

# ENTERPRISE

# AI

## Opportunities, Challenges & The C-Suite Imperative

A Strategic Research Report on AI-Powered Digital  
Transformation for Enterprise Leaders

**\$2.59T**

Global AI Spending 2026

**40%**

Enterprise Apps with AI Agents

**12%**

At Full Enterprise Deployment

- Market Overview & Growth Forecasts 2026–2030
- C-Suite Implementation Challenges & Solutions
- AI Vendor Landscape & Selection Framework
- Enterprise AI Platform Comparative Analysis
- AI Investment Tracker – 300+ Deals
- White Paper: Implementing Claude in the Enterprise



**Anuj Khanna**  
CEO, Peak State Consulting  
[anuj@peakstate.com](mailto:anuj@peakstate.com)  
[peakstate.com](http://peakstate.com)

June 2026  
Research Report

## Author's Foreword



### Anuj Khanna

CEO, Peak State Consulting

London, United Kingdom · June 2026

[anuj@peakstate.com](mailto:anuj@peakstate.com) · [peakstate.com](https://peakstate.com)

Once in a generation, a technology arrives that does not merely improve how business is done but rewrites the rules of who wins. We are living through one of those moments, and it is unfolding faster than any that came before it. In the last twelve months, more capital has flowed into artificial intelligence than in the entire previous decade of the industry combined. Q1 2026 alone saw more than \$330 billion in global venture investment, with AI companies capturing roughly 80 percent of it. By May 2026, Anthropic had raised a \$65 billion Series H, reaching a \$965 billion valuation, a figure that exceeds the GDP of many G20 nations. OpenAI raised \$122 billion in a single quarter. The numbers have grown so large they risk losing meaning. But strip away the headline figures and what remains is something more profound: a structural shift in enterprise technology as consequential as the arrival of the internet, and moving at several times the speed.

The prize is real, measurable, and being claimed right now. McKinsey estimates AI could add \$2.6 to \$4.4 trillion in annual economic value across industries through the automation of knowledge work alone. Global AI spending is forecast to reach \$2.59 trillion in 2026 (Gartner), and Peak State Consulting forecasts the total market will reach \$6.40 trillion by 2030. This is no longer a story about promise. Enterprises running AI in production are reporting productivity gains of 26 to 55 percent in targeted workflows. The first true wave of agentic AI, systems that do not merely answer questions but take autonomous action, is delivering results that were purely theoretical eighteen months ago. A global bank now processes thousands of regulatory documents every month with AI, returning hundreds of analyst hours to higher-value work. A software organisation has reached 90 percent developer adoption of AI coding tools and recorded double-digit gains in engineering velocity. These are not pilots or proofs of concept. They are production systems generating measurable, audited returns, today.

The obstacles are just as real, and most of them are not technological. Data quality remains the most stubborn constraint of all: 52% of enterprise professionals name it as their single biggest implementation barrier. Gartner predicts that through 2026 organisations will abandon 60 percent of AI projects unsupported by AI-ready data, an extraordinary attrition rate for any category of

investment. Legacy integration complexity, immature governance, and the sheer difficulty of managing organisational change at the pace AI demands are stalling programmes that enjoy strong executive sponsorship and ample budget. The uncomfortable truth is that the technology is now consistently ahead of the organisation's capacity to absorb it responsibly.

Three dynamics deserve particular attention from every enterprise leader. First, the Chinese AI challenge is far more significant than most Western enterprises yet appreciate. In just twelve months, Chinese providers have surged from under 2 percent of token traffic on OpenRouter, a leading model marketplace, to over 45 percent, a powerful signal of model-routing demand and intensifying price/performance pressure. DeepSeek, Alibaba Qwen and Baidu ERNIE now offer frontier-class performance at prices five to ten times lower than equivalent US models, a combination that is steadily reshaping the price/performance calculus for enterprise AI worldwide.

Second, the talent constraint is becoming structural. The shortage of professionals who combine domain expertise with AI product development, programme management, security engineering, legal compliance and governance skills is not a temporary market imbalance. Universities rarely produce graduates with both deep AI skills and deep domain expertise. The enterprises winning on AI are building hybrid teams deliberately, through long-horizon workforce development rather than reactive hiring. Prediction 08 reflects our conviction that this scarcity will be the single most significant constraint on enterprise AI ambition through 2029.

Third, and most important of all: the distance between AI ambition and AI execution is now the single line that separates the leaders from everyone else. We forecast that 50 percent of enterprise applications will embed AI agents by 2027. The gap that decides winners is not technological, it is organisational. Governance frameworks, data architecture, change-management discipline, and the institutional capacity to move from pilot to production at scale are what divide the organisations extracting durable value from those producing impressive demonstrations and disappointing returns. The technology is no longer the hard part. The organisation is.

A further signal now demands board attention: the private capital story is rapidly becoming a public-market one. Cerebras opened the 2026 listing wave in May, surging 68 percent on its first day. OpenAI and Anthropic have each confidentially filed to go public, and SpaceX, following its merger with xAI, completed the largest IPO in history on 12 June 2026, raising \$75 billion at a \$1.77 trillion offering valuation and trading above \$2 trillion on its first day. Together with Databricks and others, this pipeline represents over \$3 trillion in value, the most AI-concentrated IPO year on record. What public markets are about to reward, durable revenue, disciplined governance and demonstrable returns, is precisely the standard now arriving for every enterprise AI programme.

And there is a new dimension that did not exist when most boards last set their AI strategy: geopolitics has moved from the margins to the centre. At the G7 summit in Evian in June 2026, the chief executives of the world's leading AI companies sat alongside heads of state, a vivid signal of how far this technology has risen up the global agenda. In the space of a single week, the United

States, the European Union and Canada each set out competing national visions for AI; Brussels unveiled a Tech Sovereignty package, and Ottawa launched a national AI strategy built around sovereign compute. When access to the most advanced models can be reshaped overnight by a single government decision, sovereign AI ceases to be an abstract policy debate and becomes a direct operational risk on the enterprise balance sheet. A National AI Plan is fast becoming as fundamental to a country's future as its energy policy or its currency, and the enterprises that understand this shift first will navigate the next decade with a decisive advantage.

This report bridges that gap with evidence. Every statistic is sourced from a verifiable public source or clearly identified as a Peak State Consulting analytical model, and the ten predictions set out our considered view of where enterprise AI is heading. The organisations that will lead the next decade are making hard decisions right now on data infrastructure, governance, talent and vendor strategy.

I would be candid with every leader reading this about the human dimension, because it is the hardest part. As agentic AI moves from pilot to production, some degree of job displacement is no longer a possibility to be debated but a reality to be managed. Entire categories of routine knowledge work will be automated, and pretending otherwise serves no one. Yet this is not a counsel of despair. The enterprises that will be admired a decade from now are those that treat reskilling not as a cost to be minimised but as the central act of leadership in this era, redeploying their people into higher-value, more uniquely human work as the machines absorb the routine. How an organisation handles this transition will define its culture, its reputation and its ability to attract talent for years to come. Handled well, AI augments people; handled poorly, it simply replaces them. The choice is a matter of leadership, not technology.

Above all, this report has a single purpose: to help you implement AI in your enterprise and to do it well. The capital, the models and the vendors are no longer the constraint; the ability to translate them into real, governed, measurable value inside your own organisation is. Implementation is where strategy meets reality, where the gap between the leaders and the rest is won or lost, and where the productivity gains, the cost advantages and the competitive edge are either captured or surrendered. The opportunity is extraordinary, and it belongs to those who move with both urgency and discipline. I hope this report helps you do exactly that.

Please do get in touch. We will be happy to assist.

**Anuj R Khanna**

**CEO, Peak State Consulting**

*London · June 2026*

# Table of Contents

ENTERPRISE AI RESEARCH REPORT · PEAK STATE CONSULTING · JUNE 2026

---

Author's Foreword .....	2
Executive Summary.....	7
Key Predictions for 2026–2030 (10 Predictions).....	15
Market Overview: The AI Imperative.....	20
Enterprise AI Investment Tracker 2026.....	24
AI Adoption Maturity Across Enterprises.....	34
Opportunities: Where AI Creates Enterprise Value .....	37
AI and the Future of Work.....	39
Challenges for CEOs: Strategy, Governance & Change.....	45
Challenges for CTOs & CIOs: Architecture & Security .....	48
Sovereign AI & Geopolitical Strategy .....	52
The Vendor Landscape: Selecting AI Partners .....	57
Leading AI Platforms: Comparative Analysis .....	60
The Future of AI Agents in Enterprise Operations.....	65
The AI Programme ROI Journey .....	69
Peak State AI Readiness Framework .....	72
AI Operating Model Framework.....	74
The Economics of Enterprise AI .....	78
Enterprise AI Implementation Guide .....	82
AI Governance Templates & Implementation Checklists.....	87
White Paper: Implementing Claude in the Enterprise.....	90
About the Author.....	107
About Peak State Consulting .....	109
References & Sources.....	114

---

# EXECUTIVE SUMMARY

## Key Findings & Strategic Overview

The state of enterprise AI in 2026 — opportunity, execution gap, and what separates leaders from laggards

**\$2.59T**  
Global AI  
Spend 2026

**40%**  
Enterprise apps  
embedding agents

**71%**  
CEOs rate AI  
top priority

**43%**  
Agentic AI  
in active use

## Executive Summary

The enterprise AI market has crossed a threshold. Global AI spending is forecast to reach \$2.59 trillion in 2026, up 47% year-on-year (Gartner, May 2026), and Peak State Consulting forecasts the total AI market will reach \$6.40 trillion by 2030, a 25.4% compound annual growth rate from the 2026 base. The enterprise-facing segment, covering AI infrastructure and software, is growing faster still, from \$0.79 trillion in 2026 to \$2.60 trillion by 2030 at a 34.5% CAGR, the enterprise-facing subset of Peak State's \$6.40 trillion total market forecast.

Yet the headline numbers conceal the real story. The gap between the organisations extracting genuine value from AI and those still trapped in perpetual pilots is widening, and widening far faster than most boards appreciate. This report maps where the market stands, who is leading and why, the specific challenges confronting CEOs, CTOs and CIOs, a detailed platform comparison, a practical implementation roadmap, and the commercial and capital-market dynamics that will define the enterprise AI landscape through 2030. Our conclusion is direct and uncomfortable: the technology is ready, the vendors are scaling, the capital is committed, and the execution gap is now the only variable that separates the enterprises that will win from those that will not.

The scale of capital being committed to AI is without historical precedent. Q1 2026 alone saw \$330.9 billion in global venture investment, more than double the prior quarter (KPMG Venture Pulse Q1 2026). The largest AI and AI-adjacent companies, OpenAI, Anthropic, SpaceX (following its February 2026 merger with xAI and its June 2026 listing), Databricks, Cursor and Cohere, together represent over \$3 trillion in estimated market value. These are not speculative projections drawn up by optimists. They are the prices the world's most sophisticated institutional investors are paying today, and they imply a level of enterprise AI adoption over the next five years that demands an urgent strategic response from every major organisation.

### AI IPO Pipeline 2026: The Public Markets Signal

The most significant signal of AI's maturation as an asset class is the IPO pipeline forming in 2026. Cerebras Systems (CBRS) opened the AI IPO era when it priced at \$185 on 13 May 2026 and began trading on 14 May 2026, surging 68% on its first trading day to a market capitalisation of approximately \$95 billion before settling to around \$70 billion in subsequent weeks, illustrating investors' surging appetite for AI data centres and the chips running them (Bloomberg, May 2026). OpenAI confidentially filed a draft S-1 with the SEC on 8 June 2026, working with Goldman Sachs and Morgan Stanley and targeting a public debut later in 2026 (Bloomberg; CNBC, June 2026). SpaceX, which merged with xAI in February 2026, completed the largest IPO in history on 12 June 2026, pricing at \$135 per share to raise \$75 billion at a \$1.77 trillion offering valuation; it began trading on the Nasdaq the same day and closed its first session above a \$2 trillion valuation. Anthropic, valued at \$965 billion following its Series H, confidentially filed a draft S-1 in early June 2026 and is targeting a listing as early as October 2026. The 12 most-watched companies in this pipeline represent a combined estimated value of over \$3 trillion, with AI and AI-adjacent companies

accounting for roughly 92% of that figure, making 2026 the most AI-concentrated IPO year on record.



Source: Bloomberg (May–June 2026), Reuters S-1 review, SEC EDGAR, CNBC, aifundingtracker.com | Peak State Consulting Analysis | June 2026

## Implementing AI in the Enterprise: The Five Critical Success Factors

The Enterprise AI Implementation Guide later in this report provides detailed delivery guidance drawing on Peak State Consulting’s delivery framework and client intelligence. The single most important finding from that analysis is that enterprise AI programmes fail not because the technology does not work, but because the organisation was not ready for it. Data preparation alone typically consumes 40–60% of total implementation effort, a figure that consistently surprises leadership teams who expect AI deployment to begin with model selection rather than data architecture. The following five factors distinguish programmes that achieve production-scale value within twelve months from those that stall at pilot stage.

- 1. Executive alignment and board-level ownership.** Successful programmes begin with a cross-functional AI steering committee reporting directly to the board, a named executive owner with P&L accountability for AI outcomes, and a clear AI vision tied to measurable business objectives, not technology exploration. Organisations must move from “AI strategy” documents to “AI execution” infrastructure with specific milestones, ownership, and financial targets before a single use case is commissioned. See the Enterprise AI Implementation Guide (Phase 1) for the full executive alignment framework.
- 2. Data foundation as a prerequisite, not a parallel workstream.** 52% of enterprise AI practitioners cite data quality and availability as their primary barrier (PEX Network 2025/26). Organisations that attempt AI deployment on unprepared data infrastructure consistently fail to reach production. A dedicated Data Foundation Programme, covering

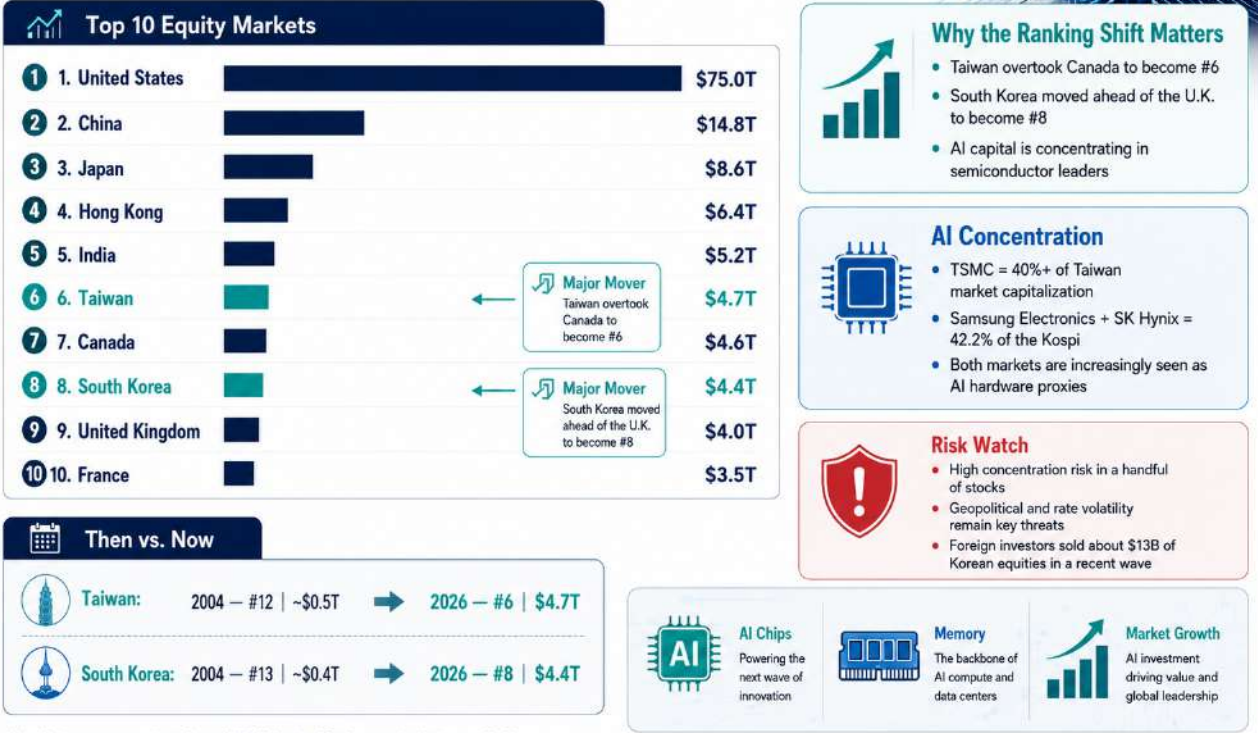
data quality, lineage, access controls, and governance, must be treated as Phase 1 infrastructure, not a concurrent activity. Total cost of ownership for enterprise AI is dominated by integration and data preparation, not inference costs. See the Peak State AI Readiness Framework section for the six-pillar data readiness assessment.

