisk eksempel kan være en virksomhed, der implementerer spamfiltre for at blokere e-mails, der hævder at komme fra deres bank, men som indeholder små stavefejl eller unøjagtige domæner.

- **Multilagsautentificering:** To-faktor-godkendelse tilføjer et ekstra sikkerhedslag ved at kræve både en adgangskode og en yderligere bekræftelse, såsom en SMS-kode eller en biometrisk scanning. Et eksempel kunne være en medarbejder, der forsøger at logge ind på virksomhedens system. Selvom deres adgangskode bliver kompromitteret via phishing, kan de ikke få adgang uden den sekundære bekræftelseskode,som sendes til en autoriseret enhed.
- **Uddannelse af medarbejdere:** Træning i at genkende troværdige e-mails kontra phishing-beskeder.

Vishing
Voice Phishing

JAN ENGELBRECHT PEDERSEN

Vishing er en avanceret form for phishing, hvor angriberen anvender telefonopkald som primært redskab til at manipulere ofre og opnå adgang til følsomme oplysninger. Typisk præsenterer angriberen sig som en autoritativ kilde, f.eks. en bankmedarbejder, en teknisk supportrepræsentant eller en kontakt fra en virksomhed, som offeret har en legitim forbindelse til. Målet er at udnytte offerets tillid og skabe en følelse af nødvendighed eller pres for at opnå de ønskede oplysninger.

Generativ AI har gjort det muligt for angribere at udføre vishing-angreb med øget realisme og effektivitet. Ved hjælp af stemmekloningsteknologier som ElevenLabs kan AI præcist efterligne en persons stemme baseret på korte lydoptagelser. Dette gør det muligt for angriberen at udgive sig for at være en velkendt autoritet, leder eller kollega med høj troværdighed. Desuden gør AI-drevne chatbots det muligt at føre flydende og overbevisende samtaler, der efterligner menneskelig kommunikation, hvilket gør det ekstremt svært for offeret at opdage, at interaktionen er falsk.

AI kan også personalisere angreb ved at integrere data fra sociale medier og offentlige kilder. Dette giver angriberen en mere målrettet tilgang, der føles autentisk for den specifikke modtager.

Eksempel på AI-drevet vishing: En direktør i en virksomhed modtager et telefonopkald fra, hvad der lyder som "deres CFO". Stemmen på den anden linje bærer alle karakteristika af CFO'ens tale – fra tonefald til sproglige nuancer – men i virkeligheden er det en angriber, der bruger stemmekloningsteknologi. Under opkaldet anmoder den falske CFO om en hurtig pengeoverførsel til en "presserende forretningsaftale", hvilket skaber en følelse af tidsmæssigt pres. På grund af stemmens autenticitet og den overbevisende fortælling, bliver direktøren narret til at foretage overførslen uden at stille spørgsmål.

Andre mulige anvendelser af AI i vishing: Ud over den klassiske telefonsamtale kan AI også bruges i andre vishing-scenarier:
- Falske voicemail-beskeder: Stemmekloning kan skabe autentiske beskeder, der virker som om, de er fra en pålidelig kilde.

- **Automatiserede opkaldskampagner:** AI-drevne chatbots kan masseudbrede troværdige opkald med skræddersyede beskeder til forskellige målgrupper.

- **Kombination med deepfakes:** Ved at inkludere videoelementer, der supplerer stemmen, kan angribere yderligere styrke deres troværdighed.

For at imødegå truslen fra AI-drevet vishing er det afgørende at implementere robuste sikkerhedsforanstaltninger:
-**Verifikationsprocedurer**: Indfør "callback"-protokoller, hvor modtagere af opkald verificerer identiteten på afsenderen gennem en separat og sikker kanal.
- **Begræns deling af stemmeprøver:** Offentlige optagelser af autoritære personer kan bruges til stemmekloning. Derfor bør sådanne data begrænses.
- **Bevidsthedstræning:** Uddan medarbejdere i at identificere potentielt manipulerende kommunikation, selv når denne fremstår overbevisende og autentisk. AI's evne til at forbedre vishing gennem stemmekloning og automatisering repræsenterer en markant trussel mod cybersikkerheden. Ved at forstå de teknologiske muligheder og risici kan virksomheder og enkeltpersoner dog tage de nødvendige forholdsregler for at beskytte sig mod denne form for digital manipulation.

