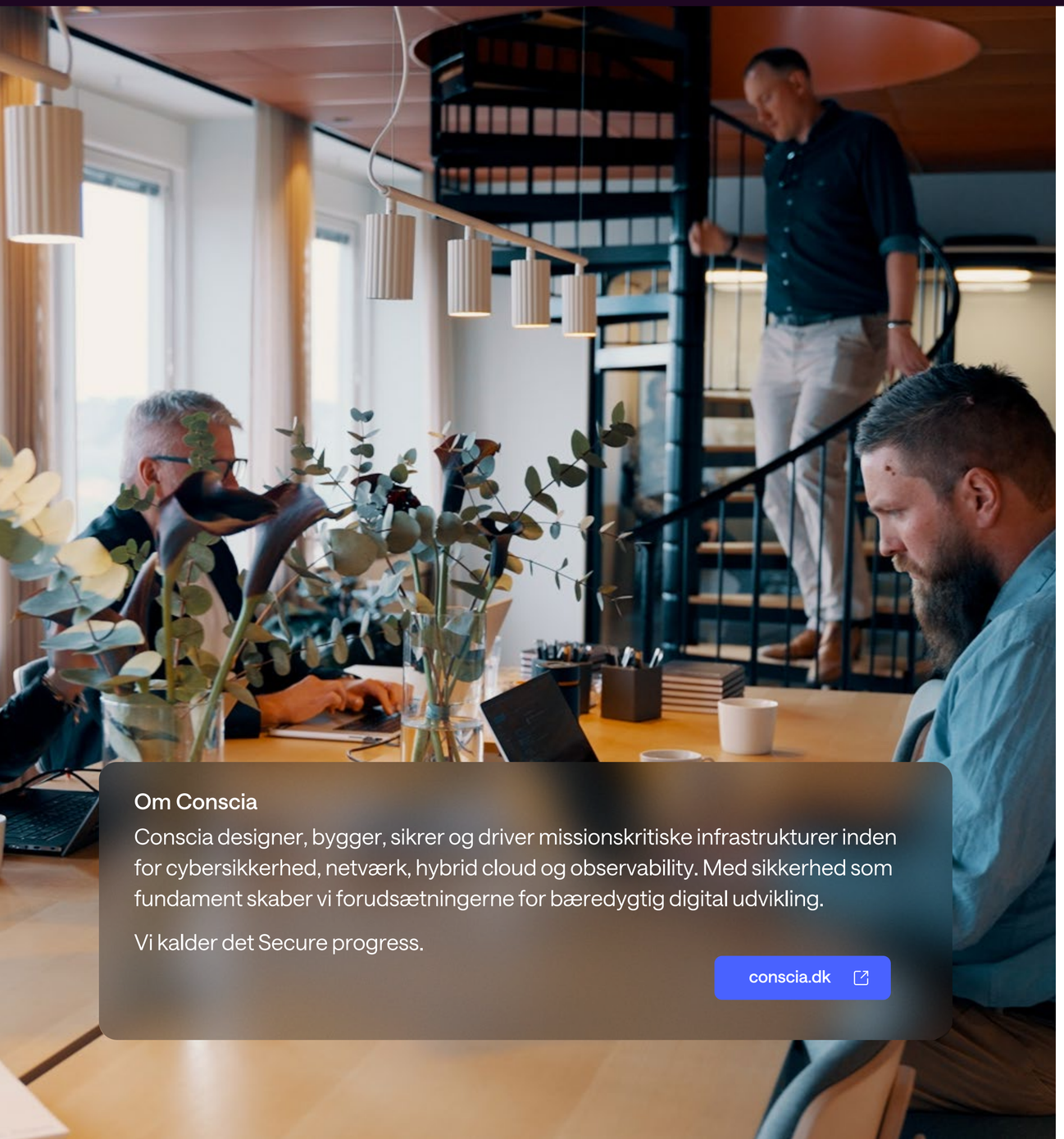
Conscia
Secure progress

Sikker AI 2026

Effektiv og fremtidssikret AI



Om Conscia

Conscia designer, bygger, sikrer og driver missionskritiske infrastrukturer inden for cybersikkerhed, netværk, hybrid cloud og observability. Med sikkerhed som fundament skaber vi forudsætningerne for bæredygtig digital udvikling.

Vi kalder det Secure progress.

conscia.dk

Rapportens indhold

Introduktion til AI **3**

Begreber & definitioner **4**

Centrale begreber kort fortalt 5

AI-udviklingen kort fortalt **6**

En teknisk udvikling 7

Agentic AI & den digitale kollega 8

En organisations AI-modenhed **9**

Modenhedsmodel for AI 9

De seks byggesten i AI-modenhed 10

Nye risici & trusler **11**

Når AI bruges imod os 12

AI i sikkerhedsdomænet **13**

AI som angrebsværktøj 14

Beskyttelse af AI - tekniske foranstaltninger i den rette rækkefølge 15

Ledelse, governance & ansvarlighed **18**

Menneskeligt ansvar 19

Governance i praksis: hvordan AI styres i den daglige drift 20

En sikker vej frem **23**

Konklusion **24**

Eksperterne og kilder **25**

Introduktion til AI

AI påvirker hele organisationen: fra beslutningstagning og risikostyring til innovationsevne og måden, hvorpå tillid opbygges – både internt og eksternt.

Sikker, effektiv og fremtidssikret AI

Kunstig intelligens (AI) er på kort tid gået fra at være et eksperimentelt værktøj til at blive en central del af, hvordan organisationer skaber værdi, udvikler produkter og træffer beslutninger. AI har bevæget sig ud af teknologiens domæne og er blevet et strategisk ledelsesanliggende.

Det høje udviklingstempo har fundamentalt ændret landskabet. Spørgsmål, der tidligere handlede om muligheder og potentiale, handler nu i lige så høj grad om ansvar, tillid og kontrol. Hvordan organisationer introducerer, styrer og anvender teknologien, vil i høj grad afgøre, om AI bliver en kilde til langsigtet værdi eller til usikkerhed og utilsigtede konsekvenser.

Mange organisationer bevæger sig nu fra nysgerrighed til ansvarlighed og fra eksperimenter til reel implementering. For de fleste er spørgsmålet ikke længere, om AI skal anvendes, men hvordan. Hvordan sikrer en organisation, at brugen af AI forbliver bæredygtig og sikker over tid? Hvordan bevarer man agilitet og innovation uden at miste kontrollen? Det kræver et helhedsorienteret perspektiv, der samler teknologi, governance, kompetencer og kultur.

Samtidig accelererer udviklingen hurtigere end nogensinde før. AI-systemer bliver mere autonome, dybere integreret i kerneprocesser og anvendes af stadigt flere mennesker. Parallelt stiger regulatoriske krav og forventninger fra kunder, partnere og samfundet generelt.

I dette miljø er teknisk ekspertise eller isolerede initiativer ikke længere nok. AI skal håndteres som den strategiske kapabilitet, det reelt er.

Om rapporten

Denne rapport henvender sig til beslutningstagere, tekniske ledere og forretningsudviklere, som ønsker at forstå, hvad AI betyder for deres organisation – ud over hypen og de enkelte teknologiske løsninger.

> *Udgangspunktet er, at AI ikke primært handler om innovation eller sikkerhed, men om governance, ansvar og langsigtet kapabilitet.*

Formålet er at vise, hvordan organisationer kan kombinere hurtig teknologisk udvikling med tydelig kontrol, og hvordan en struktureret og ansvarlig tilgang til AI kan blive en konkurrencefordel frem for en risiko. Rapporten begynder med grundlæggende begreber og udviklingen inden for AI og bevæger sig derefter videre til spørgsmål om modenhed, risiko, sikkerhed, ledelse og sikre strategiske valg for fremtiden.

Begreber & definitioner

I dette afsnit defineres og forklares centrale begreber, herunder hvordan de relaterer sig til forretningsværdi, risiko og governance. Hvad er en AI-model? Hvad menes der med en agent? Og hvorfor er det ikke længere tilstrækkeligt kun at tale om AI-chat og individuelle værktøjer?

For at anvende AI ansvarligt og kontrolleret skal teknologien forstås som et sammenhængende økosystem, hvor modeller, data, agenter, integrationer og rettigheder interagerer – og hvor hver enkelt del påvirker sikkerhed, compliance og operationel påvirkning.

Fælles forståelse tydeliggør ansvar

AI kan betyde forskellige ting for forskellige mennesker. I mange organisationer bruges begrebet om alt fra chatbots og assistenter til avancerede systemer, der analyserer information, træffer beslutninger og handler direkte i driften. Uden en fælles forståelse af, hvad AI faktisk er, bliver det vanskeligt at styre brugen, vurdere risici og tydeliggøre ansvar.

AI har bevæget sig fra eksperimenter i isolerede teams til at blive en naturlig del af hverdagens værktøjer – fra kontorstøtte og udviklingsmiljøer til analyse- og sikkerhedsløsninger. Samtidig betyder AI forskellige ting afhængigt af perspektivet.

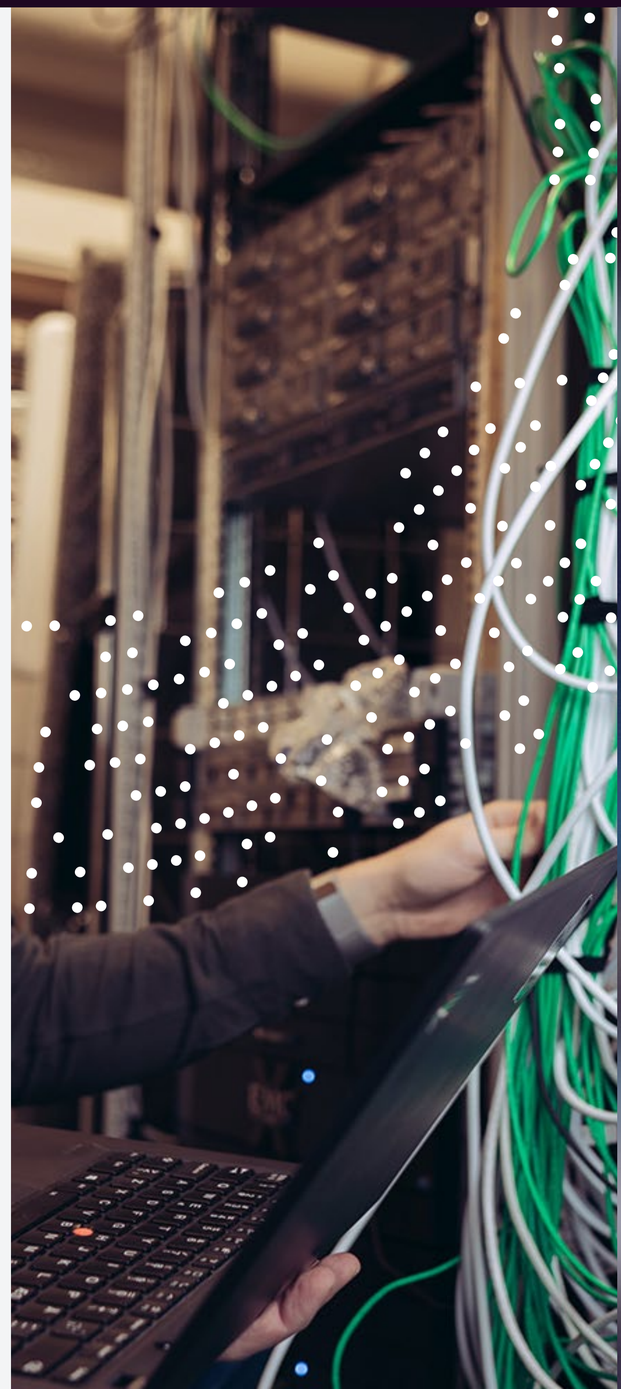
> *I dag er det ikke længere nok kun at tale om prompts og individuelle AI-værktøjer.*

For nogle er det en chatbot; for andre en kompleks infrastruktur af modeller, agenter og integrationer dybt integreret i forretningens kernesystemer.

For at kunne styre, vurdere og beskytte brugen af AI må organisationer betragte AI som et system af sammenhængende komponenter. Først når data, identiteter, modeller, agenter og værktøjer spiller sammen, bliver spørgsmål om ansvar, kontrol og reel påvirkning tydelige.