- 3. Governance framework before scale, not after.** The EU AI Act's obligations for most high-risk and transparency requirements apply from 2 August 2026, with full application by 2 August 2027. Obligation dates are staggered by system category, however, rather than landing on a single uniform schedule: prohibited-practice bans and AI-literacy duties applied from February 2025, general-purpose AI model obligations from August 2025, and certain product-safety high-risk and legacy-system obligations extend to 2027 and beyond. Enterprises deploying high-risk AI systems, including those in financial services, healthcare, HR and critical infrastructure, require conformity assessments, human oversight mechanisms, transparency documentation, and incident response protocols in place before production deployment. Governance retrofitted after deployment is two to three times more expensive than governance built in from the start. Peak State's AI Governance Templates and Implementation Checklists (later in this report) provide a ready-to-use framework aligned with the EU AI Act, NIST AI RMF, and FCA guidance.
- 4. Use case prioritisation by value and feasibility, not novelty.** The highest-ROI enterprise AI use cases in 2026 are not the most technically complex. Software engineering productivity (45% productivity gain), legal and contract intelligence (60–80% review time reduction), customer service automation (55% productivity gain), and financial close acceleration are consistently generating production-scale returns because they combine high volume, structured data, and measurable output. Organisations should prioritise two or three use cases using a rigorous value-vs-feasibility matrix before any platform selection decision. Platform selection should follow use case prioritisation, not precede it. See the Opportunities and AI Adoption Maturity sections for the full opportunity assessment and the Maturity Archetype framework.
- 5. Change management and AI literacy as core programme infrastructure.** Only 20% of executives believe their workforce is AI-ready (KPMG 2025). McKinsey data shows enterprises with structured AI upskilling programmes outperform peers by 2.4x on AI-driven productivity gains. AI literacy must be treated as a baseline competency requirement for all roles, not an optional training supplement. Enterprises that invest in structured change management, covering role redesign, workflow integration, and adoption measurement, achieve production-scale deployment in an average of nine months. Those that do not typically remain in pilot purgatory indefinitely. The detailed implementation roadmap, reference architecture, and phased delivery model are set out in full in the Enterprise AI Implementation Guide and the AI Governance Templates & Implementation Checklists sections of this report.



# AI Boom Reorders the Global Equity Market

Top 10 economies by stock market capitalization | Approximate market caps, May 2026



Artificial intelligence is no longer just reshaping technology companies. It is now reshaping the hierarchy of global capital markets.

A powerful AI-driven rally has pushed Taiwan and South Korea sharply higher in the global stock-market rankings, underscoring how investor capital is concentrating around the economies most exposed to advanced semiconductors, memory chips and AI infrastructure. Taiwan has overtaken Canada to become the world’s sixth-largest equity market, while South Korea has moved ahead of the United Kingdom into eighth place. The shift marks one of the clearest signs yet that the AI boom is changing not only corporate valuations, but the global geography of market power.

At the centre of this reordering is a narrow group of companies that have become critical to the AI supply chain. Taiwan Semiconductor Manufacturing Company now represents more than 40% of Taiwan’s total market capitalisation, reflecting the company’s central role in manufacturing the advanced chips required for AI models, data centres and next-generation computing. In South Korea, Samsung Electronics and SK Hynix together account for 42.2% of the Kospi index, turning the country’s benchmark market into a highly concentrated proxy for global AI hardware demand.

This is not a broad-based emerging-market rally. It is a focused repricing of semiconductor leadership.

The scale of the change is striking. In 2004, Taiwan's equity market was worth roughly \$500 billion and ranked 12th globally. Today, it is valued at approximately \$4.7 trillion. South Korea, worth around \$400 billion and ranked 13th two decades ago, now stands at about \$4.4 trillion. The world's five largest markets remain the United States, China, Japan, Hong Kong and India, but the rise of Taiwan and South Korea shows how quickly AI can alter long-established financial rankings.

The driver is clear: investors are chasing the infrastructure layer of artificial intelligence. As AI models become more advanced and more widely deployed, demand for chips, memory, servers and data-centre capacity has surged. The move toward more agentic AI systems — capable of performing increasingly complex tasks — is expected to accelerate demand for computing power. That has created exceptional pricing power for the companies best positioned in the semiconductor value chain.

Yet the rally also carries risk. Markets that rise on concentrated leadership can also fall on concentrated weakness. Taiwan and South Korea's stock-market gains are heavily dependent on a small number of heavyweight technology companies. That creates exposure to earnings disappointments, supply-chain disruption, geopolitics, interest-rate shifts and company-specific events. South Korean equities recently experienced sharp swings after foreign investors sold roughly \$13 billion of local shares, while Samsung's share price has been sensitive to labour negotiations and strike concerns.

For investors, the message is twofold. First, AI is becoming a structural force in global market allocation, not a short-term technology theme. Second, the winners of the AI infrastructure race are creating new forms of concentration risk. Portfolios that appear geographically diversified may, in practice, be increasingly exposed to the same semiconductor cycle. The broader implication is profound. AI is changing how markets assign value to countries, sectors and companies. Economies that sit at the heart of the AI hardware supply chain are gaining financial weight, while some established Western markets are being overtaken by smaller but more strategically positioned technology ecosystems. In the new market order, the question is no longer simply which companies will benefit from AI. It is which economies will become indispensable to powering it.

## Key Findings

---

- Peak State Consulting forecasts that 50% of enterprise apps will embed AI agents by 2027.
- The 451 Alliance independently reports 43% already have agentic AI in active use in Pilot and POC stages.
- Data quality and availability is the greatest barrier, cited by 52% of professionals surveyed (PEX Network Survey 2025/26, n=200+), consistent with IBM's finding that data complexity affects 25% of enterprises deploying AI.
- Only 27% of executives have a comprehensive AI strategy; only 20% believe their workforce is AI-ready.
- 71% of global CEOs say AI is their top investment priority for 2026 (KPMG Global CEO Outlook 2025, n=1,325 CEOs).
- The first half of 2026 saw record venture investment: OpenAI (\$122B, Q1), Anthropic (\$65B Series H, May), xAI (\$20B) and Databricks (\$7B) raised over \$210 billion combined (KPMG Venture Pulse Q1 2026; Bloomberg May 2026).
- The AI IPO pipeline is the largest in history: Cerebras (CBRS) priced at \$185 on 13 May 2026, began trading 14 May 2026 and surged 68% on Day 1. SpaceX completed the largest IPO in history on 12 June 2026, raising \$75 billion at a \$1.77 trillion offering valuation and trading above \$2 trillion on its first day. OpenAI confidentially filed a draft S-1 with the SEC on 8 June 2026, targeting a 2026 debut. Anthropic confidentially filed in early June 2026 and is targeting a listing as early as October 2026. The 2026 AI IPO cohort represents over \$3 trillion in combined estimated valuation, the most AI-concentrated IPO year on record (Bloomberg; Reuters; SEC EDGAR, June 2026).
- Sovereign AI has moved from policy debate to board-level imperative. At the June 2026 G7 summit in Evian, AI company chief executives joined heads of state as the US, EU and Canada each unveiled competing national AI strategies in a single week. 84% of organisations now factor digital sovereignty into their AI strategy, and global spending on sovereign AI systems is projected to surpass \$100 billion in 2026 (Peak State Consulting analysis; Gartner). A National AI Plan and digital strategy is becoming a baseline requirement of national competitiveness.
- Chinese AI providers have surged from under 2% to over 45% of OpenRouter token traffic in twelve months, a signal of model-routing demand and price/performance pressure; they offer frontier-class performance at prices five to ten times lower than equivalent US models (Q2 2026).

AI New Market Order, Source and date note: Based on CNBC reporting, "AI boom reshuffles global stock market pecking order as South Korea and Taiwan surge," published May 19, 2026, updated May 25, 2026. Supplementary reference: CNBC video, "Why the AI boom is reshuffling the global stock market hierarchy,". Prepared for Enterprise AI by Peak State Consulting, June 2026.

#### KEY INSIGHT

The gap between AI intention and AI execution is the defining challenge of 2026. We forecast 50% of enterprise applications will embed AI agents by 2027; the 451 Alliance independently reports 43% of organisations already have agentic AI in active pilot or POC use. These are different measures pointing to the same conclusion: deployment is accelerating, but production-scale value realisation is not keeping pace. The enterprises that close this execution gap, through disciplined programme management, data foundation investment, and structured governance, will compound structural competitive advantages over the next three years. Those that do not will face permanent strategic disadvantage in every market they operate in. The technology risk in enterprise AI is no longer whether it works. The risk is whether your organisation is capable of executing at the speed the market demands.

#### METHODOLOGY

This report combines third-party market data with Peak State Consulting's own analysis. Source hierarchy: where available, we prioritise primary sources (company announcements, SEC filings, regulator publications) and established research houses (Gartner, IDC, McKinsey, KPMG, PwC, S&P Global 451 Research, Crunchbase), with each material figure attributed at point of use. Figures described as Peak State forecasts or estimates are our own analytical models, derived from primary and secondary research, client engagements and pattern analysis, and are clearly labelled as such. Market definitions: "total AI market" covers the full AI economy (infrastructure, software, services, devices and AI-optimised semiconductors), while "enterprise-facing AI" covers AI infrastructure and software sold to enterprises; agentic "pilot/POC" usage is distinguished throughout from "production deployment." Currency basis: figures are in US dollars unless otherwise stated. Forecast approach: forward-looking figures are scenario-based and inherently uncertain (see the Forward-Looking Statements disclaimer). Data cut-off: market and capital-market data are current to the publication date of June 2026 and will move thereafter.

# KEY PREDICTIONS

## Ten Strategic Forecasts for 2026–2030

Agentic AI displacement · Governance failure · Chinese AI · Talent scarcity · Board-level ROI disclosure

**10**

Strategic  
predictions

**2028**

Agentic AI  
displacement

**2027**

Governance  
failure tipping point

**2029**

Valuation premium  
for AI governance

## Key Predictions for 2026–2030: Ten Strategic Forecasts

Peak State Consulting issues the following ten strategic predictions based on primary and secondary research, client intelligence and pattern analysis.

NO.  
**01**

### **Agentic AI will displace 50% of knowledge worker task workflows by 2028.**

IDC and McKinsey converge on roughly \$1.4 trillion in annual productivity value unlocked by AI agents by 2028. This is incremental economic value, not spending. Peak State forecasts total AI market spending will reach \$6.40 trillion by 2030, providing the infrastructure base from which this productivity gain is realised. Peak State forecasts widespread adoption of Agentic AI causing a major disruption on dependency on humans for knowledge worker task flows. This may cause considerable job losses but not necessarily replace all knowledge workers. New training and upskilling will be required.

NO.  
**02**

### **By 2027, 60% of enterprises will experience at least one significant AI-related governance failure.**

This is a Peak State Consulting forecast. We define a “significant AI-related governance failure” as an incident causing material financial loss, regulatory enforcement action, or reputational damage serious enough to warrant board-level attention. Inadequate AI governance, model hallucinations in critical workflows, EU AI Act breaches, or data privacy violations, will result in reputational and financial damage. The EU AI Act entered into force on 1 August 2024, with most high-risk obligations applying from 2 August 2026 and full application by 2 August 2027. Governance failures are highly likely in spite of all the best efforts put in by corporations and governments.

NO.  
**03**

### **Existing enterprise hyperscale platforms will control over 50% of enterprise AI infrastructure spend by 2030.**

The enterprise AI platform software segment will grow rapidly by 2030. The existing enterprise software hyperscalers will retain more than 50% market share and will also be challenged by competitors such as Anthropic that are expected to gain major market share in the enterprise sector.

NO.  
**04**

### **AI talent scarcity peaks in 2027–2028, forcing a shift to AI-augmented development.**

The shortage of qualified AI engineers and data scientists will become the most acute programme constraint. Cursor reached \$2 billion in annualised revenue by February 2026, the fastest zero-to-\$2B arc in B2B software history, demonstrating extraordinary demand for AI-augmented development tools.

NO.  
**05**

**Enterprises with mature AI governance may command 15–25% valuation premiums by 2030.**

Capital markets are already pricing AI readiness into valuations at some level. AI startups trade at a 3.2x valuation premium over traditional tech (PwC MoneyTree 2026). Boards treating AI governance as compliance-only will be penalised by investors.

NO.  
**06**

**The EU AI Act will force re-architecting of AI systems for 70% of European enterprises by end of 2027.**

High-risk AI classifications require conformity assessments, human oversight mechanisms and transparency documentation. Compliance-ready vendors are gaining measurable advantage as the EU AI Act provisional application is already in effect. Different Sovereign AI regulations will require constant updates and compliance on a global basis.

NO.  
**07**

**Chinese AI providers will capture significant global enterprise API token spend by end of 2027.**

Chinese providers already account for over 45% of OpenRouter token traffic (Q2 2026). DeepSeek, Alibaba Qwen and Baidu ERNIE offer pricing five to ten times lower per token than equivalent US frontier models, fundamentally reshaping enterprise AI economics globally. US AI companies will dominate large markets such as India unless major local models emerge. US and China are expected to be the major market players in AI followed by Europe. However, AI adoption may be higher in India and other Asian Markets.

NO.  
**08**

**Skills in AI Product Development, Programme Management, Security, Legal Compliance and Governance will be the most sought-after new hire attributes, leading to a severe shortage of trained talent and pushing up wages for skilled experts.**

Enterprises are discovering that deploying AI across an organisation requires a new generation of hybrid professionals: individuals who combine deep domain knowledge with AI fluency and the ability to govern, secure and manage AI systems in production. Demand for AI product managers, AI governance specialists, AI security analysts and AI legal compliance officers is growing at a rate that university and professional training pipelines cannot match. LinkedIn, Glassdoor and sector salary surveys (2025–2026) consistently show AI-specialist roles commanding salary premiums of 35–65% over equivalent non-AI roles.

NO.  
09

**AI will become the primary interface between enterprises and their data by 2028, replacing traditional BI dashboards and static reporting for the majority of knowledge worker interactions.**

Natural language querying of enterprise data, asking Claude or equivalent AI to "show me which regions are underperforming against Q2 targets and why" and receiving an instant, context-aware analytical response, is already in production at leading enterprises. Snowflake's Cortex, Databricks' Genie, and Anthropic's Claude integrations with enterprise data platforms are collapsing the time between data and decision from hours to seconds. Peak State forecasts that by 2028, the majority of enterprise employees will use AI-powered data interfaces as their primary analytics tool, up from less than 10% in 2024. This shift will fundamentally alter the role of the data analyst and the BI function, redirecting human analytical effort from report production to interpretation, judgement and action.

NO.  
10

**By 2030, AI programme ROI may become a mandatory board-level reporting metric, with institutional investors requiring standardised AI performance disclosures equivalent to ESG reporting.**

Peak State forecasts total AI market spending will reach \$6.40 trillion by 2030, and this scale of capital is intensifying investor scrutiny of AI return on investment. Capital markets are beginning to differentiate between AI investment that creates sustainable competitive advantage and AI spend that generates activity without measurable business value. PwC's MoneyTree 2026 report confirms that AI startups trade at a 3.2x premium to comparable non-AI technology companies, and that premium is increasingly conditional on demonstrated production deployment and measurable ROI rather than capability demonstration alone. Peak State Consulting anticipates that major institutional investors and proxy advisers will begin mandating standardised AI programme disclosures, covering AI investment levels, production deployment scale, governance maturity and quantified business impact, within the next 24 months. Enterprises that build rigorous AI measurement frameworks now will be ahead of this disclosure curve; those that do not will face uncomfortable investor questions at a time when they are least prepared to answer them.

Reference sources include Peak State Consulting Research 2026, McKinsey Global AI Survey 2025, IDC Worldwide AI Spending Guide Q1 2026, Forrester Agentic AI Wave Q1 2026, S&P Global 451 Research, KPMG Global CEO Outlook 2025, Gartner AI Report and IBM Global AI Adoption Index 2024.

### Forward-Looking Statements and Predictions Disclaimer

The ten strategic predictions set out in this section (“Predictions”) represent the forward-looking analytical views of Peak State Consulting analysts as at June 2026, based on primary research, secondary research, client intelligence, publicly available data and pattern analysis. They are provided for general informational purposes only and do not constitute, and must not be construed as, professional advice of any kind, including without limitation financial advice, investment advice, legal advice, regulatory advice, or technology procurement advice.

Forward-looking statements involve known and unknown risks, uncertainties and other factors that may cause actual outcomes, market developments, technology trajectories, or business results to differ materially from those expressed or implied in the Predictions. Factors that could cause such differences include, without limitation: changes in macroeconomic conditions, geopolitical developments, regulatory interventions, technology breakthroughs or failures, competitive dynamics, changes in enterprise investment priorities, and unforeseen events of any nature. The AI market is characterised by an exceptionally high rate of change, and the reliability of any forward-looking prediction in this domain diminishes significantly as the time horizon extends. Readers are accordingly cautioned not to place undue reliance on the Predictions contained in this section.

The Predictions reflect Peak State Consulting’s internal analytical assessment of evidential strength at the time of publication and are not representations of probability, guarantee, or assurance of any outcome. Predictions may be speculative and reflect emerging trends that Peak State Consulting believes are directionally plausible but which are supported by limited precedent or empirical evidence at the date of this report. No warranty, express or implied, is given as to the accuracy, completeness or reasonableness of any Prediction.

Peak State Consulting Limited accepts no liability whatsoever for any loss, damage, cost, expense or liability of any kind arising directly or indirectly from any reliance placed on the Predictions by any person, including without limitation any business decision, investment decision, procurement decision, technology deployment decision, or strategic planning decision made in reliance on the Predictions. This limitation of liability applies to the fullest extent permitted by applicable law. Readers and organisations wishing to make decisions informed by the subject matter of the Predictions are strongly advised to obtain independent, professionally qualified advice tailored to their specific circumstances before doing so.

*This disclaimer is supplemental to, and should be read in conjunction with, the full General Disclaimer, Independence and Governance statement, and Forward-Looking Statements notice set out in the Disclaimer section of this report.*

# MARKET OVERVIEW

## The AI Imperative: Scale, Sector & Budget Allocation

Global spending, sector adoption, and how enterprise AI budgets are being allocated across functions

**\$2.59T**

Global AI  
Spending 2026

**\$6.40T**

Peak State Total  
Market 2030

**25.4%**

Total Market  
CAGR 2026–2030

**34.5%**

Enterprise Subset  
CAGR 2026–2030

# Market Overview: The AI Imperative

## The Scale of the Opportunity

Global AI spending is forecast to reach \$2.59 trillion in 2026 (Gartner, May 2026), up 47% from \$1.76 trillion in 2025 and up from \$1.10 trillion in 2024. This figure encompasses the full AI economy: infrastructure hardware, enterprise software, services, devices and AI-optimised semiconductors. Within this, the enterprise-facing components tell their own story.