Autoritet
AI's rolle i social engineering

Jan Engelbrecht Pedersen

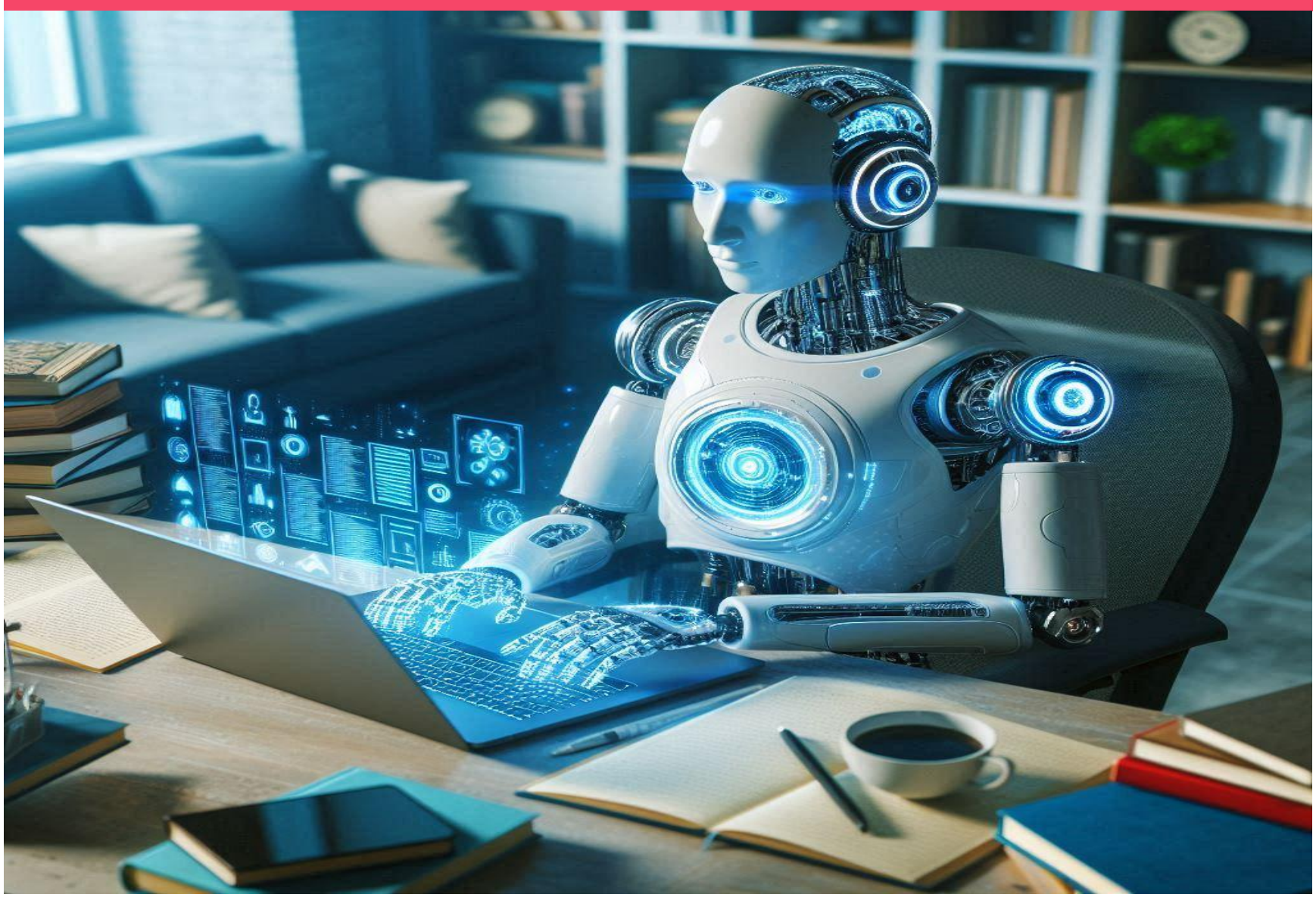
Autoritet er et kraftfuldt psykologisk fænomen, hvor mennesker naturligt følger instrukser fra personer, der fremstår magtfulde eller legitime. Dette princip kan udnyttes i social engineering, hvor angribere manipulerer ofre til at handle imod deres bedste interesser. Ved hjælp af generativ AI er mulighederne for at skabe falske autoritetsfigurer blevet både mere realistiske og farligere
Generativ AI kan anvende teknologier som **stemmekloning** og **deepfakes** til at simulere stemmer og udseender af personer i autoritetspositioner, eksempelvis virksomhedsledere, embedsmænd eller advokater.