Centrale begreber kort fortalt

- **Artificial Intelligence (AI)**
  En samlet betegnelse for teknologier, der gør det muligt for computere at analysere, ræsonnere, skabe og træffe beslutninger på måder, der minder om menneskelig intelligens.
- **Agent Identity / Non-Human Identity (NHI)**
  En separat identitet til AI-agenter, som muliggør autentificering, rettighedsstyring og sporbar adgang.
- **Agent Orchestration**
  Styrer, hvordan flere AI-agenter samarbejder, fordeler opgaver og koordinerer beslutninger mod fælles mål.
- **AI-agent (agentic AI)**
  Autonom AI, der kan planlægge, træffe beslutninger og handle over flere trin, ofte i samspil med mennesker og andre systemer.
- **AI governance**
  Rammer og roller, der styrer, hvordan AI udvikles, anvendes og overvåges med fokus på ansvarlighed, risiko og transparens.
- **AI Gateways**
  Et centralt lag til governance, overvågning og policykontrol af adgang til AI-modeller og AI-agenter.
- **AI Management System (AIMS)**
  Et ledelsessystem for AI governance, der er tilpasset standarder som ISO/IEC 42001 og integrerer politikker, risikostyring og løbende forbedringer.
- **AI-model**
  Den trænede algoritme, der fortolker data og producerer resultater, eksempelvis en sprogmodel, billedfortolker eller prediktionsmotor.
- **AI-sikkerhed**
  Tekniske og organisatoriske sikkerhedsforanstaltninger, der skal forhindre manipulation, datalækage og fejl i modeller og agenter.
- **Responsible AI**
  AI, der anvendes etisk, retfærdigt og bæredygtigt med respekt for privatliv, menneskerettigheder og samfundsmæssig påvirkning.
- **Autonomi**
  Graden af selvstændighed i AI – fra menneskeovervågede beslutninger til fuldt autonome agenter.
- **AI-chatbot**
  En AI-drevet applikation, der interagerer med brugere via tekst eller tale, besvarer spørgsmål, giver anbefalinger eller udfører opgaver baseret på input.
- **Bias**
  Systematiske skævheder i data eller modeller, som fører til uretfærdige, upræcise eller biased beslutninger.
- **Bring Your Own AI (BYOAI)**
  Medarbejdere, der anvender egne AI-værktøjer og konti i deres arbejde uden organisatorisk godkendelse eller governance.
- **Data Governance**
  Strukturer og processer, der sikrer, at data er korrekte, beskyttede, sporbare og anvendes på en kontrolleret måde.
- **Deep Learning**
  Avanceret machine learning baseret på neurale netværk, der behandler store datamængder og identificerer komplekse mønstre.
- **Generative AI**
  AI, der skaber nyt indhold såsom tekst, billeder, lyd eller kode baseret på tidligere data og mønstre.
- **Large Language Models (LLM)**
  Store sprogmodeller trænet på omfattende tekstdata, som kan generere, fuldføre og analysere tekst.
- **Machine Learning**
  Et område inden for AI, hvor modeller lærer mønstre fra data og forbedres over tid uden at være eksplicit programmeret.
- **Model Context Protocol (MCP)**
  En åben protokol, der styrer, hvordan AI-agenter får adgang til kontekst, værktøjer og ressourcer på en standardiseret måde.
- **Multimodal AI**
  Modeller, der kan forstå og kombinere flere typer data – såsom tekst, billeder og lyd – for at skabe en samlet forståelse.
- **Prompt**
  En instruktion, et spørgsmål eller en tekst, der indtastes i en AI-model for at styre dens svar eller generering.
- **Traceability**
  Evnen til at følge hele kæden fra data til beslutning og dermed vise, hvordan et AI-output er blevet skabt.



Shadow AI – en voksende realitet

Shadow AI refererer til brugen af AI-værktøjer – såsom generative tjenester, kodeassistenter og analyseplatforme – uden IT-afdelingens viden eller godkendelse. Det er AI, der anvendes uden for organisationens synsfelt. Ikke primært som et brud på reglerne, men som en konsekvens af, at teknologien er blevet tilgængelig og nyttig, før organisationen har etableret de nødvendige strukturer.

Fænomenet minder om klassisk Shadow IT. Forskellen ligger i påvirkningen – AI håndterer ikke kun data, men former også analyser, anbefalinger og handlinger. Når sådanne beslutninger træffes uden synlighed og kontrol, kan konsekvenserne blive større, end organisationen er i stand til at forudse.

De fleste ledelser har i dag begrænset eller ingen indsigt i, hvilken AI der anvendes på tværs af organisationen. Det gør det umuligt at:

- identificere værdifulde initiativer, der bør understøttes.
- etablere de rette beskyttelsesmekanismer omkring de rette data.
- sikre compliance med love og aftaler.
- udvikle en sammenhængende AI-strategi.

For at kunne håndtere AI langsigtet kræves der mere end politikker og retningslinjer. Det kræver en teknisk og organisatorisk kapabilitet bygget op omkring tre sammenhængende trin: Discover, Detect og Protect. Sammen udgør de et framework for praktisk AI governance. (Læs mere i kapitlet AI i sikkerhedsdomænet).

AI-udvikling kort fortalt

Dette kapitel giver et overblik over, hvordan AI har udviklet sig – fra de tidlige machine learning-tilgange til nutidens multimodale og selvlærende agenter. Det, der begyndte som en visionær idé om tænkende maskiner i 1950'erne, er inden for få årtier vokset til en af de mest transformative teknologier i vores tid.

De tidligste pionerer inden for AI, Alan Turing (1912–1954) og John McCarthy (1927–2011), var blandt de første til at definere koncepter for systemer, der kunne ræsonnere og lære. Men datidens teknologiske muligheder var begrænsede – både hvad angår tilgængelige data og regnekraft. I mange år bevægede området sig derfor mellem perioder med optimisme og såkaldte AI-vintre, hvor forventningerne langt oversteg, hvad teknologien faktisk kunne levere.

Skiftet i 2010'erne

Det store vendepunkt kom i 2010'erne. Kombinationen af store datamængder, stigende regnekraft og gennembrud inden for dybe neurale netværk flyttede AI fra teori til praktisk anvendelse. Pludselig kunne systemer genkende billeder, forstå sprog og forudsige adfærd.

På få år blev AI en kernekomponent på tværs af brancher – fra produktion og finans til forskning og kultur. For mange mennesker var dette tidspunktet, hvor AI blev håndgribeligt og reelt nyttigt i hverdagen takket være tilgængelige generative værktøjer, som demonstrerede teknologiens egentlige potentiale.

Den nuværende AI-æra

I dag er AI trådt ind i sin næste fase: den generative og multimodale æra. AI-systemer kan skabe tekst, kode, billeder og andre former for indhold og kan forbindes i autonome agenter, der lærer, planlægger og handler selvstændigt.

Samtidig opstår nye fænomener. Ét eksempel er BYOAI, hvor medarbejdere bruger deres egne AI-værktøjer uden for central governance.

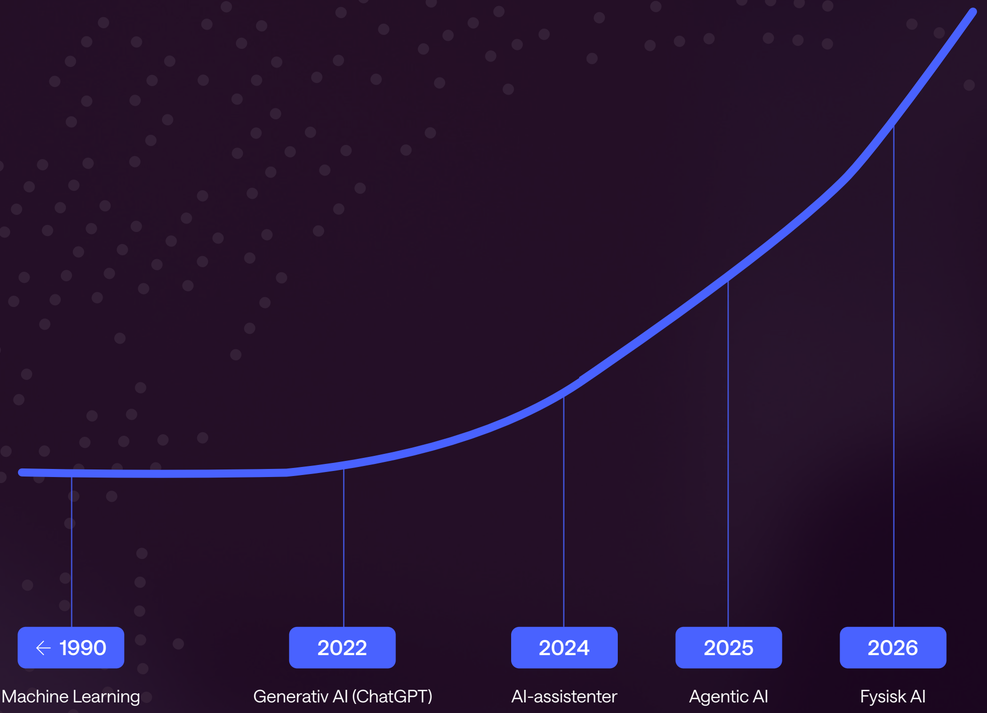
> *Historien gentager sig: teknologien udvikler sig hurtigere end governance*

Denne adfærd er en central drivkraft bag det, der ofte omtales som Shadow AI – den voksende tilstedeværelse af AI-løsninger, som introduceres og anvendes uden for officielle strukturer, politikker og sikkerhedskontroller.

Det er derfor, de følgende kapitler undersøger, hvor godt nutidens organisationer reelt er forberedt på denne nye teknologibølge, hvilke risici og trusler der opstår, samt hvilken rolle ledelse og governance skal spille, når AI bliver en naturlig del af den daglige drift. I dag handler det både om kapabilitet og kontrol – og om hvordan AI kan udvikle sig på en måde, der gavner både organisationen og samfundet som helhed.

Hvad organisationer står over for

Illustrationen nedenfor viser, hvordan udviklingen inden for AI accelererer – fra machine learning til generativ og agentbaseret AI. I dag overstiger innovationstempoet organisationernes evne til at etablere de nødvendige strukturer til at styre teknologien, hvilket understreger det voksende behov for modenhed, governance og bevidste strategiske valg.



En teknisk udvikling

AI bliver ofte opfattet som noget pludseligt – som om teknologien sprang fra simple chatbots til autonome agenter fra den ene dag til den anden. I virkeligheden er nutidens AI-agenter resultatet af en lang teknologisk udvikling, hvor hvert trin har bygget videre på det foregående. For at kunne styre, sikre og anvende AI ansvarligt er det ikke nok at forstå enkelte begreber. Det kræver en forståelse af, hvordan de hænger sammen, hvorfor de opstod, og hvad der har ændret sig undervejs.

Maskinlæring som vendepunkt

Da AI gik fra faste regler til at lære ud fra data

De første AI-systemer blev bygget på hårdkodede regler og statistiske metoder. Gennembruddet kom, da maskiner begyndte at lære mønstre ud fra data i stedet for at blive programmeret trin for trin – et skifte, der blev kendt som Machine Learning. Med Deep Learning kom det næste store spring. Neurale netværk med mange lag gjorde det muligt at identificere komplekse sammenhænge i store datamængder. AI kunne nu genkende billeder, forstå sprog og forudsige adfærd med en præcision, som tidligere havde været uden for rækkevidde. Det var her, AI-modellen blev etableret som kernen: en trænet algoritme, der fortolker data og producerer resultater.

Foundation models & generativ AI I

Da brede, generelle modeller gjorde AI tilgængelig på tværs af organisationer

I lang tid blev modeller trænet til ét specifikt formål ad gangen. Et skifte opstod, da meget store modeller blev trænet på brede datasæt og derefter tilpasset mange forskellige opgaver. Disse blev kendt som foundation models. En særligt vigtig kategori er LLM'er, som arbejder med sprog og kode. De dannede grundlaget for det, der senere blev kendt som generative AI – systemer, der både analyserer og skaber nyt indhold. Det var på dette tidspunkt, AI blev bredt tilgængeligt. Chatinterfaces, kodeassistenter og skriveværktøjer gjorde teknologien håndgribelig på tværs af hele organisationer – ikke kun for udviklingsteams. AI var dog stadig primært reaktiv og reagerede kun, når det blev promptet.

AI i samspil med systemer & data

Da modeller fik kontekst, værktøjer og systemintegrationer

Det næste skridt kom, da modeller ikke længere var isoleret fra deres omgivelser. De begyndte at få kontekst, adgang til dokumenter, data og historik samt mulighed for at anvende værktøjer og API'er. I praksis betød det, at AI-systemer kunne:

- → hente information
- → opdatere systemer
- → udløse handlinger

Det var her, AI-systemet blev vigtigere end selve modellen. Modellen var stadig hjernen, men det omkringliggende system bestemte den reelle påvirkning, AI kunne have på organisationen.

Agentic AI & øget autonomi

Når AI-systemer begynder at træffe beslutninger og igangsætte handlinger

Når modeller kombineres med mål, hukommelse, værktøjer og beslutningslogik, opstår noget nyt: AI-agenter. En AI-agent kan nedbryde en opgave i trin, beslutte den næste handling og udføre den uden, at et menneske styrer processen direkte.

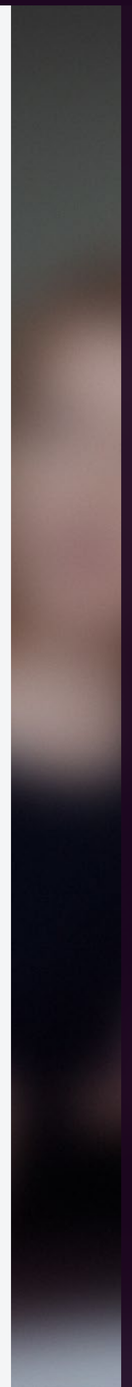
Graden af autonomi kan variere – fra tæt menneskelig overvågning til mere selvstændig drift. Det er her, AI bevæger sig fra at være et understøttende værktøj til at blive en operationel aktør, hvilket fundamentalt ændrer risikolandskabet.

Agentic AI & den digitale kollega

Med agentbaserede systemer får AI en ny rolle i organisationen – som en digital kollega, der kan tage initiativ, træffe beslutninger og interagere med andre systemer. Dette åbner for et betydeligt potentiale inden for effektivitet og innovation, men rejser også nye spørgsmål.

> *Når AI får en aktiv rolle i organisationen, skal den styres ud fra de samme grundlæggende principper som menneskelige roller: et klart formål, definerede rettigheder, mandat og strukturer for tilsyn.*

Sammen danner disse principper en ny måde at tænke AI på i arbejdslivet – hvor AI betragtes som en samarbejdspartner med tydelige krav til governance, opfølgning og løbende udvikling. Organisationer, der begynder at definere rollebeskrivelser for deres digitale kolleger – herunder mandat, formål og tilsyn – opbygger både sikkerhed og tillid. Det er dette, der skaber fundamentet for fremtidens samarbejde mellem mennesker og maskiner.