## Peak State Consulting: Five-Year AI Spending Forecast

Peak State Consulting issues two complementary forecasts. Peak State projects \$3.35 trillion in 2027, \$4.28 trillion in 2028, \$5.30 trillion in 2029 and \$6.40 trillion by 2030, a 25.4% compound annual growth rate from the estimated 2026 base of \$2.59 trillion. The second covers the enterprise-facing subset, combining AI infrastructure with enterprise software and services: this narrower but faster-growing segment is projected to grow from \$0.79 trillion in 2026 to \$2.60 trillion by 2030 at a 34.5% CAGR, the enterprise-facing subset of Peak State's \$6.40 trillion total market forecast, reflecting stronger enterprise software penetration growth than the broader market.

References: Peak State, IDC AI Infrastructure Tracker (April 2026) and Goldman Sachs AI CapEx model (May 2026), Gartner AI Report (2026).

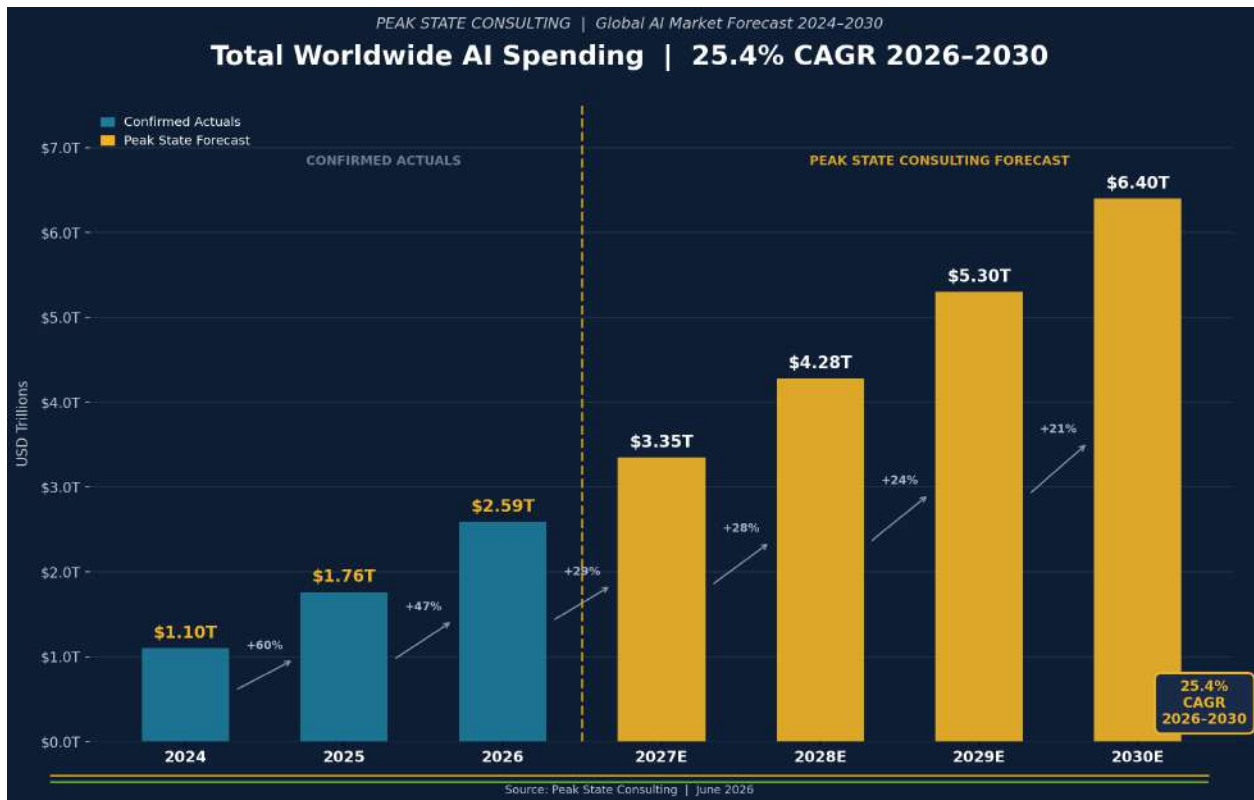


Figure 2a: Peak State Consulting, Total AI Market Forecast 2024–2030 | Actuals (2024–2025), 2026 Estimate + Peak State Forecast (2027–2030) | Source: Peak State Consulting Research, June 2026

## Peak State Consulting: Growth Trajectory 2026–2030

The total AI market crosses \$3 trillion in 2027 and \$6 trillion by 2030. The enterprise-facing subset, infrastructure plus software, crosses \$1 trillion in 2027 and \$2 trillion in 2029, reaching \$2.60 trillion by 2030. These two numbers, \$6.40 trillion total and \$2.60 trillion enterprise-facing, are measuring different parts of the same market. Both are growing. Both are credible.

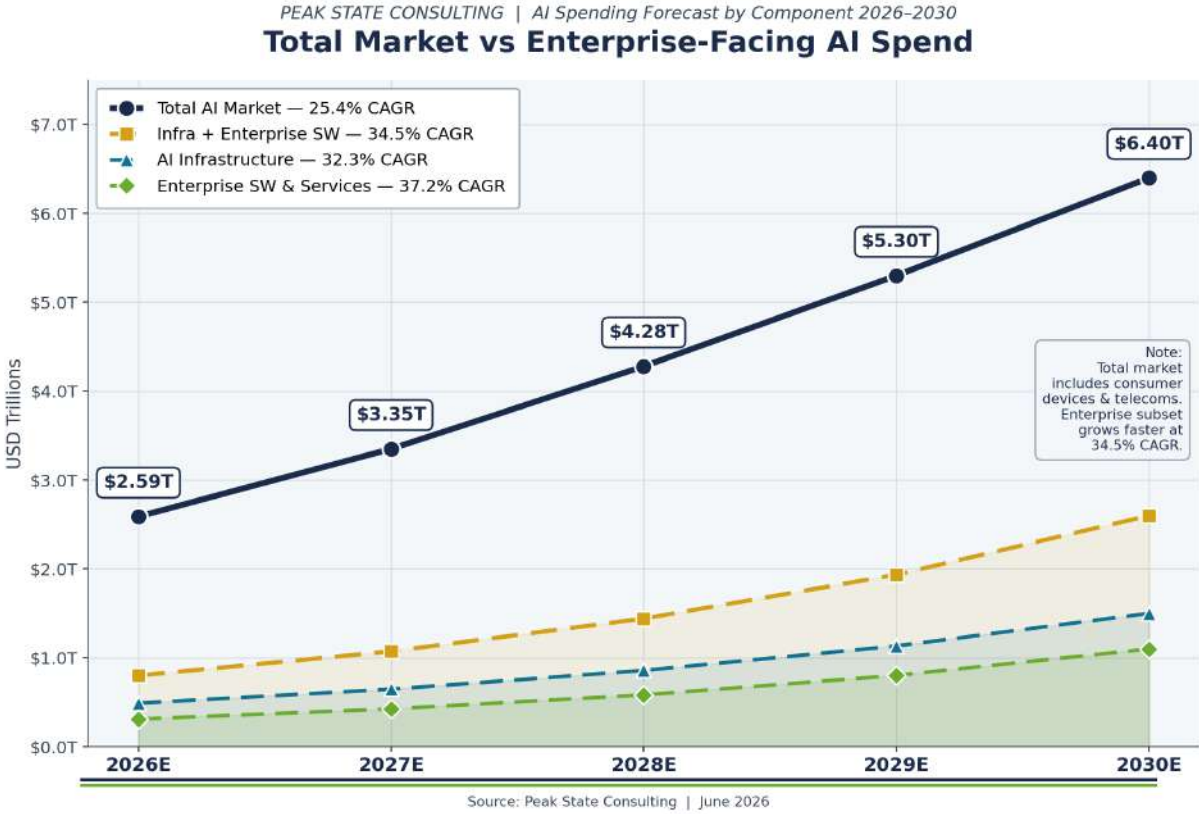


Figure 2b: Peak State Consulting, AI Spending Forecast by Component 2026–2030 | Total market vs enterprise-facing subset | Source: Peak State Consulting Research, June 2026

**Table 2b: Reconciliation of the \$6.40 Trillion Total AI Market Forecast (2030)**

Market segment	2026 base	2030 forecast	CAGR	% of 2030 total
<b>AI infrastructure</b> (compute, hardware, cloud)	\$0.49T	\$1.50T	32.3%	23%
<b>Enterprise software &amp; services</b>	\$0.31T	\$1.10T	37.2%	17%
<b>Enterprise-facing subset</b> ( <i>infra + software</i> )	<b>\$0.79T</b>	<b>\$2.60T</b>	<b>34.5%</b>	<b>41%</b>
<b>Consumer, devices, telecoms &amp; AI semiconductors</b>	\$1.80T	\$3.80T	20.5%	59%
<b>Total AI market</b>	<b>\$2.59T</b>	<b>\$6.40T</b>	<b>25.4%</b>	<b>100%</b>

*Note: The enterprise-facing subset (\$2.60T) is the sum of AI infrastructure and enterprise software & services, the two faster-growing segments charted above. The remaining \$3.80T comprises consumer AI, AI-enabled devices, telecoms and AI-optimised semiconductors. The total (\$6.40T) and the enterprise-facing subset (\$2.60T) measure different parts of the same market; they are not additive to one another. Segment totals may not sum exactly to the total due to rounding. Source: Peak State Consulting Research; IDC AI Infrastructure Tracker (April 2026); Gartner (May 2026); June 2026.*

Key market figures 2026 and Sources: Total AI market \$2.59T (+47% YoY) | AI infrastructure \$487B (IDC, +53% YoY) | Enterprise AI software and services \$307B (IDC) | Peak State total market forecast 2030: \$6.40T

### Sector Adoption Snapshot

<p><b>71%</b></p> <p>Financial services firms formally using AI (ACA/NSCP 2025)</p>	<p><b>20.77%</b></p> <p>Healthcare AI sector CAGR through 2031 (Mordor Intelligence)</p>	<p><b>\$2.6–4.4T</b></p> <p>Annual value AI could add (McKinsey Global Institute)</p>	<p><b>26–55%</b></p> <p>Productivity gains in AI-transformed operations (McKinsey 2025)</p>
---	--	---	---

### AI Budget Allocation by Business Function

Operations and supply chain leads at 21% of total enterprise AI budgets in 2025. Customer-facing applications accounted for 38.91% of 2025 AI spending overall (Mordor Intelligence). Human resources and talent workflows are forecast to grow at the fastest rate, 19.76% CAGR, through 2031.

# ENTERPRISE AI INVESTMENT TRACKER

Selected Funding Rounds Above \$20M — 2026

\$330.9B global VC investment in Q1 2026 alone — 80% directed at AI-focused companies

**\$330.9B**

Global VC  
Q1 2026

**\$122B**

OpenAI  
Q1 2026

**\$30.6B**

Anthropic  
Q1 2026

**80%**

VC funding  
directed at AI

# Enterprise AI Investment Tracker 2026

Q1 2026 shattered venture funding records, with global VC investment reaching \$330.9 billion, more than double Q4 2025's \$128.6 billion (KPMG Venture Pulse Q1 2026).

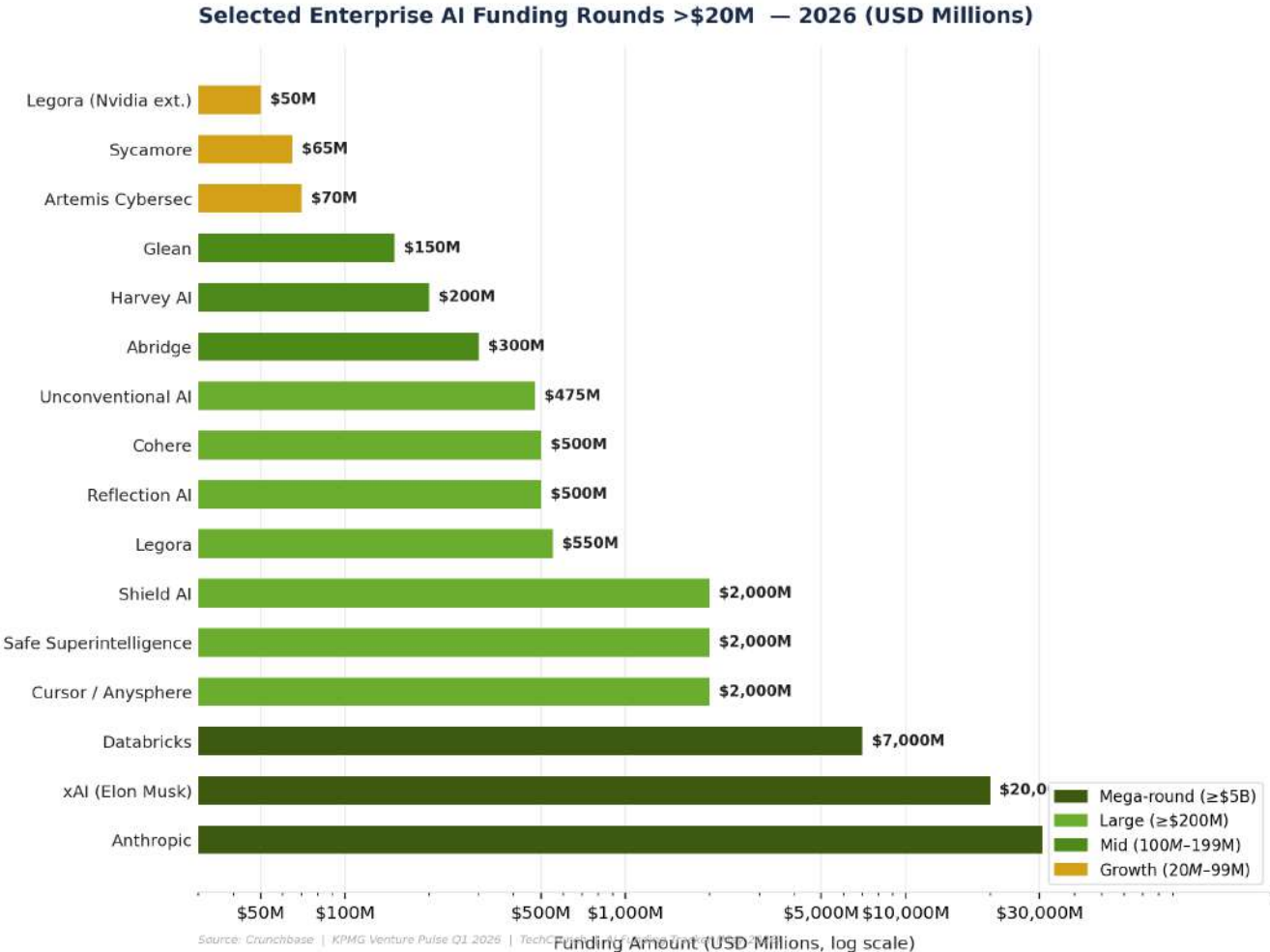


Figure 17: Selected Enterprise AI Funding Rounds >\$20M in 2026 (USD Millions, log scale) | Source: Crunchbase, KPMG Venture Pulse Q1 2026, TechCrunch, AI Funding Tracker (June 2026)

# Enterprise AI Investment Rounds Above \$20M, 2026

COMPANY	ROUND	AMOUNT (USD)	FOCUS & NOTES
OpenAI	Q1 2026	\$122,000M	Frontier AI / AGI; ChatGPT; enterprise API. Amazon/Nvidia/SoftBank; \$852B valuation.
Anthropic	Q2 2026	\$65,000M	Foundation models; Claude enterprise API. Series H led by Altimeter, Dragoneer, Greenoaks, Sequoia; \$965B valuation (May 28, 2026). Includes \$5B Amazon investment (April 2026). ARR: \$47B run rate.
xAI (Elon Musk)	Q1 2026	\$20,000M	Frontier AI / AGI; Grok models. Nvidia, Cisco, Fidelity invested; SpaceX merger Feb 2026.
Waymo	Q1 2026	\$16,000M	Autonomous vehicles; AI-powered robotics; Alphabet subsidiary.
Databricks	Q1 2026	\$7,000M	Data + AI platform; Mosaic AI; Delta Lake. Enterprise data infrastructure.
Cursor / Anysphere	Q2 2026 (talks)	\$2,000M	AI coding agent; \$2B ARR Feb 2026; fastest B2B growth arc in history.
Safe Superintelligence	Q2 2025	\$2,000M	Closed frontier research lab; \$32B valuation. Greenoaks led.
Shield AI	Q1 2026	\$2,000M	Autonomous AI for defence; Series G; \$12.7B valuation.
Legora	Q1 2026	\$550M	Legal AI (Europe); \$5.55B valuation; Series D.
Cohere	Q3 2025	\$500M	Enterprise-focused foundation models; sovereign AI; \$6.8B valuation.
Reflection AI	2026 cumul.	\$500M	Open-weight frontier model; \$2.1B raised across 3 rounds.
Unconventional AI	Q2 2026	\$475M	Neuromorphic AI compute; biology-inspired; a16z and Lightspeed co-led.
Abridge	Q2 2025	\$300M	AI clinical scribe; \$5.3B valuation; a16z and Khosla led.
Harvey AI	Q1 2026	\$200M	AI for legal; \$11B valuation; deployed in ~50% of Am Law 100 firms.
Glean	2025/2026	\$150M	Enterprise AI search & agents; \$7.2B valuation; Series F.
Artemis	Q2 2026	\$70M	Agentic cybersecurity platform; autonomous threat detection.
Sycamore	Q2 2026	\$65M	Agentic AI OS for enterprise; seed. Coatue and Lightspeed led.
Legora (Nvidia ext.)	Q2 2026	\$50M	Extension from Nvidia NVentures; legal AI vertical.

AI capital concentration has accelerated further into Q2 2026. OpenAI (\$122B, Q1), Anthropic (\$65B Series H announced May 28, 2026, at \$965B valuation), xAI (\$20B) and Waymo (\$16B) collectively raised \$223 billion.