Dette gør det muligt for angribere at sende overbevisende beskeder, e-mails eller videoer, der får det til at fremstå som om, en leder fremsætter en anmodning. Et typisk scenarie kunne være en medarbejder, der modtager en e-mail fra en "administrerende direktør," som beder om en hastende pengeoverførsel. Beskeden er yderst troværdig, med præcist sprog, professionelt udseende grafik og tilhørende vedhæftede filer. For yderligere at styrke illusionen kan e-mailen være suppleret med en **AI-genereret stemmebesked** eller en **deepfake-video**, der tilsyneladende bekræfter anmodningen.

Derudover kan AI bruges til at generere falske juridiske meddelelser eller kommunikation fra offentlige myndigheder, hvilket kan presse ofre til at handle hurtigt. Disse angreb udnytter menneskers tendens til at adlyde autoritet, særligt i hierarkiske strukturer.

For at forebygge sådanne angreb bør organisationer indføre verifikationsprocesser, som fx at bekræfte anmodninger via separate kanaler.

Uddannelse af medarbejdere til at genkende manipulation, kombineret med avancerede teknologiske løsninger til at opdage stemmekloning og deepfakes, er også nødvendigt for at beskytte sig mod denne trussel.



Pretexting
AI i Social Engineering

JAN ENGELBRECHT PEDERSEN

Pretexting er en strategisk og manipulerende metode, hvor angribere opbygger en falsk historie eller identitet for at narre personer til at afsløre følsomme oplysninger eller give adgang til beskyttede systemer. Denne teknik afhænger af angriberens evne til at udnytte tillid og skabe en illusion af legitimitet gennem en overbevisende narrativ. Ved at optræde som en autoritet eller en betroet person kan angriberen overbevise ofre om, at deling af information er nødvendigt og ufarligt. Pretexting er særligt effektivt i miljøer, hvor hierarki og autoritet spiller en afgørende rolle.

Konkrete eksempler på AI-drevet pretexting:
Falske e-mails fra HR: En medarbejder modtager en professionel e-mail fra en "HR-konsulent," der anmoder om opdatering af personlige oplysninger. Formularen og e-mailen er genereret af AI, som har analyseret virksomhedens kommunikationsstil.
Virtuelle jobinterviews: En angriber poster en falsk jobannonce og organiserer interviews via videoopkald. Under samtalen indhentes personlige data og adgangsinformation.
Falske tekniske support-opkald: En IT-medarbejder modtager en anmodning fra en "tekniker," der har brug for adgang til systemer for vedligeholdelse. AI genererer detaljerede fejlrapporter, der understøtter denne falske historie.
Finansielle svindeloperationer: E-mails fra "revisorer" eller "projektledere," der anmoder om ændringer i bankoplysninger.
Falske myndighedsmeddelelser: Dokumenter og e-mails, der udgiver sig for at være fra regeringsorganer, presser ofre til handling.
Kundebedrag: Falske beskeder fra "leverandører," der kræver fortrolige oplysninger for en påstået ordreopfølgning.

AI's rolle i pretexting

Generativ AI har revolutioneret pretexting ved at gøre det lettere at skabe realistiske og overbevisende falske historier.
Ved hjælp af AI kan angribere hurtigt producere autentiske dokumenter, såsom e-mails, rapporter og kontrakter, der efterligner en legitim kommunikation. For eksempel kan AI generere troværdige CV'er eller stillingsannoncer, der optræder som autentiske fra etablerede virksomheder.

Generativ AI forstærker pretexting ved at skabe overbevisende falske historier, identiteter og dokumenter. Angribere kan bruge AI til at generere realistiske e-mails, CV'er og deepfakes, der manipulerer tillid. For eksempel kan en falsk "HR-medarbejder" bruge AI til at anmode om følsomme oplysninger gennem troværdig kommunikation, hvilket gør pretexting mere effektivt og farligt

AI kan også manipulere visuelle og grafiske elementer. Med værktøjer som deepfakes og stemmekloning kan angribere efterligne ansigtet og stemmen af en kendt person, som f.eks. en leder eller teknisk ekspert.
Kombinationen af autentisk-synende materiale og avanceret personliggørelse gør pretexting mere overbevisende og uigennemskueligt.