Krav til en AI-kollega

Tydelig identitet
Hver agent skal have en unik identitet – præcis som enhver anden bruger. Den skal autentificeres, autoriseres, og alle interaktioner skal logges for at sikre sporbarhed og ansvarlighed.

Defineret formål
En agent skal have et klart mandat eller anvendelsesområde, som er koblet til organisationens mål og politikker. Det skal være muligt at forstå, hvorfor den handler – ikke kun hvad den gør.

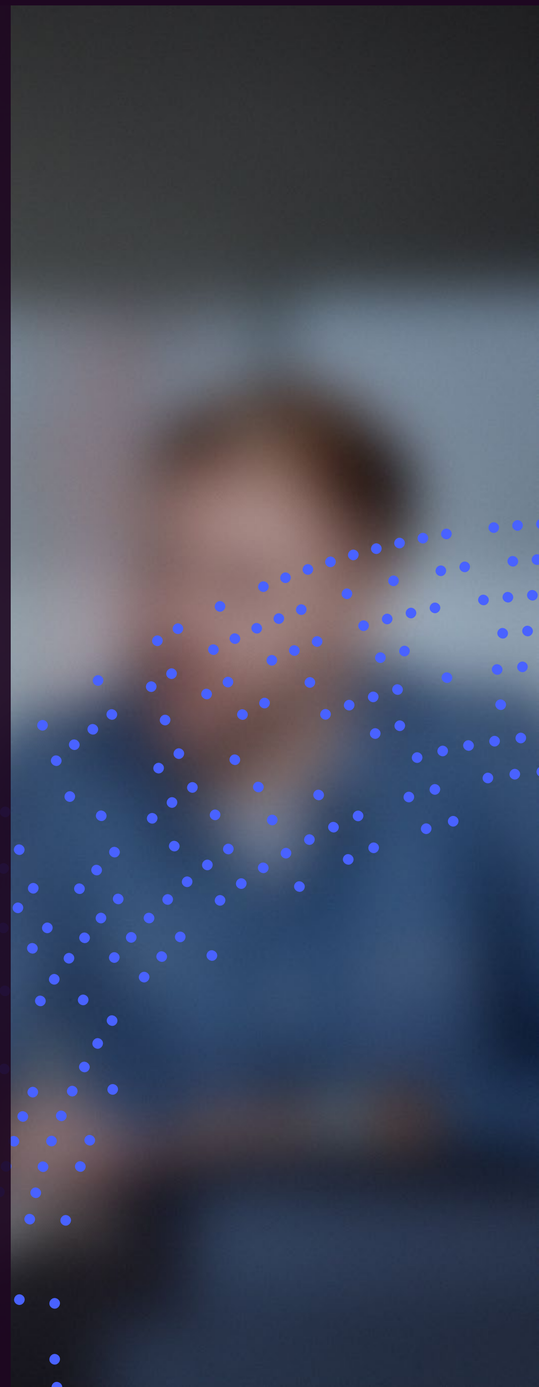
Mandat & begrænsninger
Agenten skal have fastlagte rammer for, hvad den må og ikke må gøre, hvilke systemer den kan få adgang til, hvilke data den må anvende, og hvilke beslutninger den har lov til at igangsætte.

Sporbarhed & auditérbarhed
Alle beslutninger og handlinger skal dokumenteres og kunne gennemgås. Agentens logik og interaktioner skal være tilstrækkeligt transparente til, at de kan forklares efterfølgende.

Menneskeligt tilsyn
Som AI Act kræver (se kapitlet Ledelse, governance & ansvarlighed), skal der være et menneske med ansvar for agentens adfærd og resultater – med beføjelse til at gribe ind eller stoppe den.

Etisk & juridisk alignment
En agent skal handle i overensstemmelse med organisationens værdier og regulatoriske krav – ikke optimere mod et mål uden hensyn til konsekvenserne.

Løbende evaluering
Agentens performance, adfærd og påvirkning skal testes og vurderes kontinuerligt, da dens læring over tid kan ændre måden, den opfører sig på.



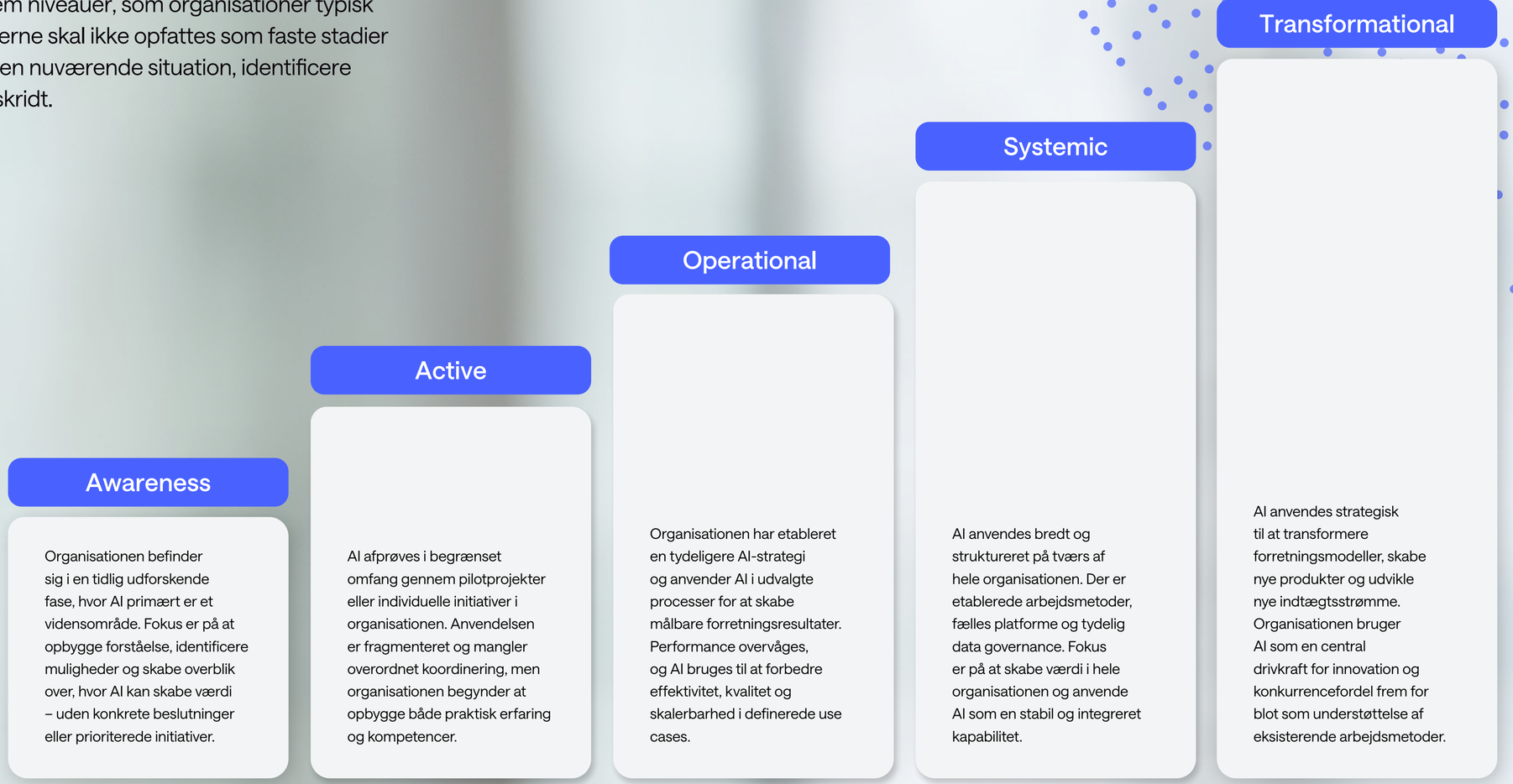
En organisations AI-modenhed

Modenhedsmodellen i dette kapitel beskriver fem niveauer, som organisationer typisk bevæger sig igennem på deres AI-rejse. Niveauerne skal ikke opfattes som faste stadier eller mål i sig selv, men som en måde at forstå den nuværende situation, identificere mangler og skabe fælles retning for det næste skridt.

AI-modenhed handler ikke om, hvor avanceret den teknologi er, som organisationen anvender. Det handler om, hvor godt AI er integreret, styret og forankret i forretningen. Mange organisationer anvender AI i en eller anden form, men forskellen mellem at eksperimentere og at skabe langsigtet værdi er ofte betydelig.

Et højere modenhedsniveau betyder ikke nødvendigvis mere AI, men en bedre balance mellem innovation, governance og ansvarlighed. Organisationer kan befinde sig på forskellige modenhedsniveauer samtidig i forskellige dele af forretningen. Modellen giver et fælles referencepunkt for dialog, prioritering og ledelse.

AI-modenhedsmodel
Fra tidlig forståelse og eksperimentering til systematisk og transformativ anvendelse. Modellen beskriver, hvordan en organisations fokus gradvist bevæger sig fra læring og test til governance, skalering og strategisk påvirkning. Niveauerne afspejler typiske mønstre i, hvordan AI anvendes, styres og skaber værdi over tid.



Kilde: AI Maturity Model (Gartner).

De seks byggesten i AI-modenhed

AI-modenhed handler i sin kerne om at balancere innovation, governance og ansvarlighed. Organisationer, der lykkes med at etablere AI som en langsigtet kapabilitet, gør det ved at opbygge et stabilt fundament af teknologi, processer og struktur.

Erfaring fra både forskning og industri viser, at seks tilbagevendende byggesten går igen i disse organisationer: strategi, infrastruktur, data, governance, kapabilitet og kultur. Sammen udgør de et framework for, hvordan AI kan implementeres, styres og udvikles på en sikker og bæredygtig måde.

De seks byggesten i AI-modenhed

- → Klar AI-strategi
- → Skalerbar og sikker infrastruktur
- → Centraliserede og kvalitetssikrede data
- → Governance med ansvarlighed og transparens
- → Intern AI-kapabilitet
- → Ansvarlig og lærende kultur

Strategy

En tydelig AI-strategi kobler teknologien til organisationens mål og værdier. Den tydeliggør formålet med brugen af AI, prioriterer initiativer og fastlægger rammerne for, hvordan teknologien må anvendes. Når strategien er forankret i ledelsen, bliver det muligt at følge op på resultater og styre innovation i den ønskede retning.

Infrastructure

AI afhænger af en robust og pålidelig infrastruktur. Platforme, netværk og sikkerhedsarkitektur skal kunne håndtere store datamængder, realtidsanalyse og høje krav til tilgængelighed og integritet.
En moderne infrastruktur muliggør udviklingen af AI-løsninger, der er skalerbare og resiliente. Hvis teknologien er motoren, er infrastrukturen vejen.

Data

Datakvalitet, tilgængelighed og sporbarhed afgør, hvor pålidelige AI-modeller bliver. Organisationer, der arbejder systematisk med data governance – fra indsamling til anvendelse og sletning – kan forbedre præcisionen i deres modeller og reducere risikoen for bias og fejlagtige beslutninger. I sidste ende handler datahåndtering om tillid.

Governance

Governance skaber rammerne for ansvarlighed, risikostyring og beslutningstagning. Med tydelige roller, dokumenterede processer og løbende opfølgning kan AI udvikles kontrolleret uden at begrænse innovation. Governance bør derfor ses som en forudsætning for bæredygtig brug af AI.

Capability

AI kræver tværgående kapabiliteter. Ledelsen skal forstå, hvordan teknologien påvirker forretning og risiko, mens tekniske roller skal forstå etiske og regulatoriske krav. Organisationer, der investerer i intern kompetenceudvikling, styrker både beslutningsevne og innovationskraft.

Culture

En moden AI-kultur er kendetegnet ved nysgerrighed, ansvarlighed og transparens. Det skal være lige så naturligt at diskutere risici og begrænsninger som muligheder og forretningsværdi. Det er i mødet mellem teknologi og værdier, at den tillid skabes, som gør AI til en langsigtet del af organisationen.

Klar til at skalere AI?

Mange organisationer befinder sig midt i denne overgang. AI-teknologien er modnet hurtigt, mens governance, ansvar og strukturer stadig er under udvikling. Hvor godt disse forudsætninger etableres, vil afgøre, om AI bliver en bæredygtig del af forretningen – eller fortsætter med at være en kilde til usikkerhed og fragmentering.

AI skaber nye muligheder for analyse, automation og beslutningsstøtte, men risikoen øges, når teknologien anvendes uden en tydelig struktur. I organisationer, der stadig udforsker AI, sker anvendelsen ofte fragmenteret og med begrænset indsigt. Initiativer drives lokalt, ansvaret er uklart, og det bliver vanskeligt at forstå, hvordan AI reelt påvirker forretningen. Risikoen ligger sjældent i teknologien i sig selv, men i manglen på kontrol over data, beslutninger og anvendelse.

> *Organisationer, der er klar til at anvende AI i større skala, har opbygget et stabilt fundament.*

Organisationer, der er klar til at anvende AI i større skala, har opbygget et stabilt fundament. Med en tydelig strategi, governance og et solidt datagrundlag skabes der transparens og ansvarlighed, hvilket gør det muligt at identificere risici tidligt og følge op på, hvordan AI anvendes. Kontrol bliver dermed en muliggører for innovation – ikke en hindring. AI kan herefter skaleres ind i mere forretningskritiske sammenhænge uden at kompromittere tillid eller compliance.

Nye risici & trusler

Dette kapitel belyser, hvordan AI introducerer nye typer trusler, der adskiller sig fra traditionelle IT-risici i både karakter, skala og påvirkning. Fokus er på at forstå dette foranderlige risikolandskab frem for på, hvordan risiciene håndteres i praksis.