## Investment Trend Analysis

---

Mega-rounds dominate: rounds above \$100 million now drive most of the dollar volume (PwC MoneyTree). The most dramatic illustration is Anthropic, which raised \$30B (Series G) in February, then \$65B (Series H, incl. \$15B prior hyperscaler commitments) on May 28, 2026, reaching a \$965B valuation in under four months. Vertical AI is winning: Harvey (\$11B), Abridge (\$5.3B), Glean (\$7.2B) and Legora (\$5.55B) demonstrate that deep domain expertise commands premium valuations. AI coding tools: Cursor's trajectory from \$100M ARR (January 2025) to \$2B ARR (February 2026) is the fastest B2B software growth in history.

## Top 10 AI IPOs to watch out for in 2026

# Top 10 AI IPOs & Public-Market Names to Watch

Verified snapshot of leading AI companies, recent funding signals, and IPO activity  
 Data checked against company announcements, SEC filings, and investor-relations releases

**1** **OpenAI** Confidential S-1

**\$852B** post-money  
 Funding round closed — Mar 2026  
 ChatGPT leader with global consumer and enterprise reach

**2** **Anthropic** Confidential S-1

**\$965B** post-money  
 Series H + draft S-1 — May/Jun 2026  
 Claude platform is a major enterprise AI contender

**3** **Databricks** IPO Watch

**>\$100B** valuation  
 Series K announced — Aug 2025  
 Data + AI platform with significant enterprise revenue scale

**4** **Scale AI** IPO Watch

**>\$29B** valuation  
 Meta investment announced — Jun 2025  
 Critical data, evaluation, and AI systems infrastructure

**5** **xAI** IPO Watch

**\$20B** Series E raised  
 Upsized funding round — Jan 2026  
 Frontier AI player with large compute ambition

**6** **Anduril** IPO Watch

**\$61B** valuation  
 Series H raise — May 2026  
 AI, autonomy, and defense technology are converging fast

**7** **Mistral AI** IPO Watch

**€11.7B** post-money  
 Series C announced — Sep 2025  
 Europe's leading frontier-model contender

**8** **Harvey** IPO Watch

**\$11B** valuation  
 Growth round announced — Mar 2026  
 One of the strongest vertical AI software stories

**9** **CoreWeave** Public since 2025

**\$40** IPO price  
 IPO priced — Mar 2025  
 Important benchmark for AI cloud and infrastructure stocks

**10** **Cerebras** Public since 2026

**\$185** IPO price  
 IPO priced — May 2026  
 Specialized AI chip company now in public markets

### KEY TAKEAWAYS

- The 2026 AI equity story spans models, data, infrastructure, cloud, chips, and vertical software
- Only two names here are already public: CoreWeave and Cerebras; the rest remain candidates or watchlist names
- OpenAI and Anthropic stand out because both disclosed confidential S-1 submissions in June 2026

Sources used: company press releases, SEC filings, and investor-relations announcements. Private-company IPO timing remains unconfirmed and subject to market conditions.

## AI IPO watch list

---

The artificial intelligence IPO pipeline is becoming one of the most important capital-market themes of 2026. After several years in which the most valuable AI companies remained private, a new phase is emerging: public-market validation. Investors are no longer evaluating AI companies purely on narrative, model capability, or user growth. They are increasingly focused on revenue quality, compute efficiency, customer concentration, infrastructure control, governance, and the ability to scale without unsustainable capital intensity.

The AI equity story is also broadening. It is no longer limited to frontier model developers. The most important names now span foundation models, AI cloud infrastructure, enterprise data platforms, model-evaluation systems, legal AI, defence autonomy, and specialist AI chips. This means the next AI IPO cycle is likely to be multi-layered, with different companies representing different parts of the AI value chain. The following ten companies stand out as the most important AI IPO and public-market names to watch.

### 1. OpenAI

OpenAI remains the most visible private AI company in the world, largely because ChatGPT has become both a consumer product and an enterprise platform. The company's latest disclosed funding round valued it at \$852 billion post-money, making it one of the largest private technology companies globally.

OpenAI has also confidentially submitted a draft S-1 registration statement to the U.S. Securities and Exchange Commission. This does not guarantee an IPO, but it gives the company optionality to move toward a public listing when market conditions, governance structure, and strategic priorities align.

For public investors, OpenAI's eventual IPO would be a major test of how the market values frontier AI. Key questions will include revenue growth, enterprise penetration, compute costs, gross margins, partnership dependence, and long-term governance.

### 2. Anthropic

Anthropic has emerged as one of OpenAI's most important competitors, particularly in enterprise AI and developer workflows. Its Claude platform is positioned around reliability, safety, and business adoption, which could make it attractive to public-market investors looking for enterprise-grade AI exposure.

Anthropic announced a Series H funding round at a \$965 billion post-money valuation and also disclosed that it had confidentially submitted a draft S-1 registration statement to the SEC. That combination makes Anthropic one of the clearest AI IPO candidates.

Its public-market case will likely depend on whether investors believe Claude can become a durable enterprise AI platform rather than simply one of several competing frontier models. Customer retention, usage intensity, cloud partnerships, and model-performance differentiation will be central to the investment debate.

### **3. Databricks**

Databricks is one of the strongest enterprise AI infrastructure candidates because it sits at the intersection of data, analytics, governance, and AI applications. Unlike some frontier model companies, Databricks has a long-established enterprise customer base and a clearer software revenue profile.

The company announced a Series K investment at a valuation of more than \$100 billion. It has also positioned the new capital around its AI strategy, including Agent Bricks, Lakebase, and broader enterprise AI workloads.

Databricks matters because most enterprise AI deployment begins with data. Companies cannot build reliable agents, copilots, analytics systems, or automation layers without trusted data infrastructure. That makes Databricks a potential bellwether for how public markets value the “data layer” of AI.

### **4. Scale AI**

Scale AI represents a different part of the AI stack: data, evaluation, and model support infrastructure. Its role has evolved beyond data labelling into systems that help AI labs, enterprises, and governments evaluate and deploy AI more reliably.

Scale announced a major investment from Meta that valued the company at over \$29 billion. The transaction reinforced Scale’s relevance to the AI ecosystem and highlighted the strategic value of high-quality data and evaluation systems.

If Scale eventually pursues an IPO, investors will focus on whether the company can convert its strategic importance into durable, diversified revenue. Customer concentration, dependence on major AI labs, and the evolution from services-heavy work toward software and platform revenue will be important issues.

### **5. xAI**

xAI is one of the most closely watched private AI companies because it combines frontier-model ambition, Elon Musk’s distribution network, and significant compute investment. The company announced that it raised \$20 billion in an upsized Series E funding round.

The investment case around xAI would likely centre on three issues: model capability, distribution through the broader Musk ecosystem, and infrastructure scale. Unlike some AI start-ups, xAI has potential access to large consumer, enterprise, and social-distribution channels.

However, xAI also remains a watchlist name rather than a confirmed IPO candidate. Its public-market attractiveness will depend on whether it can demonstrate differentiated usage, revenue generation, and a credible path to competing with OpenAI, Anthropic, Google, and other AI platforms.

## **6. Anduril**

Anduril is not a pure software AI company, but it is one of the most important AI-adjacent companies to watch because defence, autonomy, robotics, and AI are converging rapidly. The company announced a \$5 billion Series H raise at a \$61 billion valuation.

Anduril's positioning is distinct from the consumer and enterprise AI companies on this list. It operates in defence technology, autonomous systems, and advanced hardware-software platforms. As governments increase spending on AI-enabled defence capabilities, Anduril could become a major public-market story.

An Anduril IPO would test investor appetite for defence AI and autonomy. The opportunity is large, but the risk profile is different from software AI: government procurement cycles, geopolitical exposure, manufacturing execution, and regulatory scrutiny will all matter.

## **7. Mistral AI**

Mistral AI is Europe's leading frontier-model company and one of the most important non-U.S. AI companies to watch. The company announced a €1.7 billion Series C funding round at an €11.7 billion post-money valuation.

Mistral's importance goes beyond valuation. It has become a central company in the European AI sovereignty debate, where governments and enterprises want credible alternatives to U.S.-based AI platforms. Its combination of open-weight models, enterprise tooling, and European strategic relevance makes it a potential long-term public-market candidate.

For investors, the main question will be whether Mistral can build a large commercial platform while competing against much larger U.S. and Chinese AI players. Its IPO potential will depend on revenue traction, enterprise adoption, infrastructure access, and the strength of Europe's AI ecosystem.

## **8. Harvey**

Harvey is one of the clearest examples of vertical AI software. Rather than trying to build a general-purpose AI platform, Harvey focuses on legal workflows for law firms and in-house legal teams.

The company announced a \$200 million funding round at an \$11 billion valuation. This places it among the most highly valued vertical AI software companies and signals investor confidence in domain-specific AI applications.

Harvey’s IPO case would be different from frontier model companies. Public investors would likely value it more like a high-growth vertical SaaS business, with questions around annual recurring revenue, customer retention, usage expansion, gross margins, and competitive defensibility within legal AI.

**9. CoreWeave**

CoreWeave is already public, but it remains one of the most important AI public-market benchmarks. The company priced its IPO at \$40 per share in March 2025 and listed on Nasdaq under the ticker CRWV.

CoreWeave matters because it gives investors a real-time public-market reference point for AI cloud infrastructure. As frontier AI models require enormous compute capacity, infrastructure providers such as CoreWeave sit close to the centre of AI demand.

The company’s performance will influence how investors assess other infrastructure-heavy AI businesses. If CoreWeave demonstrates strong growth, margin expansion, and disciplined capital expenditure, it could support higher valuations for future AI infrastructure IPOs. If not, it may make investors more cautious about compute-intensive AI models.

**10. Cerebras**

Cerebras is also now a public-market name. The company priced its IPO at \$185 per share in May 2026 and began trading on Nasdaq under the ticker CBRS. Its listing is important because it gives investors exposure to specialist AI chip architecture outside the dominant GPU ecosystem.

Cerebras is focused on AI hardware and systems designed for large-scale model training and inference. Its IPO is a significant market test for companies attempting to challenge or complement Nvidia’s dominance in AI compute.

The key public-market question is whether specialist AI chip companies can build large, sustainable businesses in a market where customers demand performance, availability, software compatibility, and ecosystem support. Cerebras will be watched closely as a benchmark for the next wave of AI hardware companies.

**Key Themes for Investors**

The top AI IPO candidates reveal three major themes.

**First**, the AI market is becoming stack-based. The most important companies are not all doing the same thing. OpenAI, Anthropic, xAI, and Mistral represent frontier models. Databricks and Scale represent the enterprise data and evaluation layer. CoreWeave represents AI cloud infrastructure. Cerebras represents specialist chips. Harvey represents vertical AI software. Anduril represents defence autonomy and AI-enabled hardware.

**Second**, public-market readiness will require more than growth. Investors will want to see revenue quality, margin structure, customer diversification, governance, and a credible path to profitability. The cost of compute is now one of the central questions in AI investing.

**Third**, the IPO market is likely to separate AI leaders from AI narratives. Companies with strategic infrastructure, strong customer demand, and defensible business models may command premium valuations. Companies that rely only on broad AI enthusiasm may find public markets less forgiving.

**Funding round sources:** OpenAI disclosed its \$122 billion funding round at an \$852 billion post-money valuation and separately confirmed a confidential S-1 submission (OpenAI). Anthropic announced its \$965 billion Series H valuation and later disclosed a confidential draft S-1 submission to the SEC (Anthropic). Databricks announced its Series K at a valuation above \$100 billion, while Scale AI disclosed Meta's investment valuing Scale at over \$29 billion (Databricks). xAI announced a \$20 billion Series E, Anduril announced a \$5 billion Series H at a \$61 billion valuation, and Mistral announced its €1.7 billion Series C at an €11.7 billion post-money valuation (xAI). Harvey announced a \$200 million round at an \$11 billion valuation (Harvey). CoreWeave priced its IPO at \$40 per share in March 2025, and Cerebras priced its IPO at \$185 per share in May 2026 (CoreWeave). Sources: company press releases and SEC filings.

# AI ADOPTION MATURITY

## Five Enterprise Archetypes: Pioneers to Laggards

Where enterprises stand today – and what separates those extracting value from those that are not

**15%**  
AI  
Pioneers

**27%**  
AI Fast  
Followers

**31%**  
AI  
Experimenters

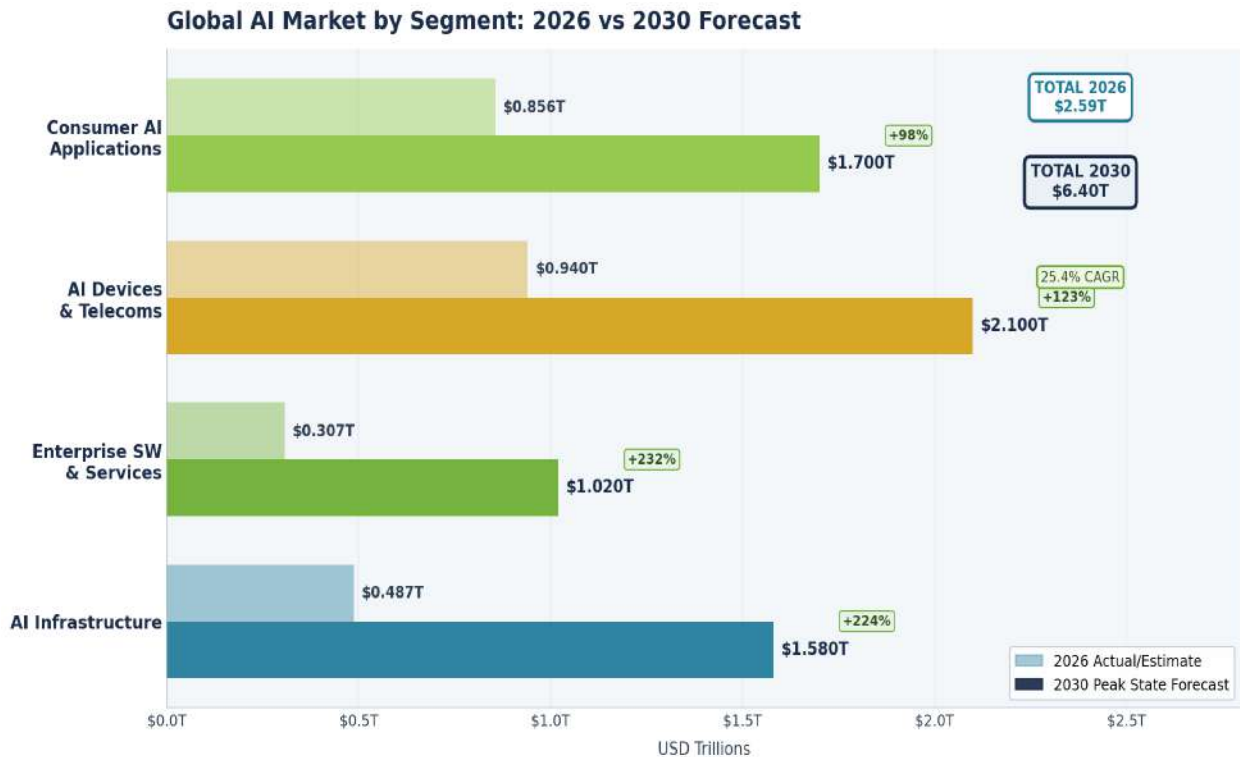
**18%**  
AI  
Observers

**9%**  
AI  
Laggards

# AI Adoption Maturity Across Enterprises

Gartner forecasts, corroborated by Peak State client intelligence and S&P Global 451 Research, that 40% of enterprise applications will embed task-specific AI agents by end-2026, up from less than 5% in 2025. The 451 Alliance reports that 43% already have agentic AI in active use. IBM reports that approximately 42% of large enterprises have actively deployed AI in business operations, while 59% of IT professionals at these organisations confirm active deployment.

Peak State Consulting | AI Market Segmentation Forecast | June 2026



Source: Peak State Consulting | June 2026

Figure 3: Peak State Consulting, Global AI Market by Segment: 2026 vs 2030 | Total \$2.59T (2026) → \$6.40T (2030) | Source: Peak State Consulting, June 2026

# The Five Maturity Archetypes

---

## AI Pioneers (15%)

Enterprise-wide AI deployment across multiple business functions, supported by dedicated governance structures, proprietary data assets and specialist AI engineering teams. Pioneers are deploying agentic AI, autonomous decision systems and AI-native products. This cohort has grown from 12% in 2025.

## AI Fast Followers (27%)

AI deployed at scale in 2–4 business functions with established AI centres of excellence, data platforms and governance frameworks. Actively scaling successful pilots, though cross-functional integration remains incomplete. Fast Followers represent the largest addressable market for enterprise AI consulting services.

## AI Experimenters (31%)

Multiple AI pilots launched but struggling to transition from pilot to production. Common barriers: data quality issues, integration complexity, change management resistance and lack of clear executive-level AI ownership. Trapped in "pilot purgatory".

## AI Observers (18%)

Exploring AI through vendor briefings and small-scale proofs of concept, but no meaningful investment commitment. High board-level scepticism; concentrated in regulated industries where compliance uncertainty creates inertia.

## AI Laggards (9%)

No meaningful AI activity. Concentrated in traditional sectors lacking digital foundations. This cohort has shrunk from 11% as even conservative organisations begin exploratory AI investment.

<b>ANALYST NOTE</b>	The execution gap remains the central challenge. We forecast 40% of enterprise apps will embed AI agents by year-end; the 451 Alliance independently reports 43% already have agentic AI in active pilot or POC use, different measures pointing to the same conclusion. Organisations must shift focus from AI feature adoption to AI programme execution, deploying governance, MLOps and change management as critical infrastructure.
-------------------------	---

# OPPORTUNITIES FOR ENTERPRISE

Where AI Creates the Most Measurable Business Value

Operational intelligence · Personalisation · Supply chain · R&D · Risk management

**\$2.6–\$4.4T**

Annual AI  
value potential

**26–55%**

Productivity  
gains in pilots

**40–65%**

Downtime  
reduction

**8–15%**

Revenue uplift  
personalisation

# Opportunities: Where AI Creates Enterprise Value

## 1. Operational Intelligence & Process Automation

McKinsey estimates AI could add \$2.6–\$4.4 trillion in annual value across industries through automation of knowledge work. Enterprise deployments report productivity gains of 26–55% in targeted processes (McKinsey Global AI Survey 2025).