Forebyggelse og modforanstaltninger

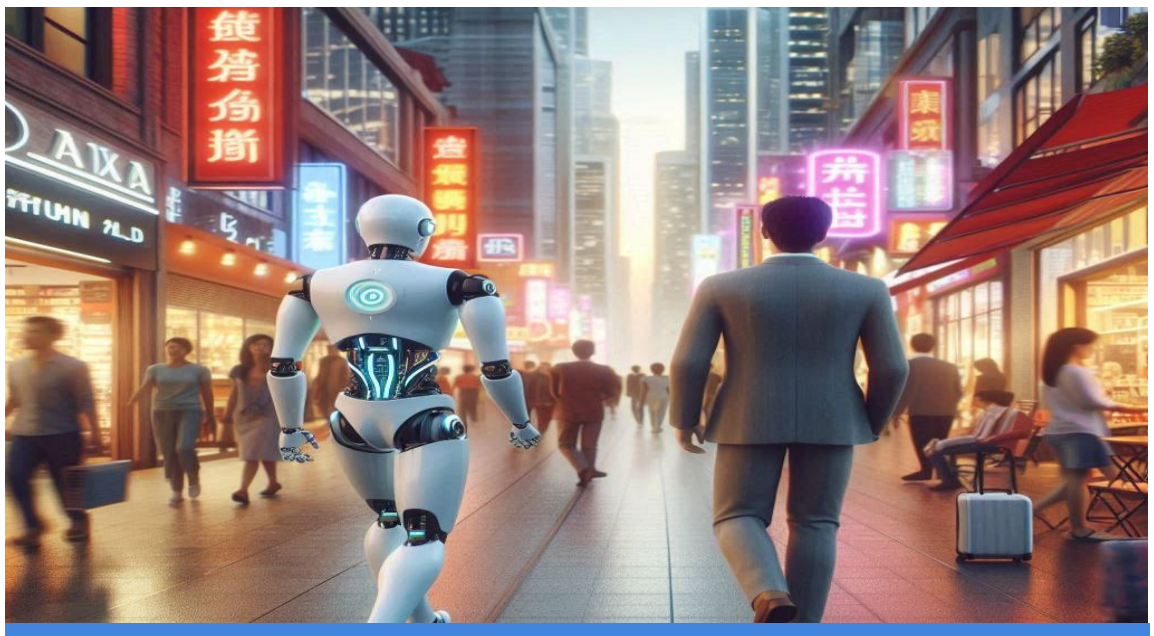
For at beskytte mod AI-drevne pretexting-angreb skal der implementeres flere tiltag:

Verifikation: Alle anmodninger skal verificeres via uafhængige kanaler, f.eks. telefonopkald eller personlige møder.

Uddannelse: Regelmæssig træning hjælper medarbejdere med at genkende og afvise manipulation. Simulation af phishing og pretexting-angreb kan være en del af træningen.

Teknologiske løsninger: Brug værktøjer til at identificere stemmekloning, deepfakes og unormal e-mailaktivitet.

AI's rolle i pretexting understreger nødvendigheden af at kombinere menneskelig opmærksomhed med teknologiske modforanstaltninger for effektiv beskyttelse mod denne trussel.



Tailgating
AI i Social Engineering

Tailgating er en fysisk form for social engineering, hvor en angriber får adgang til sikrede områder ved at følge efter autoriserede personer, som har adgang til at åbne døre eller låse systemer op. Denne metode udnytter menneskelig tillid og uopmærksomhed frem for teknologiske sårbarheder. Angriberen præsenterer sig ofte som en legitim medarbejder, gæst eller tekniker for at undgå mistanke og skabe en følelse af naturlighed i interaktionen.

JAN ENGELBRECHT PEDERSEN

Generativ AI kan spille en væsentlig rolle i at forbedre præcisionen og planlægningen af tailgating-angreb. Ved at integrere AI med GPS-tracking-teknologi, såsom AirTags, kan angriberen få adgang til realtidsdata om medarbejderes bevægelser. AI kan analysere adgangsdata, der er indsamlet fra overvågningssystemer, for at identificere mønstre i medarbejdernes rutiner og afgøre optimale tidspunkter for angreb.

Denne automatisering gør det muligt for angriberen strategisk at time angrebet for at øge chancerne for succes, mens risikoen for opdagelse reduceres.
Eksempelvis kan AI samle information om, hvornår en bestemt medarbejder ankommer til kontoret, og hvilket indgangspunkt vedkommende benytter. Angriberen kan udnytte disse data og simulere en legitim årsag til at følge efter medarbejderen ind i bygningen, såsom at bære en uniform eller udstyr, der indikerer teknisk arbejde. Dette øger angrebets troværdighed og gør det mindre sandsynligt, at medarbejderen bemærker noget usædvanligt.
Eksempler på AI-drevet Tailgating:
Ud over GPS-integration kan AI bruges til at understøtte forskellige aspekter af tailgating-angreb:
- **Simuleret tillid:** Angriberen kan udnytte AI-genererede profiler og dokumentation, såsom identifikationskort, der efterligner virksomhedens standarder.