Cybersikkerhed har traditionelt handlet om at beskytte systemer og data mod tekniske sårbarheder og brud. AI ændrer grundlæggende dette trusselsbillede. Risici opstår ikke længere kun fra isolerede fejl – de udspringer af, hvordan AI-modeller og agentbaserede systemer opfører sig over tid, hvordan de påvirkes af data og kontekst, og hvordan deres output anvendes på tværs af organisationen.

Dette markerer et skifte fra enkeltstående angreb til mere dynamiske og vanskeligere identificerbare risici. Fokus flyttes fra blot at forhindre intrusioner til at forstå, overvåge og styre, hvordan AI-systemer påvirker beslutninger, processer og tillid. Konsekvenserne bliver i stigende grad operationelle frem for rent tekniske.

Fra sårbar kode til sårbare modeller

AI-systemer er i sagens natur dynamiske. De ændrer sig, efterhånden som de trænes, opdateres og anvendes, hvilket betyder, at sårbarheder ikke længere er statiske. En model, der fungerede korrekt i går, kan opføre sig uforudsigeligt i dag.

Risikolandskabet bevæger sig derfor fra isolerede tekniske fejl til forståelsen af, hvordan modeller påvirkes og udvikler sig over tid.

Truslerne til højre illustrerer, hvordan AI-relaterede risici adskiller sig fra traditionelle IT-sårbarheder, og hvorfor de ofte er sværere at forudsige, opdage og håndtere med etablerede sikkerhedsmetoder.

Prompt injection
Angribere kan manipulere et AI-systems adfærd gennem specialudformede instruktioner. Det kan få modellen til at afsløre information eller handle på måder, der aldrig var tilsigtet.

Data poisoning
Hvis træningsdata manipuleres, kan hele modellens logik forvrænges. Det kan føre til forkerte beslutninger, bias eller gøre AI-systemet sårbart over for angreb efter implementering.

Model leakage
Mange AI-modeller risikerer utilsigtet at afsløre fortrolige oplysninger i deres svar – særligt når den samme model interagerer med flere brugere eller miljøer.

Supply chain-risici
AI udvikles sjældent helt fra bunden. Organisationer er afhængige af præbyggede biblioteker, åbne datakilder og tredjepartsmodeller. En enkelt usikker afhængighed kan derfor sprede risiko gennem hele kæden.

Autonome adfærdsmønstre
Når AI-agenter kan operere selvstændigt, opstår nye angrebsflader. En forkert konfigureret agent kan skabe kaskadeeffekter på tværs af systemer – ofte hurtigere, end et menneske kan nå at gribe ind.

De 10 mest almindelige risici ved LLM-applikationer

Kilde: OWASP LLM Top 10 2025 og Google Forecast 2025

Prompt injection
Manipulation af instruktioner.

Data leakage
Utilsigtet eksponering af følsom information.

Supply chain vulnerability
Usikre tredjepartsmodeller og biblioteker

Insecure output handling
Modeller der genererer skadeligt indhold

Unauthorized code execution
AI der udløser farlige kommandoer.

Overreliance
Blind tillid til AI-output uden verificering

Privacy violations
Utilstrækkelig databeskyttelse i træningsdata.

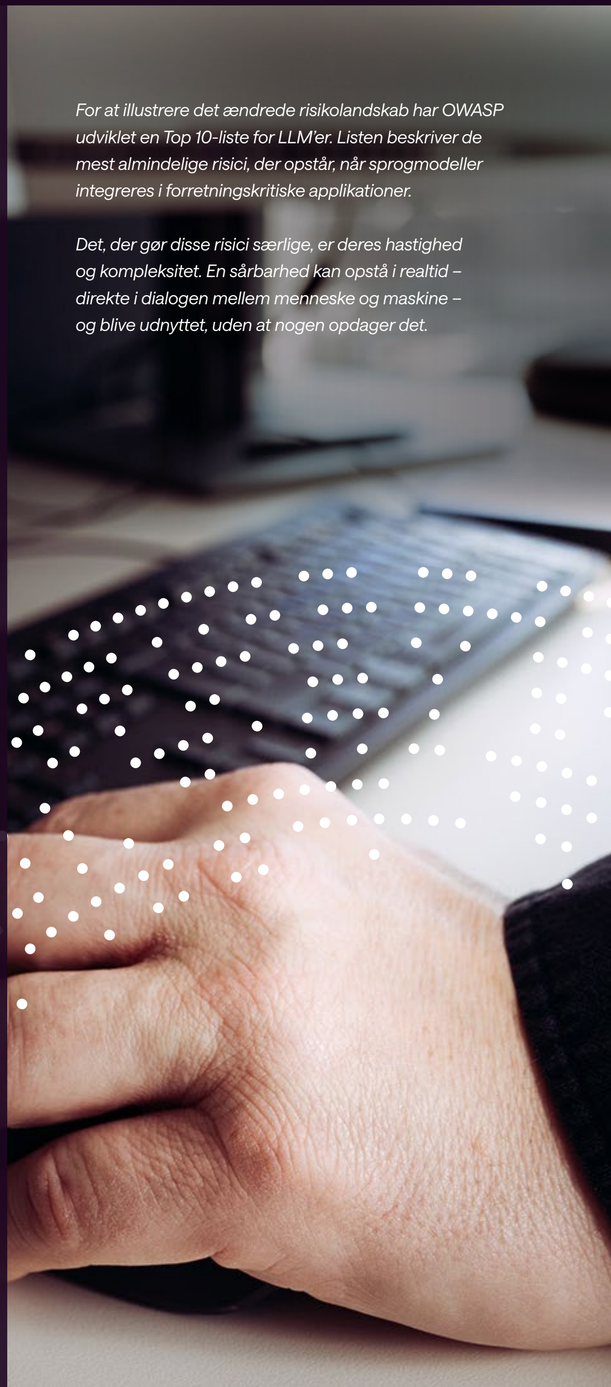
Insecure plugin integration
For mange usikrede integrationer og forbindelser.

Model theft
Tyveri af træningsdata eller modelarkitektur

Monitoring gaps
Manglende sporbarhed og logging.

For at illustrere det ændrede risikolandskab har OWASP udviklet en Top 10-liste for LLM'er. Listen beskriver de mest almindelige risici, der opstår, når sprogmodeller integreres i forretningskritiske applikationer.

Det, der gør disse risici særlige, er deres hastighed og kompleksitet. En sårbarhed kan opstå i realtid – direkte i dialogen mellem menneske og maskine – og blive udnyttet, uden at nogen opdager det.



Når AI bruges imod os

AI's påvirkning rækker langt ud over organisationen selv. Når generative modeller bruges til at producere indhold i stor skala, påvirker de, hvordan information spredes, hvordan virkeligheden opfattes, og hvordan beslutninger træffes. Informationsmiljøet bliver mere komplekst – og i nogle tilfælde bevidst manipuleret.

Manipulation & desinformation
AI bruges også offensivt – fra deepfakes til fuldt genererede informationskampagner. Når grænsen mellem det ægte og det fabrikerede udviskes, bliver konsekvenserne samfundsmæssige. Desinformation kan påvirke den offentlige opinion, markeder og tilliden til institutioner.

> *Tillid er blevet den nye angrebsflade*

For organisationer betyder det, at beskyttelse ikke længere kun handler om data, men også om opfattelsen af, hvem de er, og hvad de står for.

Fejlbehæftede beslutninger & forvrængede analyser
AI kan generere indsigter, der fremstår logiske, men som bygger på upræcise eller biased data. Når sådanne analyser bruges som beslutningsgrundlag, kan konsekvenserne være betydelige – eksempelvis fejlvurderinger af kreditrisiko, produktion eller forsyningskæder. Resultatet kan være både økonomiske tab og tab af tillid.

Dataeksponering & integritet
Generative AI-værktøjer, der anvendes uden tilstrækkelig kontrol, kan utilsigtet afsløre forretningshemmeligheder, kundedata eller kildekode. Hændelsesrapporter viser, at utilsigtet datadeling gennem AI-værktøjer er en af de førende årsager til informationslækager.

Tillidskriser
Når kunder eller samarbejdspartnere oplever, at AI-beslutninger er uretfærdige, uigennemsigtige eller forkerte, kan skaden på brandet være større end selve hændelsen. Tillid er ofte langt sværere at genopbygge end data.

Regulatoriske konsekvenser
Når AI anvendes i beslutninger med reelle konsekvenser, opstår der et tydeligt ansvar. Reguleringer som EU AI Act stiller krav om dokumentation, risikovurderinger og menneskeligt tilsyn i systemer, der klassificeres som højrisiko. Organisationer, der ikke kan dokumentere, hvordan en beslutning er blevet truffet, risikerer både sanktioner og tab af tillid.

Disse risici er ikke kun teoretiske. Organisationer bliver allerede i dag holdt ansvarlige for AI-drevne beslutninger. Det viser, at ansvar ikke kan udskydes til fremtiden, men skal indbygges fra starten.

Det autonome skifte: når risici bliver operationelle

Når AI-systemer bevæger sig fra at analysere til at handle selvstændigt, opstår nye muligheder. Agentbaserede systemer kan udføre opgaver, træffe beslutninger og interagere med andre systemer i realtid, hvilket skaber muligheder for øget effektivitet, automatisering og nye forretningsmodeller.

> *Agentic AI skaber derfor et nyt behov for governance*

Hver agent har brug for en defineret identitet, en rolle og et formål – på mange måder tilsvarende en menneskelig kollega. De skal kunne identificeres, autoriseres og overvåges. I praksis betyder det, at Zero Trust-princippet også skal gælde for autonome systemer og AI-agenter.

Fra reaktion til resiliens

De nye risici kan virke omfattende, men de følger den samme grundlæggende logik som tidligere trusler: Det, der ikke kan ses, kan ikke beskyttes. For at være forberedt skal organisationer opbygge resiliens frem for blot at reagere. Fire centrale og tilbagevendende principper er afgørende:

- **Visibility.** At vide, hvor AI anvendes, og hvordan det påvirker beslutninger.
- **Understanding.** At kunne forklare modellernes logik og begrænsninger.
- **Protection.** At etablere kontroller for data, adgang og adfærd.
- **Learning.** At evaluere og forbedre i takt med den teknologiske udvikling.

Når disse principper bliver en naturlig del af den daglige praksis, kan AI håndteres med samme disciplin som andre forretningskritiske funktioner. Målet er ikke at eliminere risici, men at forstå og håndtere dem på en struktureret og professionel måde.

> Når AI implementeres korrekt, kan det både øge hastigheden og kvaliteten af sikkerhedsoperationer og frigøre tid til det, mennesker er bedst til: at vurdere kontekst, risiko og konsekvens.

AI i sikkerhedsdomænet

AI kan fungere både som en beskyttelse og som en trussel. Dette kapitel forklarer, hvordan AI styrker sikkerhedsarbejdet, men også hvordan teknologien anvendes i angreb – fra automatiseret detektion til manipulation og agentdrevne intrusioner. Formålet er at forstå teknologiens styrke – og det ansvar, der følger med.

AI er hurtigt blevet en integreret del af sikkerhedsoperationer. Modeller analyserer logs, prioriterer alarmer og understøtter analytikere i komplekse undersøgelser. Samtidig bruger angribere den samme teknologi til at skalere phishingkampagner, generere deepfakes og identificere nye sårbarheder.

> *På den defensive side bruges AI til at effektivisere og skalere sikkerhedsarbejdet.*

Flere globale rapporter beskriver dette som anden fase af AI i sikkerhed: Teknologien er ikke længere eksperimentel, men fungerer både som forsvar og angreb.

Det aflaster et presset SOC, løfter kvaliteten af beslutningstagning og gør avancerede analyser mere tilgængelige. Samtidig betyder skiftet mod semi-autonome sikkerhedsoperationer, at systemer overtager mere af det repetitive arbejde, mens mennesker fastsætter rammerne og træffer de kritiske beslutninger.

AI anvendes blandt andet til at:

- filtrere store mængder alarmer og fremhæve det, der reelt kræver handling.
- opdage anomalier i identitets- og adgangsmønstre.
- identificere deepfakes og manipuleret indhold på tværs af kommunikationskanaler.
- levere AI-drevet understøttelse til alt fra threat hunting til rapportskrivning.

Når AI implementeres korrekt, kan det både øge hastigheden og kvaliteten af sikkerhedsoperationer og frigøre tid til det, mennesker er bedst til: at vurdere kontekst, risiko og konsekvens.

AI som angrebsværktøj

De samme egenskaber, der gør AI attraktivt for forsvarssiden, gør teknologien lige så attraktiv for angribere. Google og Microsoft beskriver, hvordan trusselsaktører anvender generative modeller til at skabe mere overbevisende phishing-, vishing- og SMS-svindel, ofte tilpasset den enkelte modtagers sprog, tone og kontekst. Det gør det muligt for angribere at skalere deres operationer både i omfang og præcision. Rapporter peger på flere tydelige mønstre:

- **Skaleret social engineering**
  AI gør det muligt at masseproducere personlige beskeder og falske profiler, herunder syntetiske stemmer og video designet til at omgå KYC*-kontroller og identitetsprocesser.
- **Hurtigere angrebskæder**
  Modeller bruges til at skrive og forbedre kode, udføre sårbarhedsanalyser og automatisere dele af intrusion- og exfiltration-flows.
- **Lavere adgangsbarriere**
  Adgang til ubeskyttede LLM'er på tvivlsomme fora gør det muligt for mindre erfarne aktører at udføre angreb, som tidligere krævede avanceret teknisk ekspertise.

Resultatet er et landskab, hvor traditionelle forsvarsmekanismer overvældes af flere, mere troværdige og hurtigere angreb, mens ansvaret for at skelne mellem ægte og fabrikeret indhold i stigende grad flyttes til modtageren.