### CASE IN POINT

A global insurance carrier deployed LLM-based claims triage automation, reducing processing time from 14 days to 2.3 days while cutting manual review costs by 38%. The programme delivered positive ROI in month seven.

## 2. Customer Intelligence & Hyper-Personalisation

Retail enterprises deploying AI-powered personalisation report revenue uplifts of 8–15% in digital channels (McKinsey, "The value of getting personalisation right", 2021). Financial services firms using AI for next-best-action recommendation engines see cross-sell conversion rates improve by 30–60%. Customer-facing AI applications represented 38.91% of enterprise AI spending in 2025 (Mordor Intelligence).

## 3. Supply Chain Resilience & Predictive Operations

Enterprises with mature AI-driven supply chains operate with inventory levels 20–30% lower than peers while maintaining equivalent service levels. Predictive maintenance programmes reduce unplanned equipment downtime by 40–65% (Deloitte and McKinsey industrial AI benchmarking, 2024).

## 4. AI-Accelerated Innovation & R&D

AI is compressing the time required for hypothesis generation, experimental design and insight synthesis. Healthcare and life sciences show the highest AI investment growth at a projected 20.77% CAGR through 2031 (Mordor Intelligence).

## 5. Risk Management & Regulatory Intelligence

AI-powered transaction monitoring systems detect fraud with 40–60% fewer false positives. The EU AI Act is driving substantial investment in AI-powered compliance monitoring, with compliance-ready vendors gaining measurable market share.

# AI AND THE FUTURE OF WORK

## Workforce Transformation & Organisational Design

How AI is reshaping roles, productivity expectations, and the organisations that deploy it

**26–55%**

Productivity gains reported

**70%**

CEOs concerned about AI talent

**35–65%**

AI specialist salary premium

**2.4×**

Upskilling outperformance

# AI and the Future of Work

AI is reshaping workforce structures across every industry. Rather than replacing all jobs outright, AI is augmenting knowledge work and redefining productivity expectations at pace. The KPMG Global CEO Outlook 2025 found that 70% of CEOs are concerned about competition for AI talent and 77% highlight workforce upskilling as a critical challenge. The organisations that proactively manage this transition will secure a durable competitive advantage.

Research consistently demonstrates that AI-assisted employees achieve productivity gains of 25–55% in targeted workflows (McKinsey Global AI Survey 2025). The nature of these gains is not uniform, they cluster in tasks involving information synthesis, document processing, code generation, and structured communication. Demand for AI product managers, AI governance specialists, and AI security professionals is increasing rapidly, while demand for routine rule-based roles is declining.

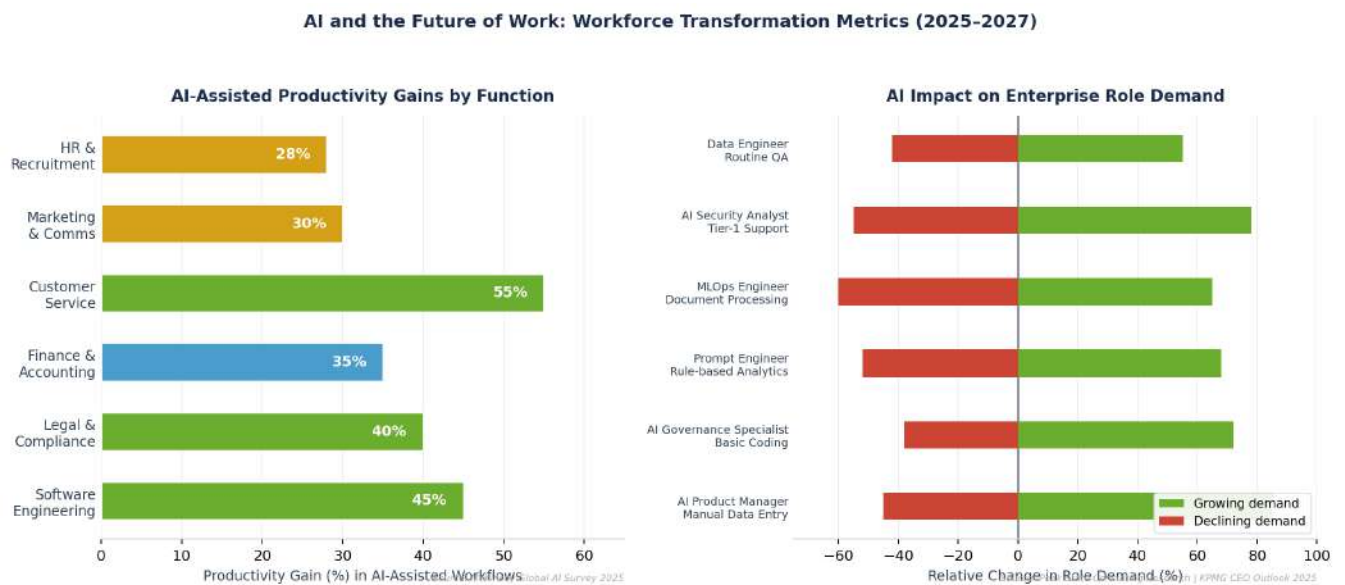


Figure 9: AI and the Future of Work, Productivity Gains by Function and Role Demand Shifts (2025–2027) | Source: McKinsey Global AI Survey 2025, KPMG CEO Outlook 2025, Peak State Consulting Research

## Confronting Job Displacement: Why Reskilling Is Now Non-Negotiable

For all the emphasis on augmentation, enterprise leaders must confront an uncomfortable truth directly: a meaningful degree of job displacement is now becoming inevitable. As agentic AI moves from pilot to production and begins to execute complete workflows autonomously rather than merely assisting with individual tasks, entire categories of routine knowledge work will be substantially automated. Roles built around rule-based processing, structured data entry, first-line query handling, document review, and standard report production are the most immediately exposed.

Peak State Consulting's assessment is that these pressures will intensify materially through 2028, and that organisations which deny or defer this reality will face the most disruptive and least manageable transitions when it finally arrives.

This does not, however, point to wholesale workforce elimination. The historical pattern of major technology transitions is one of role transformation rather than simple subtraction: jobs are redefined, new categories of work emerge, and human effort migrates up the value chain toward judgement, creativity, relationship management and oversight of AI systems. The decisive variable is how deliberately an organisation manages the shift. Enterprises that invest early and systematically in reskilling, redeploying displaced employees into higher-value work rather than simply releasing them, will preserve institutional knowledge, sustain morale, and emerge with a more capable and adaptable workforce. Those that treat workforce reduction as the primary objective of AI adoption will capture short-term cost savings while eroding the very capabilities they need to compete.

Reskilling at the necessary scale is therefore no longer an optional human-resources programme; it is core business infrastructure and a board-level responsibility. With only 20% of executives believing their workforce is AI-ready (KPMG 2025), the gap between the capabilities enterprises have and the capabilities they will need is vast, and it is widening. McKinsey data shows that organisations with structured AI upskilling programmes outperform their peers by 2.4x on AI-driven productivity gains, evidence that workforce development is not a defensive cost but a direct driver of return. The enterprises that will navigate this transition with their reputations and their talent intact are those that begin now: mapping which roles are most exposed, identifying the adjacent higher-value roles into which people can be moved, and building the AI-literacy and continuous-learning pathways that make redeployment genuinely possible. Managed with foresight, the workforce transition becomes a source of durable competitive advantage rather than a crisis to be survived.

## Workforce Transformation & Organisational Design

---

AI transformation is fundamentally an organisational transformation challenge rather than solely a technology initiative. Enterprises deploying AI in production are redesigning workflows, operating structures, leadership responsibilities, and workforce capability models. The rise of AI copilots and autonomous AI agents is changing how work is organised across finance, legal, software engineering, operations, customer service, and research functions.

### Key Workforce Transformation Trends

- AI-augmented knowledge work: AI copilots are becoming the default interface for knowledge workers across enterprise functions, not replacing human judgement but dramatically accelerating information synthesis, drafting, analysis and decision support.
- Smaller, higher-leverage teams: AI-augmented teams are achieving outputs previously requiring significantly larger headcounts. This is reshaping enterprise organisation design from function-based structures toward smaller, product-centric, AI-enabled teams.
- Increased demand for AI governance and security specialists: As AI deployment scales, enterprises urgently need professionals who can govern, audit and secure AI systems, a capability gap that is acute across all markets.
- AI literacy as a core enterprise competency: Forward-looking organisations are treating AI literacy, the ability to work effectively with AI tools and understand their limitations, as a baseline competency requirement across all knowledge worker roles.
- Continuous workforce upskilling programmes: Enterprises investing in structured AI upskilling programmes significantly outperform peers in adoption success and employee retention. This includes both technical training and broader AI fluency programmes.

### Leadership Priorities for Workforce Transformation

---

Enterprises that navigate the workforce transformation successfully share a common set of leadership practices. Transparent employee communication about how AI will change roles and responsibilities is essential, uncertainty and rumour drive resistance. Executive AI education ensures that leaders can make credible AI decisions rather than delegating entirely to technologists. Internal AI champions and communities of practice accelerate adoption by demonstrating real-world value in specific business contexts.

#### KEY INSIGHT

The organisations that combine strong AI capability with genuine workforce adaptability will create the most lasting competitive advantage. Technology alone does not differentiate, the ability to deploy, govern, and continuously improve AI across an organisation that embraces change is the actual source of sustainable competitive advantage in the AI era.

## The Two-Track Labour Market: Evidence from PwC's 2026 AI Jobs Barometer

---

PwC's 2026 Global AI Jobs Barometer, released in June 2026 and based on analysis of more than one billion job advertisements across 27 countries and territories, provides one of the most comprehensive evidence bases yet on how AI is reshaping jobs, skills, wages and productivity. Its central finding reinforces the direction set out above: rather than uniformly displacing workers, AI is increasing the premium employers place on human skills such as judgement, creativity and leadership, while the companies most able to use AI continue to expand hiring faster than their peers. For enterprise leaders, the Barometer reframes the workforce question from whether to adopt AI to how quickly an organisation can build the capability to use it well.

PwC describes the emergence of a two-track labour market. In "professionalised" roles, AI automates routine tasks so that human judgement and expertise are emphasised; in "democratised" roles, AI makes the work itself easier for non-experts to perform. According to PwC, professionalised roles such as radiologists and recruiters are seeing twice the growth in available jobs and 42% faster salary growth than democratised roles such as IT service managers and medical secretaries. As Joe Atkinson, Global Chief AI Officer at PwC, observed, the companies seeing the greatest returns are using AI to amplify human expertise rather than simply to automate, and are pulling further ahead on productivity and growth as a result.

### Key Findings from PwC's 2026 AI Jobs Barometer

1. An AI capability divide in hiring and wages: PwC reports that the companies most able to use AI are growing headcount faster than the least AI-exposed companies (52% versus 36%) and delivering higher wage growth (24% versus 17%), measured against 2018 baseline levels.
2. A "super-star" productivity effect: According to PwC, the top 20% of the most AI-exposed companies achieved average labour productivity growth of 163% relative to 2018, nearly five times higher than the most AI-exposed companies overall. More broadly, the most AI-exposed sectors recorded 34% productivity growth in 2025 relative to 2018, against 24% for the least AI-exposed.
3. A rising wage premium for AI skills: PwC finds the average wage premium for workers with AI skills has risen to 62%, up from 57% a year earlier, reaching as high as 118% in sectors such as consumer markets and around 16% in government and public sector work.
4. AI roles outgrowing the wider market: PwC reports that jobs requiring specific AI skills, such as prompt engineering or machine learning, are growing roughly eight times

- (69%) faster than the overall jobs market (9%), with the number of AI jobs almost twice as high as in 2024. Technology, media and telecommunications (11%) and professional services (6%) led AI job growth, while health sat lowest at under 1%.
5. A diverging entry-level outlook: Drawing on 2.4 million US entry-level job postings, PwC finds that entry-level roles most exposed to AI are seven times more likely to require traditionally senior-level human skills such as leadership, creativity and face-to-face interaction. These "seniorised" entry-level openings have grown 35% since 2019, while other entry-level roles declined by 10%.

These findings carry a direct implication for enterprise talent strategy. As Pete Brown, Global Workforce Leader at PwC, noted, AI is removing some of the routine work that once served as an apprenticeship while increasing demand for judgement, leadership and adaptability earlier in careers, and organisations must rethink how they develop talent accordingly. For Peak State's clients, the practical priorities are consistent with the Barometer's evidence: invest deliberately in AI skills to capture the wage and productivity premium, redesign early-career pathways now that routine tasks no longer provide the traditional training ground, and treat the AI capability divide as a competitive variable rather than a purely operational one.

# CEO CHALLENGES

## Strategy, Governance & Organisational Change

Six principal challenge clusters that determine whether enterprise AI programmes create value or consume it

**71%**

CEOs: AI is top priority

**27%**

Have a comprehensive AI strategy

**52%**

Data quality as top barrier

**30%**

Cite change resistance

## Challenges for CEOs: Strategy, Governance & Change

Only 27% of executives have a comprehensive AI strategy, and only 20% believe their workforce is AI-ready.

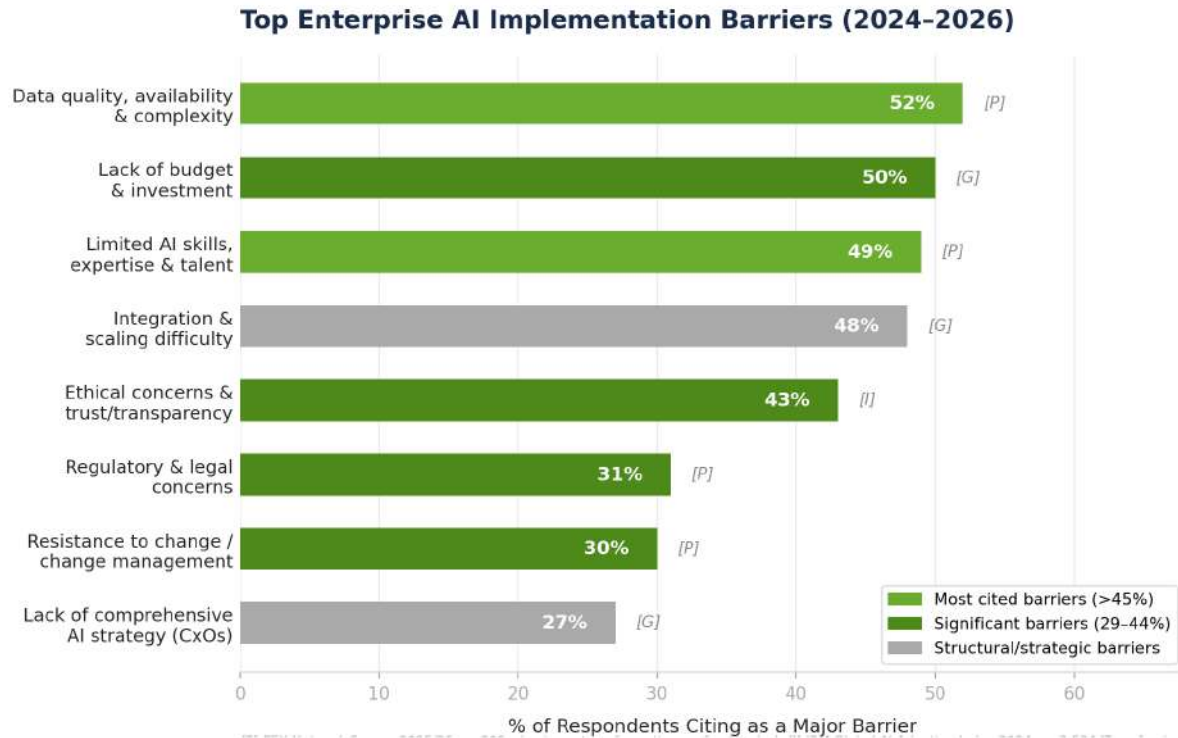


Figure 4: Top Enterprise AI Implementation Barriers (2024–2026) | Source: Peak State Consulting Research; PEX Network Survey 2025/26 (n=200+); IBM Global AI Adoption Index 2024 (n=8,584)

### Challenge 1: Defining a Coherent AI Strategy

Only 27% of executives have a comprehensive AI strategy (Gartner, December 2025). A genuine enterprise AI strategy requires specific value creation hypotheses, prioritised use cases, a realistic assessment of organisational AI readiness, a multi-year investment roadmap and a governance structure.

#### CEO ACTION POINT

Before the next budget cycle, commission an independent review, not one led by a vendor with products to sell. A credible AI strategy addresses six things: where value will be created, what capabilities are missing, whether your data is ready, how you will govern AI, how you will prioritise investments, and how you will measure success.

## **Challenge 2: Managing the Board and Investor Narrative**

---

AI startups now trade at a 3.2x valuation premium over traditional tech (PwC MoneyTree 2026). 59% of CEOs express concerns about ethical implications, 52% about data readiness and 50% about lack of regulation (KPMG Global CEO Outlook 2025).

## **Challenge 3: Building and Retaining AI Talent**

---

70% of CEOs are concerned about competition for AI talent, and 77% highlight workforce upskilling as a challenge (KPMG Global CEO Outlook 2025). Demand for AI engineers, data scientists, ML engineers, AI product managers and AI governance specialists significantly exceeds supply in every major market.

## **Challenge 4: Change Management and Cultural Transformation**

---

30% of business professionals cite resistance to change as a significant barrier to AI adoption (PEX Network Survey 2025/26). Peak State research found only 20% of executives believe their workforce is truly AI-ready. Organisations treating change management as an afterthought consistently underperform AI implementation targets by 30–50% (McKinsey, 2025).