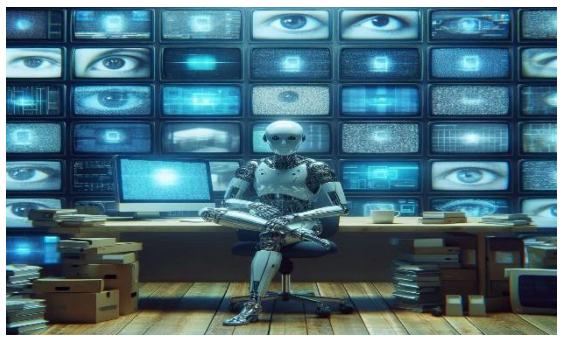
- **Timing af flere mål:** AI kan analysere bevægelsesmønstre fra flere medarbejdere samtidigt og optimere angreb baseret på de mest tilgængelige og uopmærksomme personer.

- **Kombination med andre metoder:** Tailgating kan forstærkes ved brug af social engineering-metoder som pretexting eller phishing for yderligere adgangsprivilegier.

For at beskytte mod tailgating-angreb, der involverer AI, bør organisationer implementere en række fysiske og teknologiske sikkerhedsforanstaltninger:

- **Adgangskontrol:** Uddannelse af medarbejdere i ikke at lade fremmede følge med ind i bygningen uden korrekt godkendelse.
- **Overvågningssystemer:** Brug af AI-drevet analyse til at opdage mistænkelige bevægelsesmønstre.
- **Teknologiske barrierer:** Indfør adgangssystemer, der kræver unikke godkendelser, såsom biometriske data, for hver person, der kommer ind.

Tailgating udnytter en kombination af fysisk nærhed og menneskelig tillid til at bryde ind i sikre områder. Generativ AI har gjort det muligt for angribere at planlægge og optimere disse angreb med hidtil uset præcision. Ved at kombinere teknologi, træning og opmærksomhed kan organisationer imidlertid reducere risikoen for tailgating og beskytte sig mod både traditionelle og AI-forstærkede angreb. Denne helhedsorienterede tilgang er nødvendig for at sikre, at menneskelig tillid ikke udnyttes som en svaghed.



Tillid
AI i Social Engineering

JAN ENGELBRECHT PEDERSEN

Tillid er en grundlæggende del af menneskets sociale natur og en essentiel egenskab i opbygningen af relationer og samarbejde. Mennesker er naturligt tilbøjelige til at stole på andre, især når de oplever venlighed, hjælp eller troværdighed. Denne tillid, der ofte bygger på sociale normer og tidligere erfaringer, gør os sårbare over for manipulation, især når angribere målrettet udnytter den til egne formål. Generativ AI har vist sig at være en kraftfuld ressource for angribere, der ønsker at udnytte tillid som en psykologisk mekanisme.

Ved at analysere sprogbrug, adfærd og præferencer kan AI skabe realistiske og overbevisende interaktioner, der fremstår som venlige og hjælpsomme. Gennem teknologier som AI-drevne chatbots eller automatiserede beskedsystemer kan angribere simulere en troværdig person eller tjeneste, hvilket øger sandsynligheden for, at ofre deler følsomme oplysninger eller udfører handlinger, de normalt ville være skeptiske overfor. AI's evne til at tilpasse kommunikation baseret på målgruppens reaktioner gør denne manipulation endnu mere effektiv.
For eksempel kan en chatbot designet af angriberen opføre sig som en kundeservicemedarbejder, der yder venlig og imødekommende hjælp, hvilket gør det vanskeligt for offeret at opdage bedraget.

Angribere kan yderligere optimere chatbotten ved at trække på brugerdata fra sociale profiler eller tidligere kommunikation for at skabe personaliserede oplevelser, der føles autentiske.

Eksempel på misbrug af AI

En angriber udvikler en AI-drevet chatbot, der udgiver sig for at være kundesupport for en velkendt virksomhed. Chatbotten engagerer offeret gennem et venligt og hjælpsomt tonefald, som får samtalen til at virke troværdig. Efter at have opbygget tillid anmoder chatbotten om detaljerede oplysninger, såsom loginoplysninger eller betalingskortdata, med påskud af at løse et problem. Offeret, der føler sig tryg ved den hjælpsomme kommunikation, er mere tilbøjelig til at dele de efterspurgte oplysninger.