Nye angrebsflader i AI-drevne miljøer

AI introducerer også sine egne sårbarheder. Kortlægning af cloudmiljøer viser, at mere end 85 % af alle organisationer allerede anvender en form for AI-tjeneste – ofte bygget på unge kodebaser, hurtige udviklingscyklusser og utilstrækkelige standarder. Det inkluderer eksponerede AI-databaser med logs og nøgler, sårbare GPU-miljøer og løst distribuerede frameworks, hvor én enkelt fejl kan give fuld kontrol over den underliggende infrastruktur.

> *Med Agentic AI opstår et yderligere risikolag*

Systemer, der selvstændigt kan igangsætte handlinger, interagere med eksterne systemer og træffe beslutninger over flere trin, kan skabe hurtige dominoeffekter, hvis de får for stor autonomi uden tilstrækkelige grænser og tilsyn – nogle gange uden at en angriber behøver at opnå adgang på traditionel vis.

* KYC (Know Your Customer) er en løbende proces, der skal forhindre svindel gennem kundekendskab, due diligence og kontinuerlig overvågning.

Hvad adskiller de modne organisationer fra resten?

Flere globale undersøgelser peger på den samme konklusion: De fleste organisationer anvender AI, men kun få er reelt forberedte på at gøre det sikkert. Accenture viser, at omkring tre fjerdedele mangler grundlæggende frameworks for data- og AI-sikkerhed, samtidig med at generativ AI øger både hastigheden og kompleksiteten af trusler. Cisco beskriver, hvordan en mindre gruppe frontløbere kombinerer strategi, kontroller og teknisk kapabilitet på tværs af hele AI-livscyklussen.

Organisationer, der er foran, er kendetegnet ved at:

- behandle AI-sikkerhed som en integreret del af cybersikkerhed.
- have indsigt i, hvilke AI-tjenester og modeller der anvendes, herunder BYOAI, hvilket hjælper med at minimere Shadow AI.
- anvende end-to-end-kryptering, granulær adgangskontrol og overvågning (inklusive af agentbaserede systemer).
- arbejde systematisk med AI-specifik threat modelling og test mod kendte LLM-risici.

Forskellen ligger derfor ikke i, om AI anvendes, men i hvor bevidst det anvendes – og hvor tydeligt governance, arkitektur og kapabilitet er alignet.

Beskyttelse af AI – tekniske foranstaltninger i den rette rækkefølge

Når organisationer forstår, hvor risiciene ligger, skal denne indsigt omsættes til konkrete beskyttelsesforanstaltninger. Det er her, teknologi og governance mødes. Beskyttelse af AI handler om at skabe synlighed, opdage anomalier og kunne handle præcist – ikke om blot at bygge højere mure.

AI ændrer logikken for, hvordan beskyttelse opbygges. AI-sikkerhed skal omfatte data, modeller, beslutninger og adfærd. Beskyttelsen skal samtidig være dynamisk, da AI-systemer udvikler sig over tid. For at undgå fragmenterede og reaktive tiltag kræves en tydelig rækkefølge, hvor hvert trin bygger videre på det foregående.

Discover: få indsigt i, hvor AI faktisk anvendes

Det første skridt mod kontrol er at forstå, hvor AI faktisk bliver anvendt. Uden denne indsigt er det umuligt at skabe governance. De fleste organisationer bruger allerede AI i en eller anden form, men få har et komplet overblik over, hvor teknologien anvendes, hvilke data den håndterer, og hvilke beslutninger den påvirker.

Kortlægning på tværs af organisationen bør dække tre niveauer:

- **Systemniveau.** Identificering af alle AI-relaterede systemer og tjenester – både interne og eksterne. Det inkluderer standalone AI-løsninger samt indbygget funktionalitet i eksisterende applikationer. Mange organisationer opdager, at AI allerede er til stede i deres daglige systemer uden nogen formel beslutning om det.
- **Dataniveau.** Forståelse af hvilke data AI anvender, hvor de kommer fra, og hvordan de bevæger sig. Organisationer har brug for indsigt i både træningsdata og produktionsdata – herunder klassifikation, opbevaringsperiode og adgang. Det er her, mange af de største risici opstår, såsom bias, eksponering og utilstrækkeligt samtykke.
- **Organisatorisk niveau.** Forståelse af hvordan brugen af AI påvirker processer, beslutninger og forretningsfunktioner. Synlighed handler dels om teknisk inventar, dels om at vide, hvor AI påvirker organisationen, og hvem der bærer ansvaret.

> *Sporbarhed er nødvendig gennem hele kæden: fra datakilde til model og videre til beslutning*

Når dette overblik er etableret, bliver synlighed en del af den løbende governance frem for en engangsøvelse. Det er her, teknologi, juridiske frameworks og governance mødes. Det skaber fundamentet for næste skridt: at kunne opdage, når AI-systemer begynder at opføre sig på uventede måder.

Detect: forståelse & identifikation af anomalier

Når kortlægningen er på plads, er næste skridt at forstå adfærd. Detektion af risici i AI-systemer adskiller sig fra traditionel overvågning, fordi trusler ikke altid viser sig i logs eller netværkstrafik. I stedet opstår de i modellens beslutninger, i dens fortolkning af data eller i måden, AI-systemer interagerer over tid.

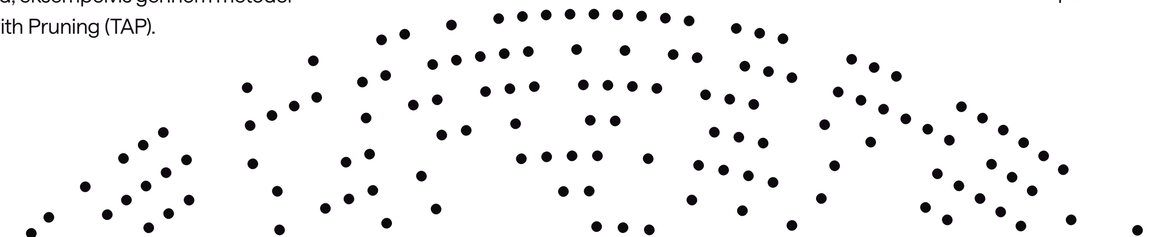
Det gør detektion til et spørgsmål om indsigt snarere end advarsler og alarmer. En model, der begynder at producere usædvanlige outputs, kan lige så vel indikere dårlig datakvalitet som manipulation, og en AI-agent kan operere inden for sit definerede scope og alligevel skabe utilsigtede konsekvenser. At identificere sådanne mønstre kræver både tekniske værktøjer og analytisk forståelse.

En moderne tilgang til detektion bygger på tre principper:

- **Kontinuerlig test.** AI-modeller skal testes gennem hele deres livscyklus – ikke kun under udvikling. Ved at validere modeller kan organisationer identificere anomalier, bias, hallucinationer eller data poisoning, før de skaber konsekvenser. Automatiserede testmiljøer bruges i stigende grad til at simulere angreb og vurdere modellernes robusthed, eksempelvis gennem metoder som Tree of Attacks with Pruning (TAP).
- **Overvågning af adfærd.** Efterhånden som AI bliver mere autonomt, er det ikke længere nok at overvåge systemer som helhed. Hver agent skal behandles som en identitet med en defineret rolle, opgave og adgangsniveau. Det betyder, at de samme principper, der gælder for mennesker – identitet, autentificering og adgangskontrol – også gælder for software, der handler selvstændigt. Ved at overvåge, hvordan en agent kommunikerer, hvilke data den bruger, og hvilke beslutninger den træffer, kan anomalier identificeres i realtid.
- **Samarbejde mellem teknologi og forretning.** Detektion er ikke udelukkende en teknisk disciplin. Uventet adfærd opdages ofte først af brugere, analytikere eller beslutningstagere, som observerer, at et AI-systems output ikke passer til konteksten. Derfor skal observationer fra forretningen indsamles, vurderes og analyseres sammen med tekniske indikatorer.

> *En effektiv detektionsstrategi samler mennesket og maskinen*

Efterhånden som AI-systemer bliver mere komplekse, bliver evnen til at forstå, hvordan "normal" adfærd ser ud, afgørende. Først da bliver det muligt at reagere på anomalier – og tage næste skridt: at opbygge beskyttelse med præcision.



Protect: Beskyttelse med præcision

Når organisationen har opnået synlighed og forståelse for sin brug af AI, kan beskyttelsen etableres de rigtige steder og i den rigtige rækkefølge. Beskyttelse handler ikke om at låse systemer bag firewalls, men om at skabe tydelige rammer for, hvordan AI må handle – med hvilke data, under hvilke betingelser og med hvilket ansvarsniveau.

I klassisk IT-sikkerhed har fokus været på at forhindre intrusioner. I AI-sikkerhed handler det i lige så høj grad om at forhindre uhensigtsmæssig adfærd. En model trænet på følsomme data, en agent der kommunikerer med andre systemer, eller en applikation der genererer kode i produktion, kan skabe skade uden, at en ekstern angriber nogensinde er involveret. Beskyttelsen skal derfor sikre, at AI-systemer forstår deres begrænsninger og forbliver inden for dem.

Effektiv beskyttelse starter med tydelighed. Hver AI-model, tjeneste eller agent skal have en defineret identitet, et formål og et adgangsniveau. Dette er fundamentet for at anvende den samme logik som i den øvrige sikkerhedsarkitektur: Zero Trust. Ingen model, bruger eller proces bør have mere adgang end nødvendigt, og alle interaktioner skal kunne verificeres og logges.

I praksis opbygges beskyttelsen i flere lag, hvor hvert lag adresserer en specifik risikodimension:

- **Formålsbaseret adgang.** AI-systemer bør kun have adgang til de data, der er nødvendige for deres definerede opgave. Eksempelvis bør kundeservicemodeller ikke have adgang til HR-data, og analyseværktøjer bør ikke kunne hente rå information direkte fra produktionssystemer.
- **Segmentering.** AI-miljøer bør adskilles fra resten af infrastrukturen. Ved at isolere trænings-, test- og produktionsmiljøer kan effekten af fejl og angreb begrænses, før dominoeffekter spreder sig.
- **Databeskyttelse & kryptering.** Data, der anvendes, genereres eller lagres af AI, skal krypteres og håndteres med samme disciplin som anden forretningskritisk information. Generative modeller bør suppleres med Data Loss Prevention (DLP)-kontroller for at reducere risikoen for utilsigtet deling af følsomme oplysninger. Dette gælder også intern brug – den mest almindelige datalækage kommer stadig indefra, ikke udefra.
- **Policy enforcement & overvågning.** Beskyttelse skal være dynamisk. Politikker for brug af AI, datadeling og modelopdateringer skal omsættes til tekniske kontroller og kontinuerlig overvågning, så systemet selv kan identificere, når en model handler uden for sit tilsigtede formål.

  Det er her, grænserne mellem sikkerhed og governance bliver tydelige. Når politikker, arkitektur og overvågning er alignet, kan beskyttelse opbygges med præcision. Målet er at gøre AI forudsigeligt og kontrollerbart – ikke at begrænse innovation. Når den balance er på plads, kan AI blive en integreret og pålidelig del af organisationen. Et segmenteret, logget og governet AI-miljø bliver et kontrollerbart aktiv – en del af organisationens struktur frem for et sideløbende eksperiment.

Fra beskyttelse til tillid

Beskyttelse af AI handler om at skabe tydeligere rammer – ikke om at bygge højere mure. Organisationer, der kombinerer synlighed med forståelse og tekniske sikkerhedsforanstaltninger, kan håndtere risici proaktivt og anvende AI med tillid. I AI-æraen handler resiliens i lige så høj grad om indsigt, ansvarlighed og tillid som om oppetid og tilgængelighed. Når beskyttelse bygger på forståelse frem for frygt, bliver teknologien et aktiv frem for en risiko.

Tekniske foranstaltninger alene er dog ikke nok. For at AI kan forblive bæredygtigt over tid, skal struktur, ansvarlighed og kultur udvikle sig sammen. Det er governance, der giver retning til teknologien – og gør det muligt at anvende AI med både styrke og kontrol.

> For at AI bliver et aktiv frem for en risiko, skal organisationer sikre, at menneskelig dømmekraft forbliver det sidste kontrolpunkt.

AI i sikkerhedsdomænet fungerer som en force multiplier i begge retninger. Det kan gøre sikkerhedsoperationer mere præcise, hurtigere og mere tilgængelige, men det kan også styrke trusselsaktører, åbne nye angrebsveje og forstærke eksisterende svagheder. Globale rapporter er tydelige: Adoptionsgraden er høj, men sikkerhed og governance har ikke fulgt med udviklingen.

For at AI skal blive et aktiv frem for en risiko, må organisationer betragte denne dobbelte kraft som et designkrav: at anvende AI offensivt i eget forsvar, beskytte modeller, data og agentbaserede systemer lige så systematisk som andre forretningskritiske aktiver – og sikre, at menneskelig dømmekraft forbliver det sidste kontrolpunkt.

Ledelse, governance & ansvarlighed

Når AI bliver en del af organisationens kerneoperationer, ændrer kravene til governance sig. Dette kapitel undersøger, hvordan organisationer kan lede, styre og tage ansvar for AI – fra strategiske beslutninger til juridiske rammer, governance og kultur.

Efterhånden som AI bevæger sig dybere ind i organisationens kerneprocesser, følger ansvaret med. Det, der tidligere blev betragtet som teknologisk innovation, er nu en integreret del af, hvordan organisationer udvikler sig, træffer beslutninger og opbygger relationer. AI driver effektivitet og produktivitet, men påvirker samtidig risikostyring, etik og tillid. Derfor er samtalen flyttet ud over udviklingsteams og blevet et strategisk anliggende for hele organisationen.