## **Challenge 5: Measuring AI ROI and Value Attribution**

---

AI programmes frequently deliver value across multiple dimensions simultaneously, cost reduction, revenue enhancement, risk mitigation, quality improvement, making clean attribution to individual AI initiatives difficult. Peak State Consulting recommends a balanced AI value scorecard tied to pre-agreed baselines established before AI deployment.

## **Challenge 6: Navigating the Regulatory Landscape**

---

The EU AI Act entered into force on 1 August 2024; most high-risk and transparency obligations apply from 2 August 2026, with full application by 2 August 2027. Banking, financial services and insurance held 23.67% of enterprise AI market share in 2025 (Mordor Intelligence) and face the most immediate regulatory pressure.

# CTO & CIO CHALLENGES

## Architecture, Security & Agentic AI Risk

Data quality, legacy integration, MLOps, compute costs, and the new cybersecurity threat surfaces of agentic AI

**52%**  
Data quality  
top barrier

**48%**  
Integration  
difficulty

**60%**  
Projects abandoned  
poor data

**50%**  
Cite budget  
as constraint

## Challenges for CTOs & CIOs: Architecture & Security

CEOs worry about strategy and governance. CTOs and CIOs have a different set of problems, equally difficult but more technical in nature.

### The Data Quality Crisis

---

Data quality and availability is the greatest barrier to AI adoption, cited by 52% of professionals (PEX Network Survey 2025/26, n=200+). Gartner predicts that through 2026, organisations will abandon 60% of AI projects unsupported by AI-ready data.

#### CTO INSIGHT

Treat data as the foundational asset of your AI strategy. A dedicated Data Foundation Programme, encompassing data quality remediation, metadata management, lineage documentation and governance policy, should be a prerequisite workstream for any enterprise AI programme of material scale.

### Legacy System Integration Complexity

---

Integration difficulties are cited by 48% of infrastructure and operations leaders as a top AI adoption challenge. Most enterprise IT environments combine 15–30 year-old systems of record with more recent digital platforms.

### Model Governance, MLOps and Explainability

---

Enterprise AI systems must be governable, auditable and, in regulated contexts, explainable. 45% of organisations with high AI maturity keep AI projects operational for three or more years, compared to only 20% in low-maturity organisations (Gartner, 2025). The enterprise AI platform market is growing at 27.7% CAGR (Verdantix).

### Compute, Cost and Cloud Architecture

---

NVIDIA captured above 80% of AI accelerator shipments in 2025 and reported record full-year fiscal 2026 revenue of \$215.9 billion, a 65% year-on-year increase (NVIDIA, February 2026). Lack of budget is cited by 50% of I&O leaders as a top AI adoption challenge (Gartner I&O Survey 2025).

# AI Cybersecurity Risks in the Agentic AI Era

Agentic AI systems introduce cybersecurity attack surfaces that go well beyond traditional software security concerns. Unlike passive AI models that only respond to queries, agentic AI systems can take actions, calling APIs, executing code, browsing the web, managing files and interacting with enterprise systems. This autonomous capability dramatically expands the attack surface that security teams must defend.

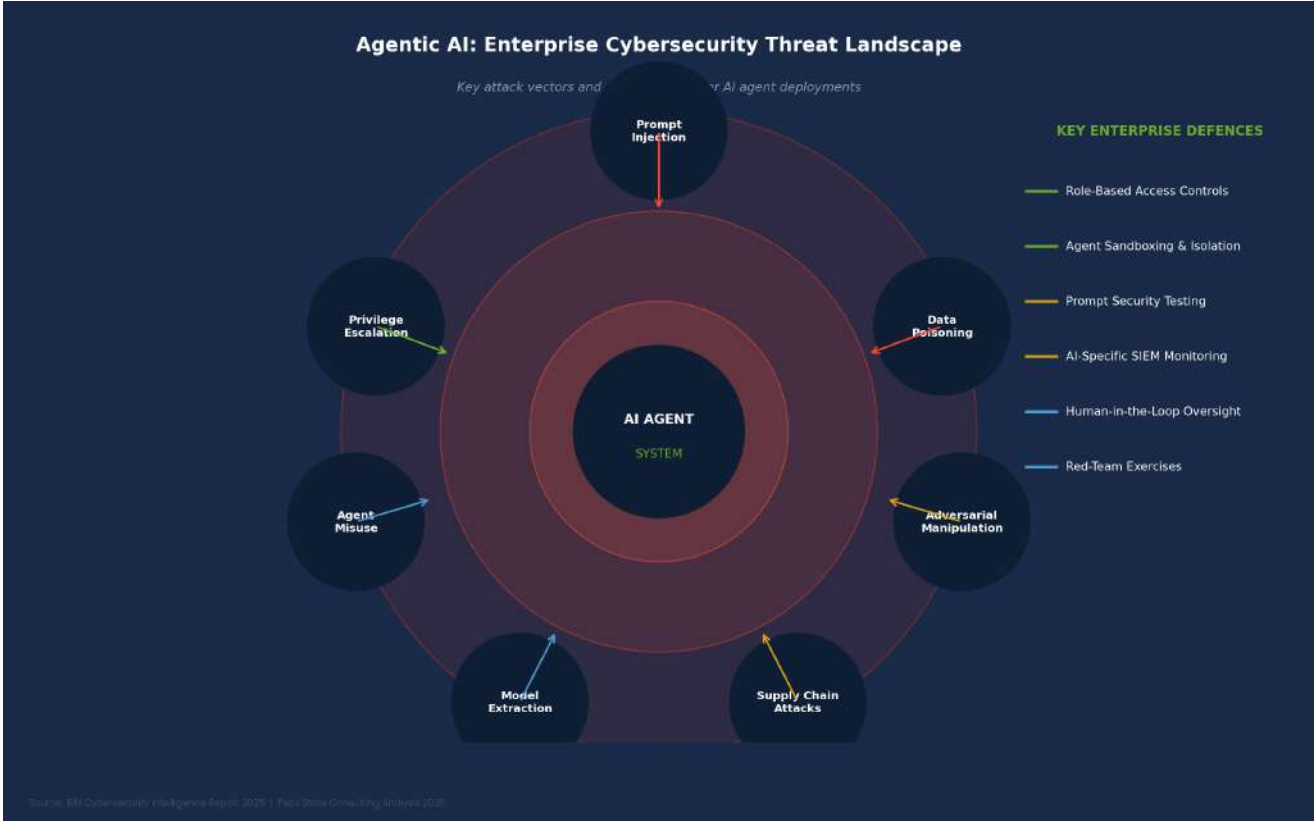


Figure 10: Agentic AI Cybersecurity Attack Surfaces | Source: IBM Cybersecurity Intelligence Report 2025, Peak State Consulting Analysis 2026

## Primary Attack Vectors in Agentic AI Deployments

1. Prompt injection: Malicious instructions embedded in data processed by the AI agent, designed to override the agent's system prompt and redirect its actions. Particularly dangerous in agents with access to email, documents or web content.
2. Data poisoning: Corrupting the training data or knowledge base that an AI system relies on, causing the model to produce systematically biased, incorrect, or harmful outputs that are difficult to detect.

3. Autonomous agent misuse: Exploiting an AI agent's authorised access to systems to perform actions outside its intended scope, including data exfiltration, privilege escalation, or lateral movement within enterprise systems.
4. Model extraction: Using carefully crafted queries to reverse-engineer a proprietary AI model's underlying parameters, effectively stealing the model through its API.
5. Adversarial manipulation: Subtle perturbations to inputs that cause AI systems to make systematically incorrect classifications or decisions, particularly relevant in AI-assisted fraud detection or security monitoring.
6. Supply chain attacks: Compromising AI models, datasets, or tooling in the software supply chain before deployment, analogous to traditional software supply chain attacks but with the additional risk of the AI system actively participating in the attack.

### Enterprise Security Architecture for Agentic AI

Peak State analysis, consistent with IBM research, shows AI-related cybersecurity spending is expected to increase by more than 40% annually through 2028. Enterprises deploying AI agents must implement layered trust architectures. Role-based access controls should govern which actions each AI agent is authorised to take, applying the principle of least privilege to AI agents as rigorously as to human users. Sandboxing isolates AI agents from production systems during testing and limits blast radius during incidents. AI-specific security monitoring, including prompt logging, action auditing, and anomaly detection on agent behaviour, provides the visibility needed to detect and respond to AI-specific attacks. Human-in-the-loop oversight for high-stakes actions remains a non-negotiable architectural requirement in regulated environments.

#### CTO SECURITY NOTE

Treat AI agents as privileged users, not trusted tools. Apply the same access governance, monitoring, and incident response processes to AI agents that you apply to human privileged users. Anthropic's Constitutional AI architecture provides a layer of built-in refusal behaviour for harmful requests, but enterprise security architecture should never rely on model-level safety as the sole defence.

# SOVEREIGN AI & GEOPOLITICAL STRATEGY

## Data Localisation, Regulation & Geopolitical Risk

Why sovereign AI is now a board-level strategic issue – and how global enterprises must respond

**70%**

EU enterprises must re-architect by 2027

**45%**

OpenRouter traffic now Chinese AI

**25%**

Chinese API share by 2027

**4**

Major regulatory regimes in conflict

# Sovereign AI & Geopolitical Strategy

The age of treating AI as a purely commercial purchase is over. In the space of a single extraordinary week in June 2026, the United States, the European Union and Canada each unveiled competing national visions for artificial intelligence, and at the G7 summit in Evian the chief executives of the world’s most powerful AI companies took their seats alongside heads of state. Artificial intelligence has arrived at the very top of the geopolitical agenda, and it is not leaving. For enterprise leaders, the geopolitical dimension of AI is no longer a distant background consideration. It has become a frontline strategic and operational challenge that can reshape vendor access, data flows and compliance obligations overnight.

Governments now regard AI infrastructure, semiconductor supply chains, cloud capacity and foundation models as matters of core national strategic importance, on a par with energy security and defence. The rapid rise of Chinese AI providers, including DeepSeek, Alibaba Qwen, Baidu ERNIE and ByteDance Doubao, has intensified the competition, while Western governments tighten controls on semiconductor exports, advanced compute access and infrastructure concentration. The stakes were made vivid in June 2026, when access to the most advanced frontier models was abruptly restricted by government order, demonstrating that a single political decision in one capital can sever an enterprise’s access to critical AI capabilities anywhere in the world. This is the new reality every board must now plan for.



Figure 11: Sovereign AI, Global Regulatory Landscape & Enterprise Implications | Source: European Commission, NIST, UK ICO, PRC AI Regulations, Peak State Consulting Research

## The G7 Inflection Point: AI Takes the Geopolitical Stage

---

June 2026 will be remembered as the moment AI sovereignty stopped being a theoretical concern and became hard strategy. At the G7 summit in Evian-les-Bains, France, the chief executives of the three most powerful Western AI companies, OpenAI, Anthropic and Google DeepMind, joined heads of state for a working session on the safe and effective deployment of artificial intelligence. Their presence at the table, alongside presidents and prime ministers, was itself the signal: technology leaders now wield geopolitical influence once reserved for nation states. Behind closed doors, those same leaders reportedly urged governments to build a coordinated international coalition to govern advanced AI, even as each major economy raced to assert its own independent position.

The week surrounding the summit saw an unprecedented cascade of competing national announcements. The United States issued an executive order on 2 June setting out a framework for oversight of the most advanced AI systems. The European Commission published its long-awaited European Tech Sovereignty package on 3 June, with proposals to boost EU-owned chip manufacturing, sovereign cloud infrastructure and open-source AI to reduce dependence on US technology giants. Canada launched its national AI strategy, “AI for All,” on 4 June, anchored on sovereign compute, a public AI supercomputer and a new Sovereign Technology Alliance. The message from Evian was unmistakable: in the AI era, no serious economy can afford to be a passive consumer of someone else’s technology.

## Why Every Nation Now Needs a National AI Plan

---

A National AI Plan and accompanying digital strategy is fast becoming as fundamental to a country’s future prosperity as its energy policy, its education system or its currency. This is no longer an aspiration for technologically advanced states alone; it is a baseline requirement of national competitiveness and security. Global spending on sovereign AI systems is projected to surpass \$100 billion in 2026, and government is on track to be the single fastest-growing AI-spending vertical, with public-sector AI investment compounding at roughly 19% a year. The nations setting the pace are not waiting for consensus. They are building.

The reference programmes are striking in their ambition. The United Arab Emirates, the first country to adopt a national AI strategy and appoint a dedicated AI minister, targets AI contributing up to 45% of national GDP by 2031, with Abu Dhabi running more than 100 AI use cases in production across over 40 government entities under a multi-billion-dollar digital strategy and a stated goal of 100% sovereign cloud adoption. France has committed €109 billion to AI infrastructure through 2030 in what President Macron has framed explicitly as a fight for strategic autonomy, backing national champion Mistral AI and a sovereign compute stack. The United Kingdom’s AI Opportunities Action Plan and its £500 million Sovereign AI Unit aim to drive both economic growth and national security. The European Union’s AI Continent Action Plan, Cloud and AI Development

Act, and EUROPA frontier-model programme together pursue a compliance-first sovereign ecosystem, while Canada's "AI for All" strategy targets up to 250,000 new AI-enabled jobs by 2031. These are not policy papers gathering dust. They are funded, operational national programmes, and they are reshaping where AI capital, talent and capability will concentrate over the next decade.

For enterprise leaders, the implication is direct and urgent. The countries in which you operate are now actively shaping the AI landscape you must navigate, through data-residency rules, sovereign-cloud mandates, procurement preferences and export controls. Those positioned within strong national AI ecosystems will enjoy privileged access to compute, talent and frontier capability; those that are not may find themselves on the wrong side of a widening divide. The question facing every government, and by extension every enterprise that depends on one, is no longer whether to have a National AI Plan, but how quickly it can build one that works.

## Key Sovereign AI Challenges for Enterprise Leaders

---

1. **Data localisation requirements:** Governments in the EU, China, India, Saudi Arabia and other jurisdictions require that certain categories of data be stored and processed within national borders. For AI systems, this means that cloud-hosted AI inference must be performed within the relevant jurisdiction, creating infrastructure complexity for global enterprises.
2. **Export control restrictions:** US government export controls now restrict the sale of advanced semiconductors (including NVIDIA H100/H200/Blackwell chips) to certain countries. For AI vendors and enterprise customers in restricted regions, this limits access to the most capable AI infrastructure.
3. **Regional AI regulation conflicts:** The EU AI Act, US executive orders on AI, China's generative AI regulations, and emerging frameworks in India, UAE and Brazil create a patchwork of conflicting obligations. A single AI deployment must satisfy multiple overlapping regulatory regimes for global enterprises.
4. **AI infrastructure concentration risks:** Heavy dependency on a single AI vendor or cloud provider creates operational risk. A vendor security incident, geopolitical event, or regulatory enforcement action against a provider could disrupt enterprise AI operations.
5. **Cloud sovereignty requirements:** Several national governments require that enterprise AI workloads involving sensitive data be processed on nationally-operated or regionally-sovereign cloud infrastructure.

## Strategic Response: Multi-Region AI Architecture

---

Many multinational enterprises are adopting multi-region AI strategies that include regional AI deployment architectures with separate infrastructure and governance configurations by geography; hybrid cloud AI infrastructure combining global cloud providers with regional sovereign cloud services; localised data residency controls that route sensitive data to the appropriate regional AI endpoint; and vendor diversification strategies that avoid over-dependence on any single AI provider or geopolitical bloc.

Boards and CIOs increasingly recognise AI sovereignty as both a commercial and an operational-resilience issue, and the events of June 2026 have removed any remaining doubt. Peak State Consulting recommends that every global enterprise with operations in more than three jurisdictions commission a Sovereign AI audit without delay, to identify current compliance gaps, infrastructure-concentration risks, single points of geopolitical failure, and the vendor-diversification options available to them. In an environment where model access can change with a single government decision, resilience is no longer optional.

**BOARD PRIORITY** Sovereign AI is a board-level risk, not just a technology procurement decision. Enterprises operating across multiple jurisdictions must establish a clear policy on data residency, AI vendor concentration, and regulatory compliance across their full AI technology stack. The cost of reactive compliance significantly exceeds the cost of proactive strategy.

# THE VENDOR LANDSCAPE

## Selecting the Right AI Technology Partners

Hyperscalers, foundation model leaders, Chinese AI providers – and how to evaluate them rigorously

# The Vendor Landscape: Selecting AI Partners

The AI vendor market is moving faster than most procurement cycles. Microsoft, Google Cloud, AWS, Oracle, SAP, Salesforce, Baidu and a growing number of pure-play foundation model companies are all competing for enterprise budgets, and the differences between them matter enormously.