AI-Genereret kommunikation
AI i Social Engineering

JAN ENGELBRECHT PEDERSEN

Generativ AI har udvidet mulighederne for cyberkriminelle til at manipulere med kommunikation og målrette ofre mere præcist end nogensinde før. Ved hjælp af avanceret teknologi kan AI skabe troværdige og skræddersyede interaktioner, der udnytter tekst, lyd og video til at narre ofre. Denne fleksibilitet gør det muligt for angribere at tilpasse og raffinere deres angreb i realtid baseret på deres måls reaktioner, hvilket gør disse metoder både dynamiske og svære at opdage.

AI's værktøjer til kommunikationsmanipulation:

Generativ AI tilbyder en række værktøjer, der effektivt udnytter kommunikationens potentiale som angrebsstrategi. Disse inkluderer:

- **Personlige phishing-mails**: Ved at analysere modtagerens online adfærd og præferencer kan AI generere overbevisende e-mails med grammatisk korrekt sprog og indhold, der er designet til at virke autentisk. Disse e-mails er skræddersyede til modtagerens kommunikationsstil og har større succesrate end traditionelle phishing-forsøg.

- **Stemmekloning**: Generativ AI kan efterligne stemmer af kendte personer, såsom chefer, kolleger eller familiemedlemmer, med præcision. Dette gør det muligt for angribere at opbygge tillid hos ofrene gennem troværdig lydkommunikation og til at anmode om følsomme oplysninger.

- **Deepfake-videoer**: Ved hjælp af AI kan angribere skabe overbevisende videoer, der manipulerer ansigter og stemmer for at narre ofre. Disse videoer kan bruges til at manipulere beslutningstagere, true eller presse individer ved at vise dem falske, kompromitterende situationer.

En af de mest sofistikerede aspekter af AI-genereret kommunikation er dens evne til at tilpasse sig offerets svar i realtid. Ved at analysere modtagerens adfærdsmønstre og svar under interaktionen kan AI dynamisk ændre sine beskeder for at virke endnu mere overbevisende. Dette inkluderer justering af tonen i samtalen, tilføjelse af relevante detaljer baseret på målgruppens reaktion eller anvendelse af strategier, der styrker modtagerens tillid.
Eksempel på AI, der tilpasser sig i realtid
En medarbejder modtager en phishing-e-mail, der udgiver sig for at være fra IT-afdelingen. AI står bag kommunikationen og opfordrer medarbejderen til at klikke på et link for at "opdatere deres sikkerhedsindstillinger." Efter at medarbejderen tøver og svarer, at de er usikre på legitimiteten af e-mailen, genererer AI øjeblikkeligt en personlig besked som svar: "Vi forstår din bekymring. Du kan ringe til vores sikre linje på det nummer, der er vedhæftet. Her kan vi verificere din konto." AI producerer endda en stemmeklonet telefonbesked fra "IT-afdelingen," der bekræfter historien, hvilket overbeviser medarbejderen om, at kommunikationen er autentisk. Denne dynamiske justering gør manipulationen langt mere effektiv

AI-manipulation

Jan Engelbrecht Pedersen

En angriber bruger generativ AI til at lave en personlig phishing-mail, der ligner kommunikation fra modtagerens bank. Mailen inkluderer et link, hvor modtageren bliver bedt om at angive følsomme oplysninger for at "løse et sikkerhedsproblem." Derudover anvender angriberen stemmekloning til at efterligne bankens kundeservice via en troværdig telefonbesked fra en "bankrepræsentant," eller ved hjælp af deepfake-teknologier skabes en video, hvor en "bankdirektør" forklarer vigtigheden af at følge instruktionerne. Disse detaljer gør angrebet meget troværdigt og sværere at opdage.

Lån og finansiering: En person modtager en e-mail, som ser ud til at komme fra en finansiel rådgiver, der tilbyder et favorabelt lån. Med AI-genererede dokumenter, der ser ægte ud, og en stemmebesked fra "rådgiveren" bliver offeret narret til at overføre penge som en "administrationsgebyr."

Jobansøgninger: En falsk HR-rekrutterer bruger AI til at sende e-mails til jobsøgende og inviterer dem til virtuelle interviews. Her bliver kandidater bedt om at dele detaljer som CPR-numre og bankoplysninger under påskud af at afslutte ansættelsesprocessen.