Beslutninger, ansvarlighed & konsekvenser

Hvem bærer ansvaret, når et AI-system foreslår en beslutning, der påvirker mennesker eller økonomi? Hvordan sikres korrekt anvendelse af data? Og hvem vurderer logikken bag systemer, der påvirker kundeoplevelser eller risikovurderinger?

Disse spørgsmål er ikke længere hypotetiske – de er blevet centrale for, hvordan virksomheder leder, udvikler sig og tager ansvar i en tid, hvor teknologi kan træffe beslutninger hurtigere, end mennesker kan nå at reflektere over dem.

Når AI kræver governance

I de senere år har mange organisationer fokuseret på at eksperimentere med AI og udforske potentialet. Men efterhånden som AI integreres i forretningsprocesser og beslutningsflows, ændrer forventningerne sig. Innovation skal ledsages af tydelig governance. Kreativitet skal ikke begrænses – men den kræver rammer.

Governance handler ikke om at modarbejde ny teknologi, men om at sikre, at alle initiativer gennemføres kontrolleret, hvor risici forstås, og beslutninger kan spores.

Fra "Kan vi?" til "Bør vi?"

Mange organisationer mangler stadig et klart overblik over, hvor AI anvendes. Værktøjer introduceres lokalt, data bruges bredt, og ansvaret er fordelt på tværs af funktioner. Resultatet er en voksende skygge af uregistreret brug – Shadow AI – initiativer drevet af gode intentioner, men uden synlighed, governance eller sikkerhedsmæssige beskyttelsesmekanismer.

> *Det er her balancen mellem innovation og kontrol opstår – og hvor ansvaret for AI's fremtid for alvor begynder*

At bevæge sig fra fragmentering til koordinering kræver ledelse, der er villig til at stille spørgsmålet "Bør vi?" frem for blot "Kan vi?".

Organisationer der lykkes, er dem, som ser AI som et fælles anliggende. Her væves forretning, teknologi, sikkerhed og kultur sammen i den samme samtale.

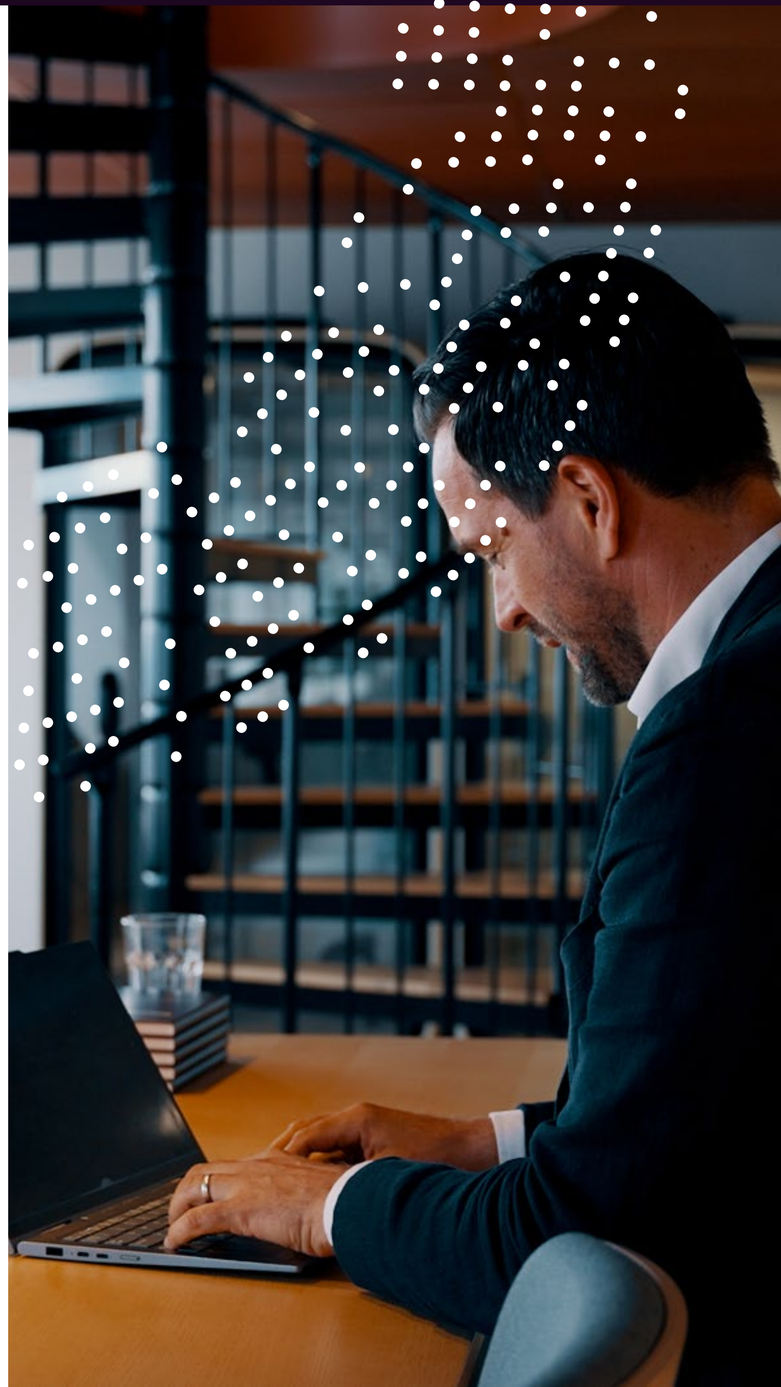
Nutidens lederskab handler både om at lede teknologien og om at forstå, hvordan teknologien påvirker os.

Menneskets evne til at styre AI

Når AI bliver en væsentlig del af beslutningsmaskineriet, må ledelse fungere som dets kompas. Det kræver forståelse for teknologiens muligheder – men også dens begrænsninger.

Organisationer, der kombinerer nysgerrighed med ansvarlighed og teknologi med tillid, vil være bedre rustet til at håndtere risici og samtidig fortsætte udviklingen.

AI ændrer ikke kun, hvordan vi træffer beslutninger – det ændrer også, hvem der træffer dem. I sidste ende handler AI's fremtid om menneskets evne til at styre teknologien i den rigtige retning – ikke om teknologiens kapabiliteter alene.



Menneskeligt ansvar

AI har fået en tydelig plads i bestyrelseslokalet. Teknologien påvirker nu brand, compliance, økonomi og tillid. Den kan ikke længere behandles som et isoleret IT-anliggende. Ligesom cybersikkerhed og bæredygtighed er blevet naturlige fokusområder på ledelsesniveau, er AI nu en central del af strategisk governance.

Forståelse som udgangspunkt

Ansvarlig ledelse begynder med forståelse – erkendelsen af, at AI ikke er en neutral teknologi. Modeller vælges, trænes, overvåges og anvendes af mennesker. Hvert trin i denne kæde former beslutninger, data og relationer. Derfor er det nødvendigt, at organisationer tydeligt definerer ansvar, mandat og roller – præcis som man længe har gjort inden for informationssikkerhed og regulatorisk compliance.

Brugen af AI skal samtidig være sporbar. Ligesom i finansiel rapportering skal det være muligt at forklare, hvordan en AI-drevet beslutning er blevet til: hvilke data der indgik, hvordan modellen fortolkede dem, og hvilke faktorer der blev taget højde for.

> *Sporbarhed er et grundlæggende princip i både god governance og kommende regulering*

Ansvar kan ikke delegeres

Selvom systemer kan analysere og handle, er det altid mennesker, der bærer ansvaret. AI kan ikke holdes juridisk ansvarligt, men det kan skabe konsekvenser, som ledelsen må stå til ansvar for. Denne ansvarsfordeling kan beskrives på tre niveauer:

- **Ledelse.** Fastlægger mål, risikoniveauer og principper for brugen af AI.
- **Forretningsejere.** Sikrer, at teknologien anvendes i overensstemmelse med formål, etik og juridiske krav.
- **Tekniske funktioner.** Implementerer, overvåger og rapporterer risici i praksis.

Samtidig er formelle beslutninger og ansvarsfordelinger ikke tilstrækkelige for at håndtere AI effektivt. Ledelse har brug for nye kapabiliteter for at kunne styre AI: forståelse for hvordan AI påvirker beslutningstagning og risiko, strukturer der fastlægger rammer og ansvarlighed, samt en kultur der understøtter transparens og tillid. Først da kan ansvaret løftes fuldt ud.

Regulering som fælles retning

AI har udviklet sig hurtigere end noget andet teknologisk skifte i moderne tid. På få år er teknologien gået fra forskning og eksperimenter til at påvirke forretningsbeslutninger, offentlig debat og myndighedskontrol. Denne hurtige udvikling har gjort behovet for regulering presserende – ikke for at bremse innovation, men for at skabe tillid og forudsigelighed.

AI Act – et risikobaseret rammeværk
Med EU's AI Act* har Europa taget det første skridt mod en fælles model for ansvarlig anvendelse af AI. Lovgivningen bygger på en risikobaseret logik. Jo større påvirkning et system har på mennesker og samfund, desto højere krav stilles der til transparens, kontrol og menneskeligt tilsyn.

> *Formålet er at skabe balance mellem udvikling og tryghed, så AI kan anvendes bredt – men ikke ukontrolleret.*

I løbet af 2025 så vi de første større retssager om AI-beslutninger. Et AI-baseret rekrutteringssystem blev anklaget for diskrimination, mens en anden model blev kritiseret for fejlagtige kreditvurderinger. Flere domstole, både i EU og USA, fastslog princippet om, at ansvaret altid ligger hos mennesket – ikke algoritmen.

Dette er et af de mest grundlæggende principper i både AI Act og GDPR: at automatiserede beslutninger altid skal kunne gennemgås, forstås og revurderes af et menneske.

Når lovgivning bliver et strategisk anliggende
Dette vil præge hele 2026. Etik, transparens og ansvarlighed bevæger sig op på bestyrelsens agenda som en central del af organisationens risikostyring. Virksomheder, der mangler dokumentation, test og menneskeligt tilsyn, risikerer både bøder og tillidskriser.

AI Act vil påvirke alt fra produktudvikling til data governance og supply chains. Reguleringen påvirker både compliance og måden organisationer opbygger tillid på. Organisationer, der kan dokumentere, at deres AI er sikker, sporbar og etisk, vil opnå en strategisk fordel – særligt i partnerskaber og offentlige udbud.

***AI Act** – ofte omtalt som AI-forordningen eller AI-loven er verdens første omfattende AI-lovgivning. Formålet er at skabe bedre rammer for udvikling og anvendelse af AI.*

Kilde: Europa-Parlamentet.

Governance i praksis: hvordan AI styres i den daglige drift

Når de tekniske sikkerhedsforanstaltninger er på plads, er næste skridt at skabe langsigtet stabilitet. Governance er den struktur, der omdanner midlertidige tiltag til varig sikkerhed og gør AI håndterbart over tid – det er skiftet fra at håndtere AI til aktivt at styre det.

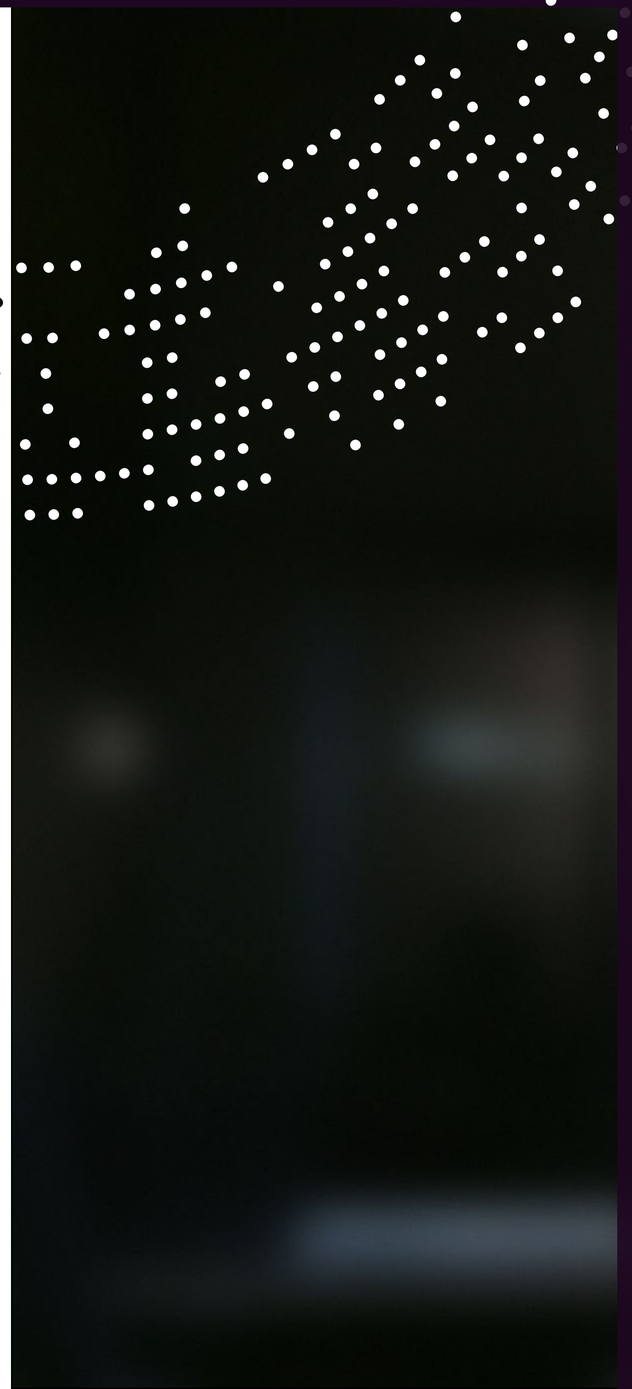
Efterhånden som AI er blevet en del af organisationens kerneoperationer, har kravene til governance også ændret sig. Sikkerhedskontroller eller individuelle politikker er ikke længere tilstrækkelige. Organisationer har brug for et framework, der samler teknologi, data, mennesker og beslutningstagning i ét samlet system. Frameworket skal kunne besvare tre centrale spørgsmål.