Enterprise AI Vendor Landscape Map (2026)

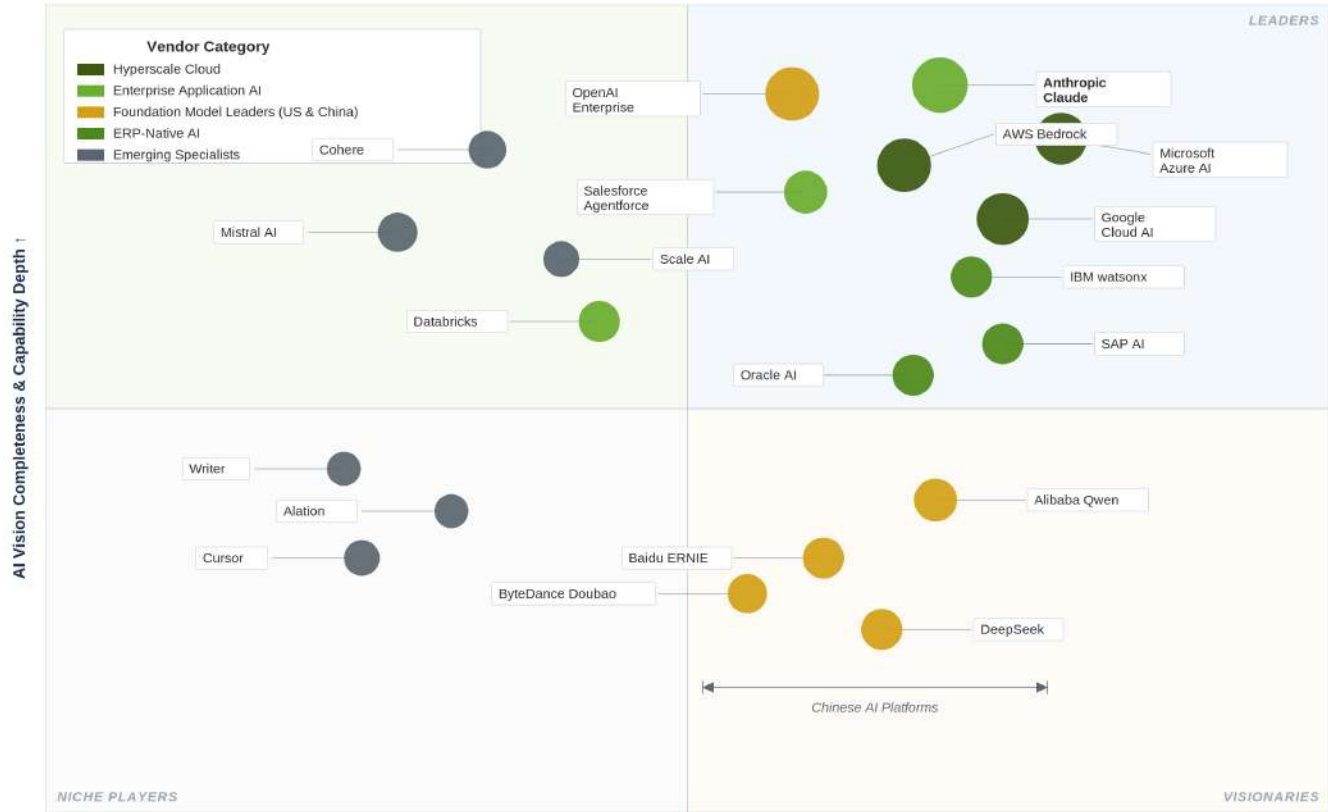


Figure 5: Enterprise AI Vendor Landscape Map (2026) | Source: Peak State Consulting Research.

## Hyperscale Cloud Platforms: Microsoft, Google Cloud, AWS

---

The three major cloud hyperscalers are consolidating dominance. Microsoft Azure AI, backed by its deep OpenAI investment, is the most widely deployed enterprise AI platform.

### CRITICAL WARNING

Vendor selection decisions made under time pressure, driven by vendor-led demonstrations or without independent technical validation frequently result in platform migration programmes within 24 months. The cost of AI platform migration typically exceeds the cost of the original implementation. Invest in rigorous independent evaluation upfront.

## Pure-Play Foundation Model Leaders

---

OpenAI and Anthropic dominate the enterprise foundation model API market. Anthropic raised \$30 billion in a Series G round in February 2026 and a further \$65 billion in a Series H round announced 28 May 2026, valuing the company at \$965 billion, making it the most valuable private AI company in the world (Bloomberg, CNBC, May 28, 2026). Anthropic's Claude models are valued for safety, instruction-following reliability and long-context performance.

## The Chinese AI Challenge

---

Chinese AI providers now account for over 45% of OpenRouter token traffic (Q2 2026 data), up from less than 2% one year earlier. DeepSeek's architectural innovations have demonstrated frontier-class performance at API pricing five to ten times lower per token than equivalent US frontier models.

## The Peak State Vendor Evaluation Framework

---

- Capability breadth and depth: Does the vendor's AI offering cover priority use cases with demonstrated accuracy in comparable production environments?
- Enterprise readiness: Are the security controls, compliance certifications, SLA commitments, support structures and data residency options adequate?
- Integration ecosystem: How well does the AI platform integrate with the enterprise's existing technology stack, data platforms and development workflows?
- Commercial model sustainability: Is the pricing model transparent, predictable and sustainable at the scale of the enterprise's anticipated AI usage?
- Strategic roadmap alignment: Does the vendor's development roadmap align with anticipated future requirements? Is the vendor financially stable?
- Partner and consultant ecosystem: Is there a mature ecosystem of implementation partners with demonstrated experience in similar enterprise contexts?

# LEADING AI PLATFORMS

Comparative Analysis: US and Chinese Coding & Development Tools

Claude Code · GitHub Copilot · Cursor · Windsurf · Gemini · DeepSeek · Qwen · ERNIE · Doubao

**87.6%**

Claude Code  
SWE-bench (Opus 4.7)

**\$2B**

Cursor ARR  
Feb 2026

**50×**

DeepSeek price  
advantage vs US

**49%**

Chinese AI coding  
token share Q2 2026

## Leading AI Platforms: Comparative Analysis

The AI developer tools and coding platform category represents one of the fastest-growing segments of the enterprise AI market. Cursor reached \$2 billion in annualised revenue by February 2026, the fastest zero-to-\$2B arc in B2B software history. The competitive landscape has rapidly evolved from AI-assisted autocomplete to fully autonomous agentic coding systems.

### US AI Coding & Development Platforms

---

#### **Claude Code (Anthropic)**

Claude Code achieves 87.6% on SWE-bench Verified with Opus 4.7 (Q1 2026); Claude Fable 5 is now selectable in Claude Code and exceeds this benchmark. Integrates natively with VS Code and JetBrains. Pricing: Claude Pro \$20/month, Claude Max \$100/month (5x usage), Max \$200/month (20x usage), or API usage at approximately \$3/M tokens (Sonnet 4.6) or \$5/M tokens (Opus 4.8).

#### **GitHub Copilot (Microsoft / OpenAI)**

The enterprise default for AI-assisted development. Pro tier \$10/month; Pro+ \$39/month includes Claude Fable 5, GPT-5.5 and o3. Enterprise pricing \$39/seat/month with IP indemnification. AI Credits system active from June 2026 for managing premium model usage.

#### **Cursor (Anysphere)**

Rebuilt VS Code around AI with native agent mode. \$20/month (Pro) and \$40/month (Business). Cursor reached \$2 billion ARR by February 2026 from \$100M in January 2025, and is in advanced talks to raise \$2 billion at a \$50+ billion valuation.

#### **Windsurf (Cognition)**

Formerly Codeium, acquired by Cognition on 14 July 2025. At \$15/month, Windsurf undercuts both Cursor and Copilot Pro while offering competitive agentic capabilities.

#### **Gemini Code Assist (Google)**

Integrates deeply with Google Cloud, BigQuery and Google Workspace. Powered by Gemini 3.1 Pro (current GA model) and Gemini 3.5 Flash (launched May 2026 at \$1.50/\$9.00 per million tokens). Google is migrating Code Assist to its new Antigravity multi-agent platform from June 18, 2026. Gemini now bundled in Google Workspace Business Standard, Plus and Enterprise at no extra charge.

### AI Coding Platform Comparison — Key Enterprise Dimensions (2026)

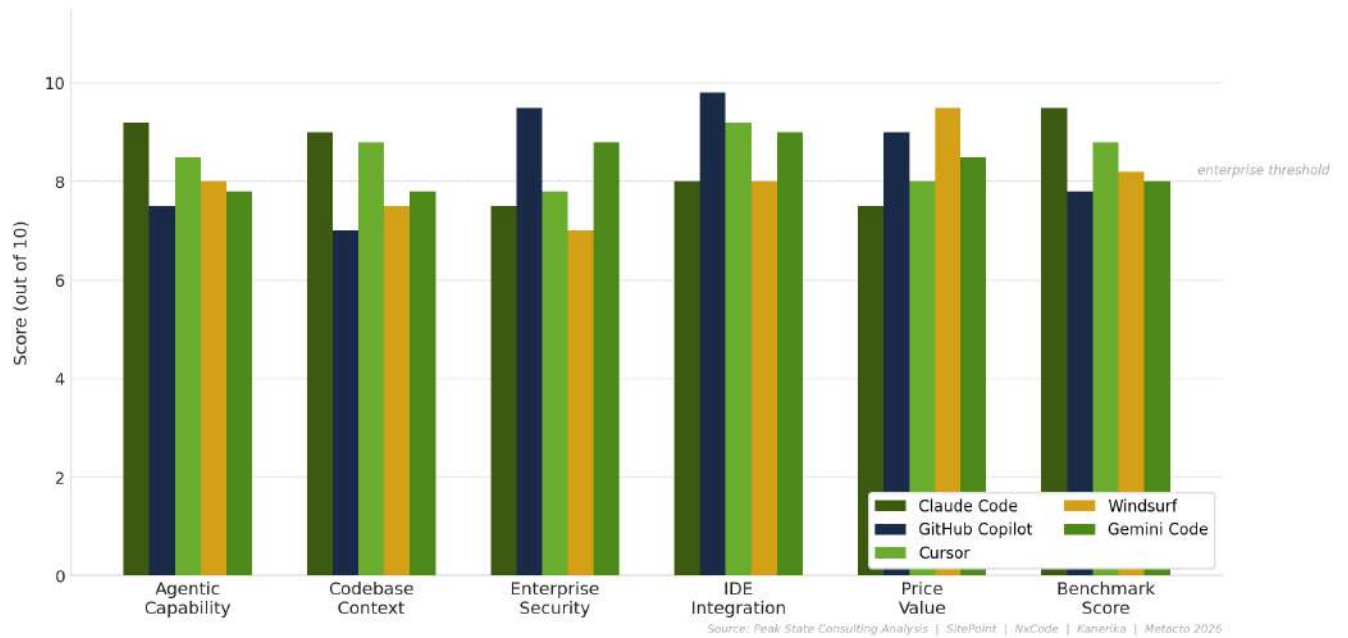


Figure 6: AI Coding Platform Comparison, Key Enterprise Dimensions (2026) | Source: Peak State Consulting Analysis, SitePoint, Kanerika, NxCode, Metactio 2026

# Enterprise Coding Platform Comparison Matrix

PLATFORM	AGENTIC MODE	SECURITY	PRICE / USER / MONTH	BEST USE CASE
<b>Claude Code (Anthropic)</b>	Terminal-native; 87.6% SWE-bench (Opus 4.7); Fable 5 now selectable	API-based; no training on customer data; SOC 2 Type II	Pro \$20 · Max \$100/\$200 · API ~\$3–5/M tokens	Multi-file refactoring, large codebase reasoning, autonomous tasks
<b>GitHub Copilot (Microsoft)</b>	Issue-to-PR workflow; native GitHub	Highest: SSO, audit, IP indemnity, org policy	Pro \$10 · Pro+ \$39 · Enterprise \$39/seat	Enterprise teams, GitHub-native workflows, compliance-heavy orgs
<b>Cursor (Anysphere)</b>	IDE-native background agents; model flexible	SAML SSO, privacy mode (no code storage)	Pro \$20 · Business \$40 · Enterprise custom	AI-native IDE, daily coding velocity, modern dev teams
<b>Windsurf (Cognition)</b>	Cascade agent; SWE-1.5; high-speed inference	Standard enterprise controls	Pro \$15 · Enterprise custom	Price-performance leader; Cursor alternative
<b>Gemini Code Assist (Google)</b>	Agent mode; multimodal	Enterprise grade; Google Cloud native	Custom Workspace Enterprise	Google Cloud ecosystems, multimodal tasks

## Chinese AI Coding & Enterprise Platforms

Chinese AI providers have made remarkable inroads into the developer tools market. Chinese models accounted for approximately 49% of all coding tokens on OpenRouter by Q2 2026 (Digital Applied Research).

### DeepSeek

DeepSeek V3.2 (March 2026) delivers frontier-class performance at approximately \$0.55/million input tokens, roughly five to ten times cheaper than equivalent US frontier models.

### Alibaba Qwen

Qwen3.5 (February 2026) claims more than 100 million monthly active users. The Apache 2.0 licence on Qwen 2.5 enables enterprise self-hosting and derivative models.

**Baidu ERNIE**

ERNIE 5.0 (released early 2026) is the standard enterprise AI choice for companies operating in the Chinese market.

**ByteDance Doubao**

Doubao has 155 million weekly active users, making it one of the most widely used consumer AI assistants in China. ByteDance's regulatory position in Western markets creates significant enterprise risk for non-China deployments.

<p><b>GEOPOLITICAL RISK NOTE</b></p>	<p>Western enterprises deploying Chinese AI models must carefully evaluate data sovereignty, export control risks, EU AI Act regulatory constraints and reputational considerations. A two-tier AI architecture, US/EU models for regulated workloads, Chinese open-weight models for cost-optimised back-office processing, is emerging as a pragmatic approach.</p>
--	---

# THE FUTURE OF AI AGENTS

## Autonomous AI in Enterprise Operations

How AI agents are moving from assistants to autonomous workflow actors — and what governance they require

**40%**

Enterprise apps  
with agents end-2026

**43%**

Have agentic AI  
in use now

**\$1.4T**

Agent spend  
by 2027

**4**

Architecture layers  
for enterprise agents

# The Future of AI Agents in Enterprise Operations

AI agents have moved well beyond chatbots. They now perform multi-step reasoning, use tools, call APIs, and execute tasks without human involvement at each step. Gartner forecasts that 40% of enterprise applications will embed task-specific AI agents by the end of 2026, up from less than 5% in 2025. This transition represents the most significant shift in enterprise software architecture since the adoption of cloud computing.

The 451 Alliance reports that 43% of enterprises already have agentic AI in active use. The gap between current deployment and projected adoption illustrates the execution challenge, the technology capability exists but organisational readiness, governance frameworks, and integration infrastructure are the binding constraints on deployment pace.



Figure 12: The Enterprise AI Agent Ecosystem, Four-Layer Architecture | Source: Peak State Consulting Research 2026, Anthropic

## Emerging Enterprise Use Cases for AI Agents

---

The most valuable near-term enterprise agentic AI deployments cluster around tasks that combine information retrieval, structured reasoning, and system interaction, workflows that are too complex for simple automation but sufficiently structured for AI agents to execute reliably within defined parameters.

### **Autonomous Procurement Workflows**

AI agents can monitor supplier performance, identify cost optimisation opportunities, generate purchase orders, manage approval workflows, and alert procurement teams to exceptions, reducing procurement cycle times and freeing specialist staff for relationship management and strategic sourcing.

### **AI-Assisted Financial Reconciliation**

In financial services and corporate finance functions, AI agents can reconcile transactions across multiple systems, identify discrepancies, generate exception reports, and draft regulatory submissions, a workflow currently consuming significant specialist finance staff time with high potential for AI-driven efficiency gains.

### **Customer Support Orchestration**

Multi-agent customer support architectures combine a triage agent (routing queries to the appropriate specialist), a knowledge retrieval agent (surfacing relevant documentation), a resolution agent (drafting responses), and a quality monitoring agent (reviewing outputs before delivery). This architecture delivers near-human quality at a fraction of the cost for routine Tier-1 and Tier-2 support interactions.

### **Intelligent Software Engineering**

Claude Code and comparable agentic coding tools are already delivering measurable engineering productivity improvements at scale. The next generation of coding agents will move beyond individual file edits to managing entire features, from requirements interpretation through implementation, testing, documentation and deployment, with human engineers providing direction, review and escalation handling.

# Governance and Trust Architecture for Agentic AI

---

The next phase of enterprise AI competition will focus on orchestration, governance, and operational trust, not just model capability. The enterprises that deploy AI agents effectively will have invested in three critical organisational capabilities: clear agent authorisation frameworks defining exactly what actions each agent is permitted to take and under what conditions; robust monitoring and anomaly detection that identifies unexpected agent behaviour before it causes material harm; and clear human escalation paths that ensure consequential decisions always have a human accountable for the outcome.

<b>IMPLEMENTATION PRIORITY</b>	Enterprises deploying their first AI agents should start with a narrowly scoped agent with read-only access to systems and clear approval requirements for any write actions. Build governance infrastructure, authorisation frameworks, monitoring, and escalation paths, before expanding agent capabilities. The operational trust built in early-stage deployments determines the speed at which agentic capability can be scaled responsibly.
------------------------------------	--

# THE AI PROGRAMME ROI JOURNEY

A Realistic 36-Month Investment and Return Trajectory

How Pioneers achieve 4× returns by month 36 – and why Late Adopters may never break even

**4×**

Pioneer ROI  
over 3 years

**Month 9**

Pioneer  
break-even

**3×**

Fast Follower  
ROI

**5**

Programme  
phases to value

## The AI Programme ROI Journey

One of the most common and damaging misconceptions about enterprise AI programmes is that they should deliver measurable financial returns quickly. 45% of organisations with high AI maturity keep AI projects operational for three or more years, compared to only 20% in low-maturity organisations (Gartner, 2025). Peak State Consulting's model, based on pattern analysis of simulated enterprise AI programme outcomes, charts the estimated ROI trajectory by adoption archetype.