Falske fakturaer: En virksomhed modtager en e-mail med en vedhæftet AI-genereret faktura, der påstås at være fra en kendt leverandør. E-mailen opfordrer virksomheden til at betale en stor sum til en falsk bankkonto, komplet med dokumentation, som AI har designet til at ligne den ægte.

Konto-verifikation: En besked, der ligner officiel kommunikation fra en streamingtjeneste, fortæller brugeren, at deres konto er ved at blive lukket, medmindre de "verificerer" deres betalingsoplysninger via et link. AI genererer en troværdig hjemmeside, der indsamler deres kreditkortoplysninger.

Anvendelse af menneskelig psykologi i cyberangreb

JAN ENGELBRECHT PEDERSEN

Generativ AI udgør en revolutionerende udvikling inden for cyberkriminalitet, der åbner nye muligheder for angribere til at udnytte menneskets fundamentale psykologiske reaktioner.

Disse reaktioner, som omfatter tillid, autoritet, social proof og frygten for at gå glip af noget (FOMO), bliver målrettet manipuleret på en skalerbar og effektiv måde ved hjælp af AI. Den teknologis avancerede evner gør det muligt for angribere at individualisere deres strategier, hvilket gør traditionelle metoder til cybersikkerhed utilstrækkelige.

Skalering og personaliserede angreb

En af de mest fremtrædende styrker ved generativ AI er dens evne til at skræddersy angreb til både individuelle ofre og grupper. Dette inkluderer skabelsen af detaljerede phishing-e-mails, som afspejler modtagernes sprogbrug og interesser, samt brug af deepfakes og stemmekloning til at efterligne realistiske kommunikationsformer. Personaliserede angreb reducerer ofrenes modstandskraft ved at gøre manipulationen mere autentisk og vanskelig at afsløre.

Generativ AI giver cyberkriminelle en hidtil uset mulighed for at udnytte psykologiske mekanismer i sociale angreb. Med en kombination af oplysningskampagner, teknologisk innovation og stringente procedurer kan vi dog styrke forsvaret mod disse trusler. Ved at anerkende de udfordringer, som generativ AI udgør, og handle proaktivt, kan vi skabe en sikrere digital fremtid for både organisationer og enkeltpersoner.

Mange traditionelle sikkerhedsforanstaltninger – såsom spamfiltre, simple adgangskoder og manuel overvågning – kæmper med at imødegå de avancerede angreb, der anvender generativ AI. Disse værktøjer er ofte ineffektive over for dygtigt udførte, AI-drevne manipulationer. Det bliver derfor nødvendigt at opdatere og forbedre sikkerhedsprotokoller samt investere i teknologiske løsninger, der specifikt kan opdage og forhindre angreb med afsæt i generativ AI.

Integreret forsvar: Kombination af opmærksomhed og teknologi
For at tackle udfordringerne fra generativ AI er det afgørende at implementere en flerdimensional tilgang til cybersikkerhed:

- **Uddannelse:** Uddannelse af medarbejdere og individer er afgørende, så de kan genkende tegn på psykologisk manipulation og identificere AI-drevne angreb.

- **Avancerede teknologiske løsninger:** Implementeringen af avancerede detektionsteknologier, som kan identificere deepfakes, stemmekloning og andre AI-genererede trusler.

- **Proceduremæssige sikkerhedsforanstaltninger:** Udvikling af omfattende protokoller, såsom flerlagsautentificeringssystemer og klare guidelines for håndtering af følsomme oplysninger.

I takt med at generativ AI bliver mere sofistikeret, vil angrebsmetoder fortsat udvikles og udfordre eksisterende forsvarsmekanismer. Dette kræver en vedvarende indsats for at justere forsvarsstrategier og implementere de nyeste teknologier. Samtidig er der behov for, at brugen af generativ AI reguleres og overvåges nøje for at minimere dens potentielle misbrug i cyberkriminalitet.