> *Hvem må anvende AI – og til hvilket formål?*
>
> *Hvordan overvåges risici, beslutninger og resultater?*
>
> *Hvad gør vi, når noget går galt?*

Governance sikrer, at disse spørgsmål får klare og konsistente svar på tværs af organisationen. Organisationer, der integrerer disse svar i deres strategi, vil være bedre rustet end dem, der blot følger regelbogen. AI governance bliver dermed en måde at styre udviklingen på, hvor ansvarlighed er en integreret del af beslutningstagningen. På lang sigt afgøres tillid ikke af datamængden eller modellernes hastighed, men af hvor ansvarligt teknologien anvendes.

Et sammenhængende framework

Et effektivt framework for AI governance skal være integreret i organisationen. Det skal kunne udvikle sig i takt med organisationen og afspejle både den teknologiske udvikling og de regulatoriske krav, der følger med. Internationale standarder som ISO/IEC 42001 og AI Act skaber et fælles fundament for, hvordan governance kan struktureres.



Fire grundlæggende søjler i effektiv AI governance

- **Policy & retning.** AI-politikker bør beskrive formål, værdier og risikoniveauer samt tydeliggøre ansvar og mandat. En velformuleret politik fungerer som et kompas for hele organisationen.
- **Roller & ansvar.** Alle funktioner, der påvirker AI – fra udvikling og drift til etik og juridiske forhold – skal have et klart defineret mandat. Roller og titler som AI Governance Lead, AI Security Architect og Data Steward spiller en vigtig rolle her.
- **Risikostyring & dokumentation.** AI brug af AI skal være sporbar og forklarbar. Det gælder både beslutninger truffet af modeller og de risikovurderinger, der foretages før implementering. Sporbarhed er afgørende for at kunne demonstrere ansvarlighed.
- **Kontinuerlig forbedring.** Governance er ikke et projekt med en slutdato. AI-systemer udvikler sig over tid, og governance-frameworket skal kunne udvikle sig sammen med dem. Regelmæssige reviews, audits og opdateringer af politikker er afgørende for at opretholde kontrollen.

Sammen udgør disse søjler et governance-system, der minder om dem, man længe har anvendt inden for informationssikkerhed – men med fokus på beslutningstagning frem for alene databeskyttelse.

AIMS som en del af ledelsessystemet

AI governance er på vej til at blive formaliseret som et dedikeret ledelsessystem. Internationale standarder introducerer begrebet Artificial Intelligence Management System (AIMS), som fungerer som et supplement til eksisterende frameworks såsom Information Security Management System (ISMS) inden for informationssikkerhed og Quality Management System (QMS) inden for kvalitet.

AIMS bygger på den samme grundlæggende logik: at planlægge, implementere, følge op og forbedre – men med fokus på livscyklus, etik, risiko og transparens i AI.

For de fleste organisationer betyder dette ikke, at de skal starte forfra, men at de skal udvide deres eksisterende governance til også at omfatte AI. Det er en vigtig pointe. AI governance er et naturligt næste skridt i en udviklingsrejse, som mange organisationer allerede er i gang med – fra informationssikkerhed og databeskyttelse mod ansvarlig teknologisk governance.

På denne måde bliver governance et integreret ledelsessystem, der samler sikkerhed, kvalitet og ansvarlighed.

Hvordan governance fungerer i praksis

I praksis betyder governance, at teknologi, mennesker og processer styres som en helhed. Et velfungerende framework indeholder mekanismer til:

- **Governance af AI-livscyklussen.** Styring af AI fra udvikling til udfasning.
- **Incident management & eskalering.** Tydelige procedurer for, hvad der skal ske, når noget går galt.
- **Etisk review.** Uafhængig vurdering af modeller og use cases.
- **Rapportering & audit.** Løbende opfølgning til ledelse og bestyrelse.

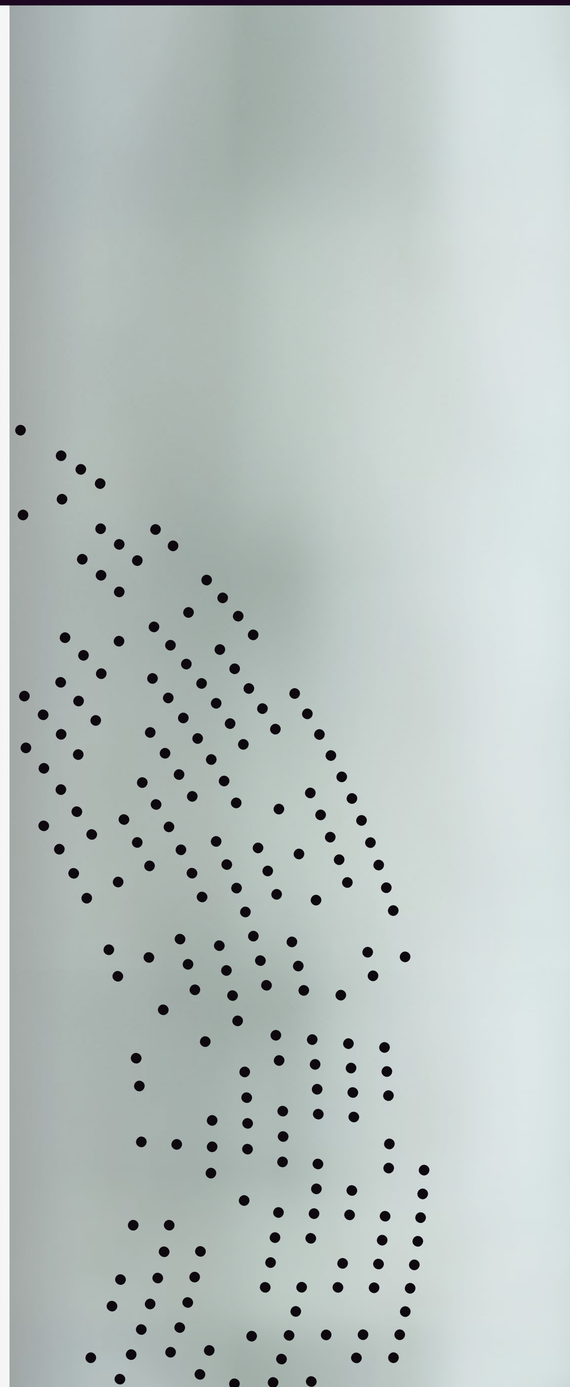
Kravene til transparens, sporbarhed og ansvarlighed er steget i takt med, at AI er blevet mere udbredt, og organisationer, der allerede arbejder med ISO 27001, GDPR eller bæredygtighedsrapportering, kan ofte bygge videre på deres eksisterende strukturer. Governance af AI handler i høj grad om at udvide det samme mindset til en ny teknologisk virkelighed.

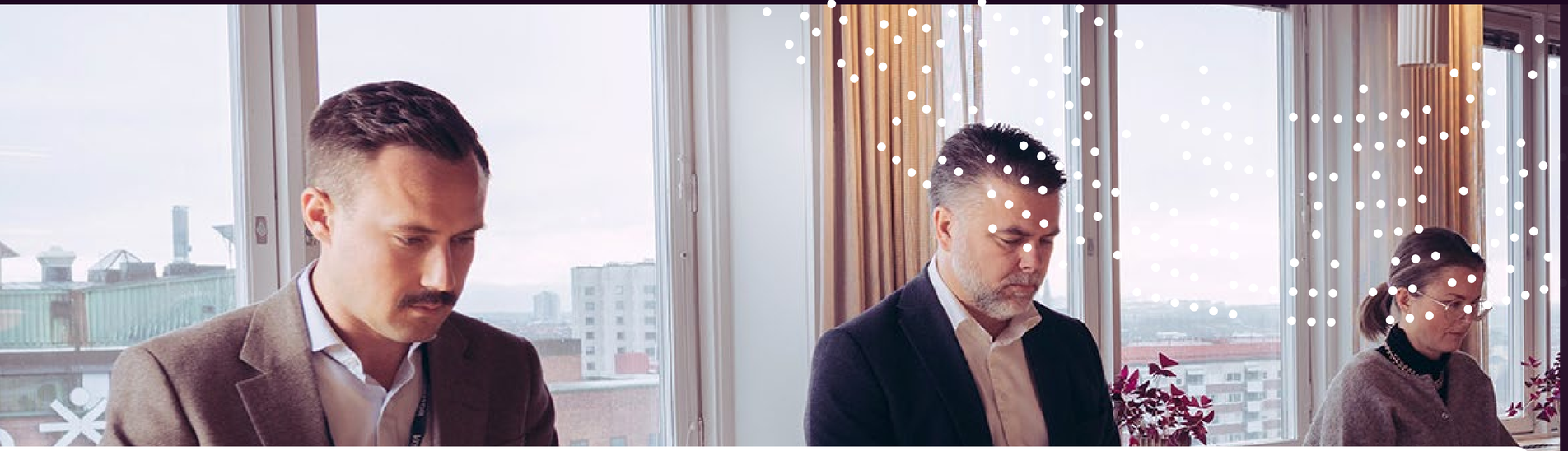
En forbindelse mellem teknologi & tillid

Governance skaber rammerne for, hvordan AI må anvendes, hvordan ansvar fordeles, og hvordan værdi beskyttes. Et stærkt framework sikrer, at innovation sker under kontrol, og at risici håndteres proaktivt frem for reaktivt. Teknologi kan skabe muligheder, men governance skaber tillid.

> *Vejen frem kræver ikke en ny begyndelse. Mange organisationer kan bygge videre på det, der allerede eksisterer.*

At udvide informationssikkerhed til også at omfatte AI-sikkerhed, governance til at inkludere AI governance og etik til at indeholde konkret menneskeligt tilsyn er ikke en genstart. Organisationer, der etablerer tydelige processer for ansvarlighed, styring og forbedring oven på deres eksisterende governance-strukturer, kan fortsætte med at udvikle sig med AI uden at miste kontrollen.





En fælles vidensbase for AI

Ledelse i AI-æraen kræver forståelse for, hvordan teknologien påvirker mennesker, drift og beslutningstagning. For at AI kan styres effektivt, har ledelsesteams og nøglefunktioner brug for den nødvendige viden til at stille de rigtige spørgsmål, vurdere risici og træffe beslutninger med både etik og forretningslogik for øje.

Bestyrelser og direktioner har derfor behov for at opbygge en fælles videnbase om AI – hvordan modeller trænes, hvordan de kan manipuleres, og hvordan beslutningslogik kan påvirkes. Fokus bør være på at kunne føre kvalificerede dialoger med dem, der udvikler og vedligeholder teknologien. Mange organisationer etablerer dedikerede roller til at koordinere dette arbejde:

- → *AI Governance Lead koordinerer policy, risiko og compliance.*
- → *AI Security Architect har ansvar for identitet, adgang og databeskyttelse.*
- → *AI Ethics Officer vurderer påvirkning på mennesker og transparens.*
- → *Data Steward sikrer datakvalitet og sporbarhed.*

Den vigtigste kapabilitet er dog ikke en titel, men et mindset: erkendelsen af, at AI er noget, der skal ledes.

Organisationer, der deler dette syn, flytter AI fra eksperiment til strategi, fra teknologi til ansvarlighed og fra risiko til konkurrencefordel.

Kultur bringer governance til live

Et framework uden kultur bliver hurtigt en papirøvelse. Governance fungerer kun, når det understøttes af en fælles forståelse af, hvorfor det eksisterer.

AI berører hele organisationen – teknologi, jura, kommunikation, HR og forretningen som helhed. Derfor skal governance være fælles forankret frem for ejet af én enkelt funktion.

Kultur er forbindelsen mellem strategi og daglig praksis. Den afgør, om principperne i frameworket omsættes til reel adfærd. Organisationer, der opbygger en kultur præget af ansvarlighed og nysgerrighed, skaber fundamentet for både sikker AI og fortsat innovation.

Kendetegn ved en stærk AI-kultur

Transparens

Medarbejdere skal forstå, hvordan AI anvendes, hvilke data der ligger til grund for beslutninger, og hvor de kan henvende sig med spørgsmål eller bekymringer.

Læring

AI governance kræver løbende kompetenceopbygning. Nye modeller, værktøjer og reguleringer betyder, at træning skal være en kontinuerlig indsats frem for en engangsøvelse.

Tillid

Når governance opleves som tydelig og retfærdig, vokser tilliden. Mennesker bliver mere motiverede til at anvende AI på den rigtige måde og føler sig trygge ved at rapportere afvigelser og bidrage til forbedringer.