Indholdsfortegnelse

Side	Kapitel	Indhold
1		Forord - En rejse ind i generativ AI´s univers
2		Generel introduktion og AI-assistenter & kvantecomputere
3		AI - nu og i fremtiden
4	1	Hvad er chatbots? Forstå de nye værktøjer, fordele og udfordringer
5	1	Perspektivering: Chatbots i fremtiden
6	1	Chatbots : Claude (Anthropic)
7	1	Chatbots : DeepSeek - Den kinesiske udfordrer & Perplexity AI - Informationssøgning i samtaleform
8	1	Chatbots : Microsoft CoPilot - Integration i produktivitetssuiten
9	1	Chatbots : Grok 3 - X's frittalende AI
10	1	Chatbots : Sammenligning af funktioner og anvendelsesområder & Fremtidsperspektiver og teknologisk dybde
11	2	Hvad er AI ? En bred definition & Træning af neurale netværk
12	2	Optimeringsalgoritmer - At finde de bedste vægte
13	2	Large Language Models (LLM) - Sprogets mestre & Træning af LLM'er - Data og skalering
14	2	Forståelse og generering af naturligt sprog & Der findes tre hovedtyper af maskinlæring
15	2	Naturlig Sprogbehandling (NLP)
16	2	Computer Vision & Computer Vision - Personlige forslag drevet af kunstig intelligens
17	2	Anbefalingssystemer - Anbefalingssystemer & Forklaring af centrale AI-termer i anbefalingssystemer
18	3	Fra LLM til Chatbot - Integrering af komponenter
19	3	Prompt Engineering - Kunsten at kommunikere med AI & Prompt Engineering - Kunsten at styre AI
20	3	Finjustering og træning af chatbots til specifikke formål
21	3	Hukommelse - Kontekstforståelse i samtaler
22	3	Sikkerhed og bias & Arkitekturer
23	4	Ansvar for generativ AI-output - Hvem har ansvaret? & Culpa-ansvar ved brug af Generativ AI
24	4	Ophavsretlige udfordringer - Ophavsretlige udfordringer
25	4	EU-Regulering af AI - AI-loven og dens Implikationer & Regulering i USA & Dansk Lovgivning
26	4	GDPR - Privatlivsproblematikker ved brug af generativ AI & AI Act - I Danmark
27	4	Kriminelles anvendelser af generativ AI - Deepfakes og misbrug
28-35	5	Oversigt over alle generative AI værktøjer og chatbots
36	6	Tekstgenerering - Artikler, blogs, kreativ skrivning m.m. & Betydningen af kvaliteten af prompts
37	6	Billedgenerering - Kunst, design og visualisering & Eksempler på AI-billedgenereringsværktøjer
38	6	Musikgenerering - Komposition og produktion
39	6	AI-lydgenerering - Teknologi, metoder og anvendelse & AI-lydgenerering - Avancerede teknikker i kort form
40	6	Talegenerering - Voiceovers, syntetiske stemmer og tilgængelighed & Kvalitet og etik
41	6	Søgning efter Information - Intelligent informationssøgning
42	6	Generering af dokumenter - Rapporter, præsentationer & Eksempler på funktioner og anvendelser
43	6	Andre innovative - anvendelsesområder for AI & Generativ AI - Nye roller
44	7	AI-agenter - Den næste evolution inden for AI
45	7	Eksempler på AI agenter
46	7	Eksempler på AI agenter & Fremtidens AI-agenter
47	7	Artificial General Intelligence (AGI) - Drømmen om menneskelignende intelligens
48	7	Etiske og samfundsmæssige konsekvenser - Fremtidens AI & Etiske og samfundsmæssige konsekvenser
49	8	Prompt Engineering - Kunsten at kommunikere med AI & Hvad er en prompt? - Det grundlæggende
50	8	Grundprincipper for opbygning af prompts - Klarhed, koncished og kontekst
51	8	Brug af stil, tone, format og teknikker til optimering af prompts & Format og Struktur
52	8	Kreativitet og analyse & Instruktioner og Begrænsninger
53	8	Avanceret Prompting
54	8	Role-Playing og iterativ prompt & Constraint Prompting , prompt- decomposition & Selv-konsistens prompting
55	8	Relationer og sammenligninger - Kommandoer til identifikation og beskrivelse af relationer &
56	8	Logik og ræsonnement - Kommandoer til styring af AI's analytiske tænkning
57	8	Begrænsninger og udelukkelser - Kommandoer til kontrolleret output fra AI & Rollebaseret promptning
58	8	A/B-Testning - Evaluering af prompt performance & Brugerfeedback - Inddragelse for forbedring
59	9	Introduktion til Social Engineering - Generativ AI som redskab
60	9	Phishing - AI & Social Engineering & Udvidede scenarier og risici - Phishing & Modforanstaltninger og beskyttelse
61	9	Vishing - Voice Phishing & Autoritet - AI's rolle i social engineering
62	9	Pretexting - AI i Social Engineering
63	9	Tailgating - AI i Social Engineering & Tillid - AI i Social Engineering
64	9	AI-Genereret kommunikation - AI i Social Engineering & AI-manipulation
65	9	Anvendelse af menneskelig psykologi i cyberangreb