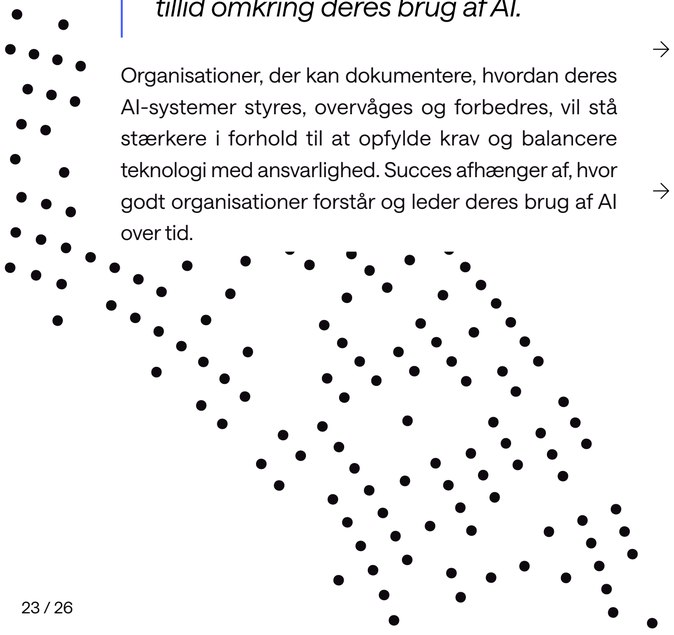
En sikker vej frem

AI bevæger sig ind i en ny fase, hvor teknologien bliver en stadig mere integreret del af organisationers drift. Skiftet mod mere agentbaserede systemer gør ansvarlighed, transparens og governance afgørende. Som følge heraf bliver AI i praksis et spørgsmål om ledelse, sund dømmekraft og tillid.

Reguleringen er begyndt at indhente teknologien, selvom teknologien selv fortsat bevæger sig hurtigere end tidligere forudset. I overgangen fra generative modeller til agentbaserede systemer får AI en mere aktiv rolle i driften og samarbejder med mennesker i langt højere grad. Det betyder, at ansvar og governance bliver lige så vigtigt som innovation. AI berører derfor spørgsmål, der rækker langt ud over teknologi alene.

> *De kommende år vil blive præget af, hvordan organisationer opbygger tillid omkring deres brug af AI.*

Organisationer, der kan dokumentere, hvordan deres AI-systemer styres, overvåges og forbedres, vil stå stærkere i forhold til at opfylde krav og balancere teknologi med ansvarlighed. Succes afhænger af, hvor godt organisationer forstår og leder deres brug af AI over tid.



Fremtidens risikolandskab

Fremtidens risici handler mindre om angreb på systemer og mere om påvirkning af beslutninger, data og tillid. Efterhånden som AI bliver mere integreret i kommunikation, forretningslogik og offentlig debat, udviskes grænserne mellem intern sikkerhed, informationspåvirkning og etisk ansvar. Tre udviklingstendenser står allerede tydeligt frem:

- **Desinformation & fabrikeret påvirkning.** AI anvendes til at forme narrativer og holdninger i et omfang, der tidligere ikke var muligt. Det gør informationspåvirkning til en strategisk risiko – også for organisationer.
- **AI i supply chain.** Organisationer skal vurdere både egne AI-modeller og dem, der anvendes af leverandører og partnere. Sårbarheder i tredjepartssystemer bliver en del af organisationens samlede risikoprofil.
- **Regulering som konkurrenceparameter.** Compliance med reguleringer som AI Act og ISO 42001 bliver en del af organisationers langsigtede konkurrenceevne. Efterhånden som dokumentation og sporbarhed bliver standard, vil tydelig governance skabe stærkere fundament for partnerskaber og offentlige udbud.

Fra ansvar til handling

AI har ændret måden, mennesker arbejder, kommunikerer og træffer beslutninger på. Den største forandring ligger dog i, hvordan teknologien anvendes og integreres i praksis. For organisationer er udfordringen ikke at forstå, hvad AI kan gøre, men hvordan det bør anvendes for at forene innovation med sikkerhed og ansvarlighed.

Erfaringer fra de seneste år viser, at governance, sikkerhed og innovation ikke er modsætninger. De forstærker hinanden. Når struktur, tillid og ansvarlighed bygges ind i teknologien, bliver AI en muliggørende faktor frem for en risiko. Vejen frem defineres af tre centrale skift:

Fra eksperimenter til økosystemer.
AI kan ikke drives som isolerede projekter eller standalone værktøjer. Det skal integreres i organisationens processer, governance og kultur – på samme måde som andre strategiske initiativer.

Fra tilsyn til samarbejde.
Fremtidens AI skal arbejde i samspil med mennesker – ikke blot overvåges. Ved at definere roller, mandat og tilsyn for digitale kollegaer kan sikker autonomi opnås.

Fra regulering til tillid.
Compliance er fundamentet, men ikke målet i sig selv. Organisationer, der kan vise, hvordan deres AI fungerer, hvordan beslutninger kan forklares, og hvordan risici håndteres, vil opnå noget, der ikke kan lovgives frem: tillid.

Et fælles ansvar

AI forandrer mange ting, men ikke det mest grundlæggende: ansvaret ligger altid hos mennesker. Teknologien kan understøtte, foreslå og forudse, men det er mennesker, der beslutter, hvordan den skal anvendes, og hvilke værdier den skal understøtte. Ansvarlig brug af AI bygger på sund dømmekraft i balance med regler og teknologi.

- Det betyder, at organisationer må tage stilling til vanskelige spørgsmål:
- Hvilke beslutninger bør automatiseres, og hvilke bør altid forblive menneskelige?

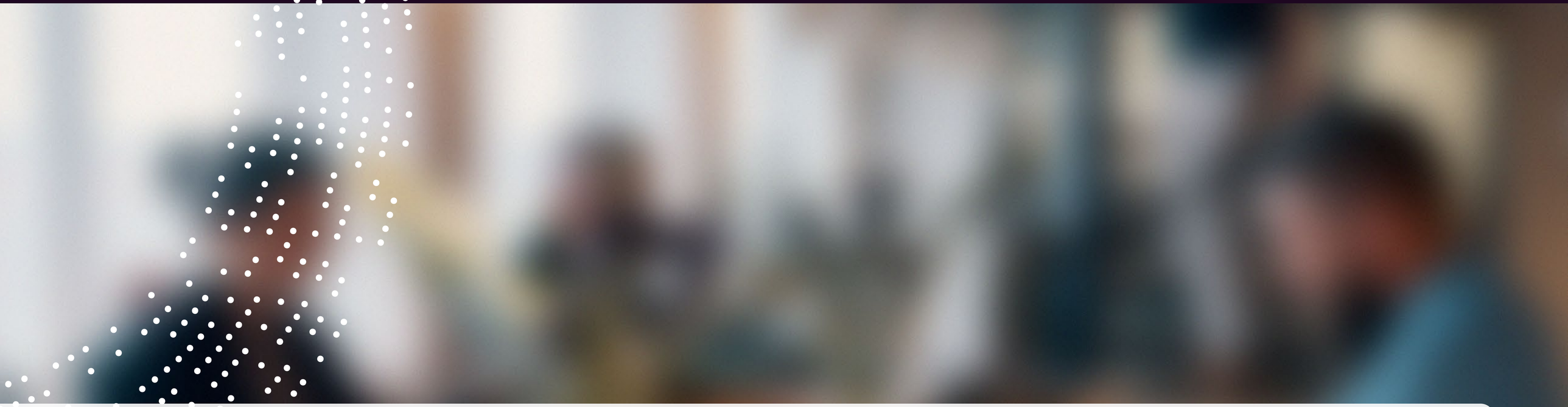
Hvilken rolle skal AI spille i organisationen, og hvordan sikrer vi, at det skaber mere værdi end skade?

Organisationer, der aktivt arbejder med disse spørgsmål, skaber bedre forudsætninger for at håndtere AI på lang sigt. De styrker deres tekniske kapabiliteter og opbygger den tillid, der bliver nødvendig, efterhånden som AI får en større rolle i driften.

At lede AI over tid

Alt dette markerer begyndelsen på en ny æra. AI handler grundlæggende ikke om, hvad teknologien teoretisk kan blive til, men om hvordan den ledes og anvendes i praksis. Succes vil blive målt på, hvor klogt organisationer formår at styre og tage ansvar for AI over tid.

Organisationer, der forener innovation med ansvarlighed og sikkerhed med tillid, skaber bedre forudsætninger for at håndtere fremtidige teknologiske skift – og vil samtidig være med til at forme dem. I sidste ende handler det om at træffe bevidste valg og styre udviklingen med sund dømmekraft. Det handler ikke om at løbe hurtigst, men om at bevæge sig i den rigtige retning.



Konklusion

Ansvaret ligger altid hos mennesker

Artificial intelligence bevæger sig ind i en ny fase. Det, der engang håndteredes som tests, pilotprojekter og isolerede initiativer, former nu, hvordan organisationer træffer beslutninger, driver drift og opbygger tillid. Efterhånden som AI udvikler sig fra generative modeller til mere autonome og agentbaserede systemer, får teknologien en operationel rolle på tværs af organisationen. AI er fortsat teknisk i sin kerne, men er samtidig blevet et strategisk ledelsesanliggende. Det stiller større krav til governance, transparens og ansvarlighed.

Samtidig ændrer risikolandskabet sig. Risici opstår fra tekniske sårbarheder, hvordan beslutninger påvirkes, hvordan data anvendes, og hvordan tillid udvikler sig over tid. Informationspåvirkning, afhængigheder i supply chains og stigende krav til dokumentation og sporbarhed betyder, at sikkerhed må betragtes i et bredere perspektiv. Når AI anvendes til at styrke sikkerhed og samtidig muliggøre mere effektive angreb, bliver evnen til at forstå, overvåge og kontrollere brugen afgørende. I denne virkelighed bliver tillid et strategisk aktiv.

For at omsætte ambitioner til praksis kræves struktur. AI skal integreres i organisationens processer, governance og kultur på samme måde som andre forretningskritiske kapabiliteter. Det kræver tydelige roller, ansvar og metoder til opfølgning samt kompetence og dømmekraft til at afgøre, hvordan teknologien bør anvendes.

I sidste ende handler det om ledelse og governance. Organisationer, der leder AI med sund dømmekraft, skaber grundlaget for bæredygtig udvikling og langsigtet tillid. Den vigtigste indsigt er, at ansvaret altid ligger hos mennesker.

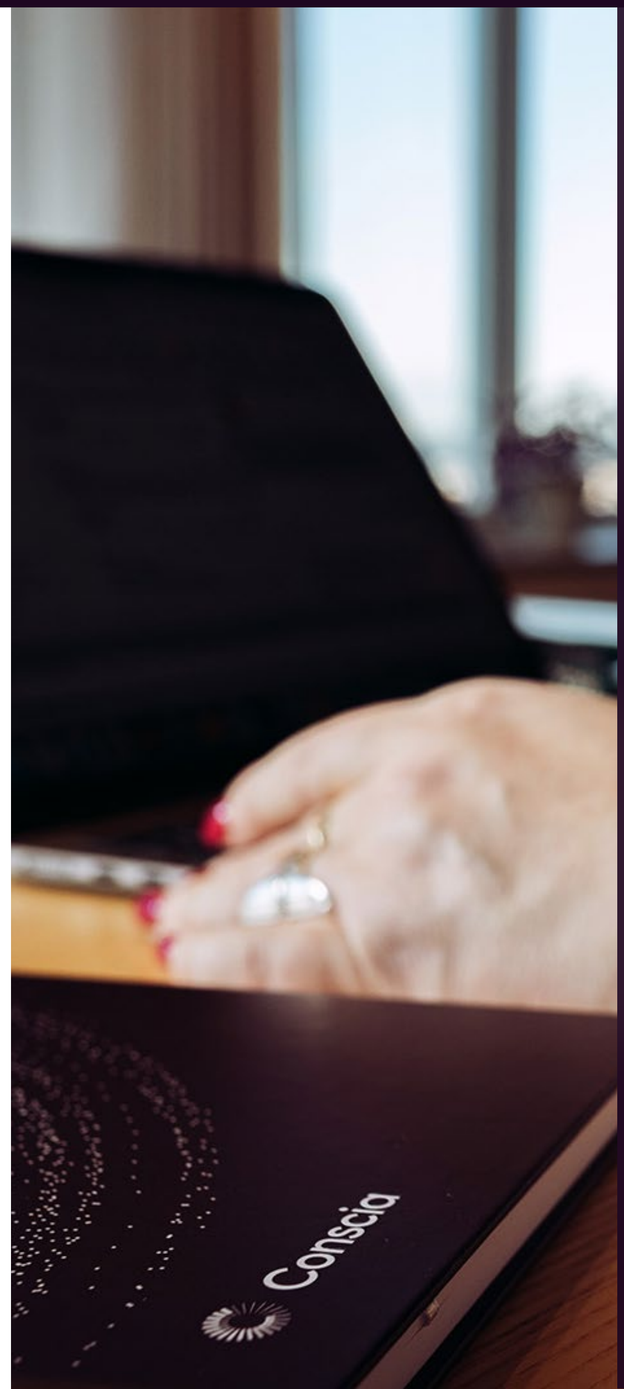
Tal med vores AI- og cybersikkerheds-eksperter

Vil du vide mere om, hvordan AI kan implementeres sikkert og ansvarligt i din organisation?

Vores specialister hjælper virksomheder med AI governance, cybersikkerhed, cloud, observability og moderne IT-infrastruktur – fra strategi og compliance til implementering og drift.

Thomas Grønne
Technical Director, Conscia Danmark
tg@conscia.com

Nicolaj Wichmann
Sales Director, Conscia Danmark
nw@conscia.com



Kilder & referencer

Rapporten bygger på konsolideret viden og erfaring suppleret med eksterne rapporter, analyser og branchedata.

Accenture, *State of Cybersecurity Resilience, 2025*

Cisco, *AI Readiness Index, 2025*

Europaparlamentet, *europarl.europa.eu*

Gartner, *gartner.com*

Google, *Cybersecurity Forecast, 2025*

Microsoft, *Digital Defense Report, 2025*

OWASP, *owasp.org*

Trapets, *trapets.com*

Wiz, *The State of AI in the Cloud, 2025*

Conscia

Secure progress