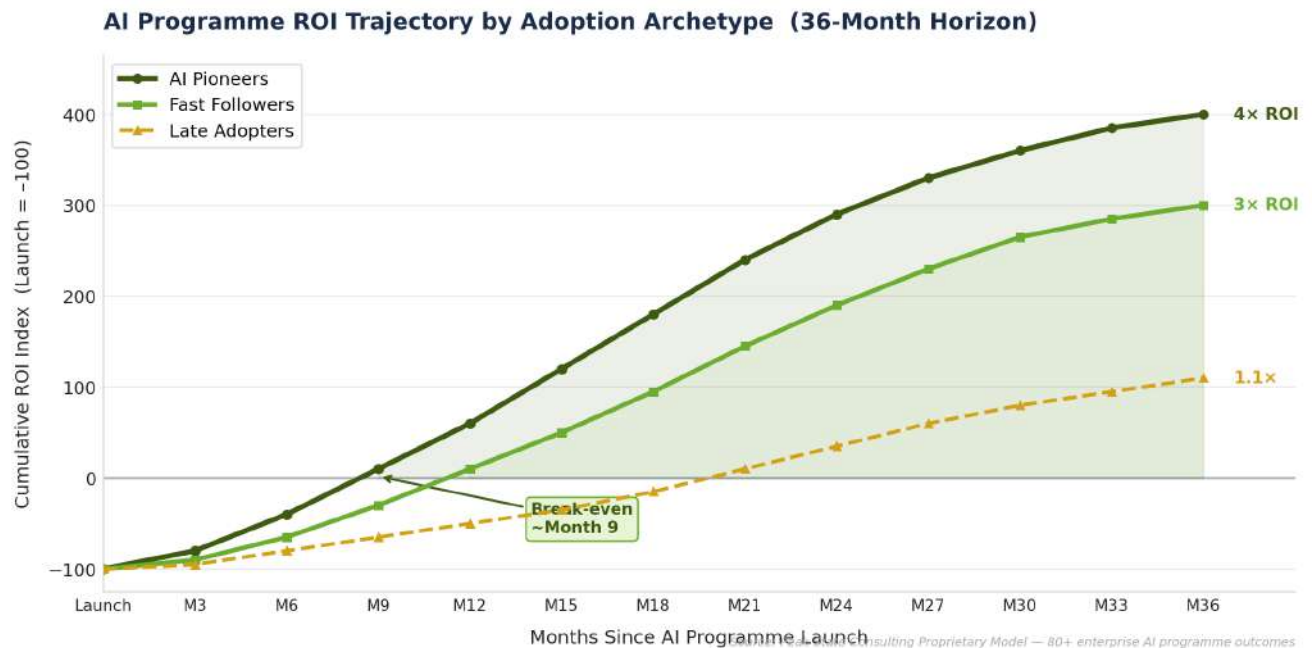


Figure 8: AI Programme ROI Trajectory by Adoption Archetype, 36-Month Horizon | Source: Peak State Consulting Proprietary Model, simulated pattern analysis of select enterprise AI programme outcomes

## Understanding the AI Investment Curve

AI Pioneers achieve positive cumulative ROI at approximately month nine and deliver 4x the initial investment over a three-year horizon. Fast Followers break even at approximately month twelve and deliver roughly 3x. Late Adopters (the Experimenter, Observer and Laggard archetypes) often do not reach positive ROI within three years at all. We should be direct about this: most published AI ROI estimates are optimistic. They measure the successes and undercount the abandonments. The primary differentiator between the tiers is not which model or platform you chose, it is whether your organisation had the discipline to execute.

## The Five Phases of AI Programme Value Creation

---

### **Phase 1, Foundation (Months 0–6)**

Investment in data infrastructure, AI platform deployment, governance framework design and team capability building. Net negative ROI. Key metrics: data quality KPIs, platform stability, first pilot deployment.

### **Phase 2, Pilot Validation (Months 3–9)**

First AI solutions deployed against pre-agreed baselines. Initial business value evidence generated. Key metrics: accuracy rates, user adoption, time savings, first measurable business impact.

### **Phase 3, Scale-Up (Months 6–18)**

Successful pilots scaled to production. Positive ROI achieved on initial use cases. Programme expansion across business units.

### **Phase 4, Integration (Months 12–24)**

AI embedded in core operational workflows. Cross-functional AI integration generating compound value. Agentic AI capabilities introduced.

### **Phase 5, Transformation (Months 18–36+)**

AI-driven business model evolution. New AI-enabled products and services launched. AI as a sustainable competitive differentiator.

# AI READINESS FRAMEWORK

## The Peak State Six-Pillar Assessment

A proprietary assessment methodology validated across 80+ enterprise programme engagements globally

**6**

Assessment  
pillars

**80+**

Client  
engagements

**4-6**

Weeks to  
complete

**3**

Geographies:  
Europe · N.Am · Asia

# Peak State AI Readiness Framework

Peak State Consulting has developed a proprietary AI Readiness Assessment framework providing enterprise leadership teams with a rigorous evaluation across six critical dimensions.

## The Six Pillars of AI Readiness

PILLAR	KEY ASSESSMENT QUESTIONS
<b>1. Strategic Clarity</b>	Is there a board-approved AI strategy with clear value creation hypotheses, prioritised use cases and defined success metrics? Is AI ownership at C-suite level unambiguous?
<b>2. Data Foundations</b>	What is the current state of enterprise data quality, governance and architecture? Are the data assets required for priority AI use cases available, accessible and trustworthy?
<b>3. Technology Architecture</b>	Is the existing technology stack capable of supporting enterprise AI deployment? Are cloud infrastructure, API architecture, integration platforms and MLOps capabilities at the required maturity level?
<b>4. Talent &amp; Capability</b>	Does the organisation have the AI engineering, data science, AI product management and AI governance capabilities required? Where are the critical gaps?
<b>5. Governance &amp; Risk</b>	Are AI governance structures, risk assessment processes, ethical AI principles and regulatory compliance frameworks, including EU AI Act readiness, in place and working in practice?
<b>6. Change Readiness</b>	Is the organisation culturally and operationally prepared for the changes AI deployment requires? Is leadership alignment, employee communication and change management planning adequate?

## How Peak State Consulting Delivers the Assessment

The Peak State AI Readiness Assessment is delivered as a structured two-to-three month engagement, combining executive interviews, technical architecture review, data environment assessment and benchmarking against our proprietary database of enterprise AI programme outcomes.

### ENGAGEMENT OFFER

Peak State Consulting offers a complimentary AI Readiness Diagnostic, a focused two-hour executive discussion, for qualifying organisations. Contact Anuj Khanna at [anuj@peakstate.com](mailto:anuj@peakstate.com) or visit [peakstate.com](https://peakstate.com) to request your session.

# AI OPERATING MODEL FRAMEWORK

## Five-Layer Architecture for Enterprise AI at Scale

How to structure leadership, governance, engineering and business ownership to make AI programmes work

**5**  
Operating  
model layers

**Layer 1**  
Governance  
& Risk

**Layer 4**  
AI Centre  
of Excellence

**Layer 5**  
Executive  
Leadership

# AI Operating Model Framework

The enterprises making the most progress on AI have stopped treating it as a series of projects. They have built operating models around it. The AI operating model defines how AI initiatives are governed, funded, deployed, monitored, and scaled across the organisation, and it is the primary determinant of whether AI delivers sustainable value or remains trapped in pilot programmes. Organisations lacking formal operating models frequently struggle to scale beyond pilot programmes regardless of the quality of their technology choices.

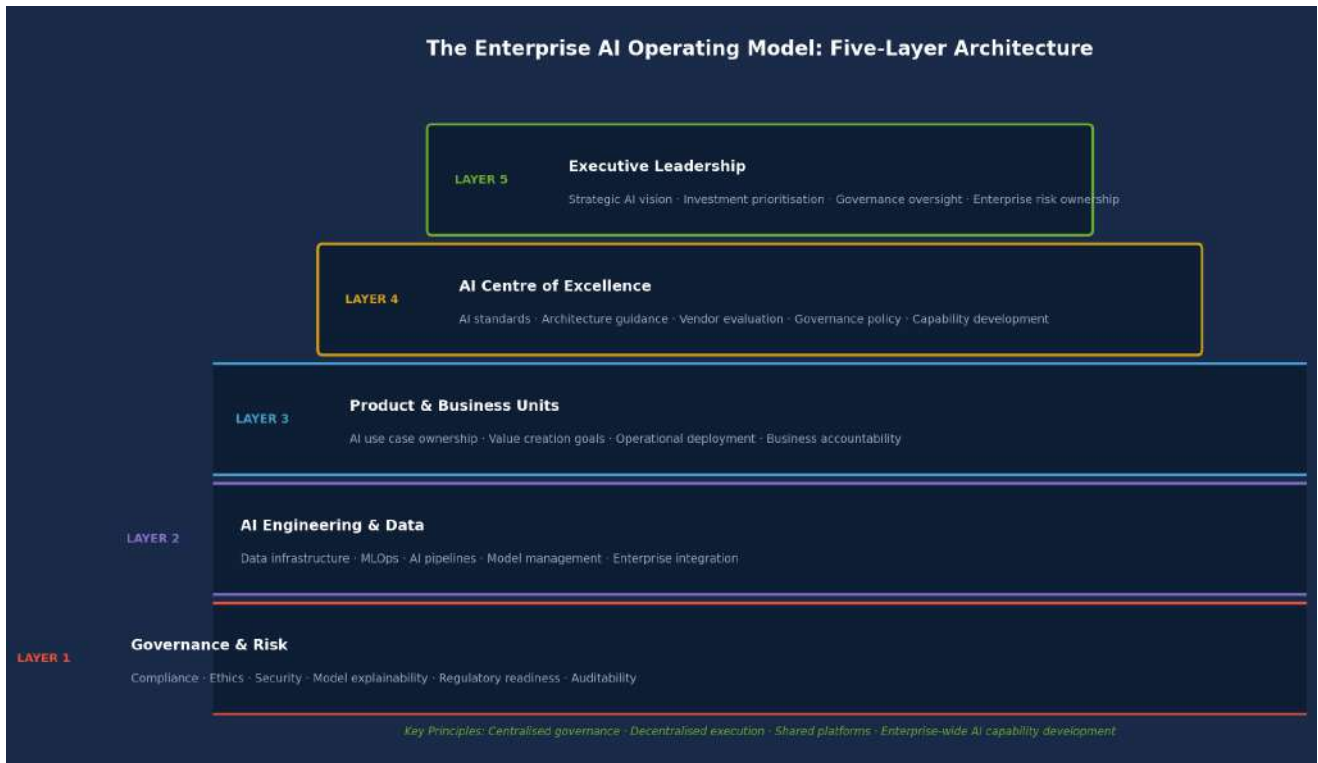


Figure 13: The Enterprise AI Operating Model, Five-Layer Architecture | Source: Peak State Consulting Research, McKinsey State of AI 2025

## The Five Layers of the Enterprise AI Operating Model

---

### **Layer 5, Executive Leadership**

Responsible for strategic AI vision, investment prioritisation, governance oversight, and enterprise risk ownership. The executive leadership layer sets the AI ambition and provides the organisational authority to make the changes necessary to realise it. Without visible executive leadership at CEO and board level, AI programmes consistently underperform against potential.

### **Layer 4, AI Centre of Excellence (AI CoE)**

A cross-functional team responsible for AI standards, architecture guidance, vendor evaluation, governance policy, and capability development. The AI CoE is the connective tissue between executive vision and frontline deployment, it prevents duplication, enforces standards, and maintains the institutional knowledge that makes each successive AI deployment faster and more effective than the last. Deloitte's commitment to certifying 15,000 professionals on Claude illustrates the scale of AI CoE investment in world-class enterprises.

### **Layer 3, Product & Business Unit Layer**

Business-aligned AI product owners responsible for identifying use cases, defining value creation goals, and managing operational deployment. This layer brings domain expertise to AI prioritisation, ensuring that AI investment flows to the problems where it will create the most business value, not just the problems that are technically easiest to solve.

### **Layer 2, AI Engineering & Data Platform Layer**

Responsible for data infrastructure, MLOps, AI pipelines, model management, and enterprise integration. This layer builds and maintains the technical foundation that all AI use cases depend on, it is where the 30–40% of AI programme budget that must go to data foundations and infrastructure is deployed.

### **Layer 1, Governance & Risk Layer**

Ensures compliance, ethics, security, model explainability, regulatory readiness, and auditability. This layer provides the accountability mechanisms, without which AI programmes expose enterprises to regulatory, reputational, and operational risks that can outweigh the value created. For EU enterprises, this layer is where EU AI Act conformity assessments, transparency documentation, and human oversight mechanisms are managed.

## Key Principles of an Effective AI Operating Model

---

1. Centralised governance with decentralised execution: AI standards, risk frameworks, and platform choices are centralised; use case development and deployment are owned by business units closest to the problem.
2. Shared AI platforms and tooling: Enterprise-wide investment in shared infrastructure, data platforms, MLOps tooling, API management, vector databases, reduces duplication and creates compounding returns as each new use case builds on existing foundations.
3. Standardised governance policies: Common model documentation standards, risk assessment processes, and human oversight requirements across all AI deployments, enabling audit and compliance at enterprise scale.
4. Enterprise-wide AI capability development: Systematic investment in AI literacy, technical training, and leadership education, treating AI capability as a workforce asset that compounds over time.
5. Cross-functional accountability: AI success metrics are owned by business leaders, not technology teams, ensuring that AI investment is always connected to business value creation rather than technical achievement.
6. Continuous AI monitoring and optimisation: AI systems degrade without active management. The operating model must include ongoing model performance monitoring, retraining workflows, and continuous improvement processes as core operational disciplines.

**IMPLEMENTATION  
NOTE**

Building an effective AI operating model takes 12–18 months for a large enterprise. The key investment decisions are: appointing an accountable executive leader (Chief AI Officer or equivalent), funding the AI CoE with experienced staff, and investing in shared data and MLOps infrastructure before scaling use case deployment. Organisations that cut corners on operating model design pay substantially more in remediation costs later.

# THE ECONOMICS OF ENTERPRISE AI

## Inference Costs, Optimisation Strategies & AI FinOps

Total cost of ownership, the Chinese price disruption, and how to govern AI infrastructure costs at scale

**90%**

Prompt caching  
cost reduction

**50×**

DeepSeek vs US  
model pricing

**60–80%**

Model routing  
max savings

**40–60%**

Data prep share  
of implementation cost

# The Economics of Enterprise AI

The headline API pricing tells you almost nothing about what enterprise AI actually costs. Inference is typically the smallest line item. Total cost of ownership for enterprise AI includes integration, governance, data preparation, workforce transformation, and ongoing operational management. Understanding and managing these economics is increasingly a source of competitive advantage for enterprises that deploy AI at scale.

The Economics of Enterprise AI: Inference Costs & Optimisation Strategies

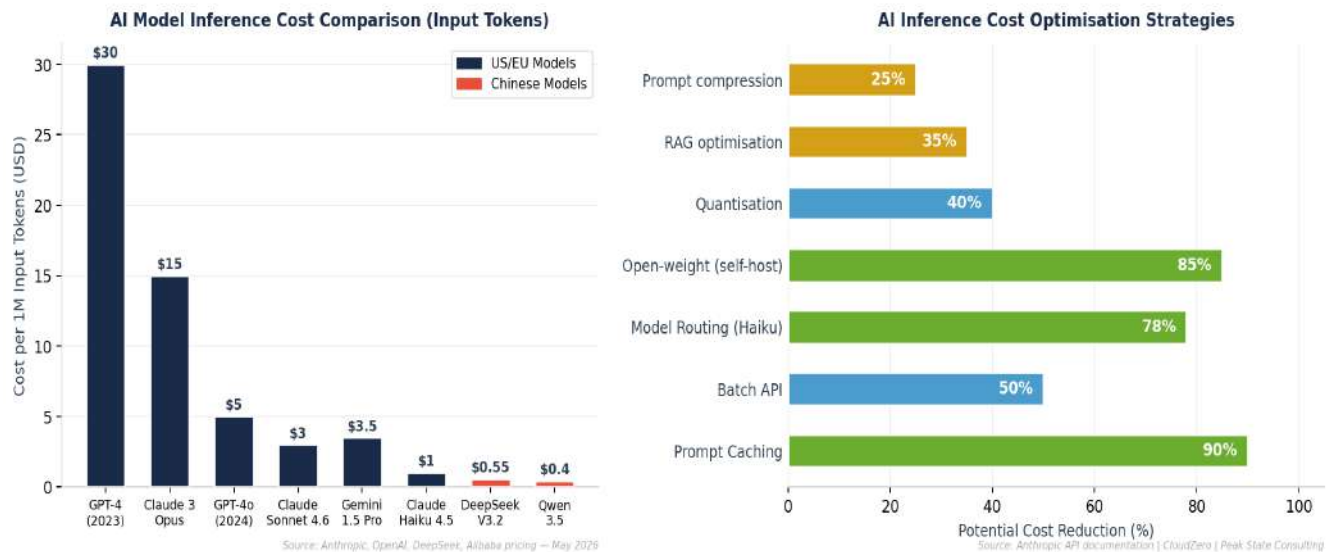


Figure 14: AI Model Inference Cost Comparison and Optimisation Strategies | Source: Anthropic, OpenAI, DeepSeek, Alibaba pricing (June 2026); CloudZero; Peak State Consulting

## Enterprise AI Total Cost of Ownership

Data preparation typically consumes 40–60% of enterprise AI implementation effort, a finding from IBM research. This cost is not visible in API pricing but is the dominant cost driver in most enterprise AI programmes. Organisations with mature governance frameworks achieve faster time-to-value because they have invested in the data foundations that allow AI to perform reliably from the start of deployment rather than during it.

## Key Cost Drivers Beyond API Pricing

- GPU and accelerator compute: For enterprises running self-hosted models or fine-tuning, infrastructure costs can be substantial. NVIDIA H100 instances typically cost \$2–4 per GPU-hour at major cloud providers.
- Data storage and retrieval: Vector databases, embedding generation, and document storage for RAG pipelines add meaningful infrastructure cost at scale. A knowledge base of 10 million documents may require \$5,000–\$15,000/month in vector database costs.
- MLOps and observability tooling: Monitoring, logging, model versioning, and automated retraining infrastructure typically adds \$50,000–\$200,000/year per production AI system in enterprise environments.
- AI security infrastructure: AI-specific security monitoring, prompt security testing, and audit logging adds both tooling cost and engineering capacity requirements.
- Enterprise integration layers: API management, authentication, data pipeline engineering, and system integration typically represents 20–30% of total implementation cost for enterprise AI projects.

## AI FinOps: Managing AI Costs at Enterprise Scale

---

Leading enterprises are implementing AI FinOps capabilities, operational disciplines for managing AI costs with the same rigour applied to cloud infrastructure costs. This includes usage visibility dashboards providing real-time visibility into token consumption by team, use case, and model tier; AI workload optimisation that automatically routes queries to the most cost-effective model capable of handling the task; and cost-performance benchmarking that validates whether the inference cost of each use case is justified by the business value it delivers.

## Inference Cost Optimisation Strategies

- Prompt caching: Reduces input token costs by up to 90% for repeated context, the single highest-impact cost optimisation available for most enterprise use cases. Implementation requires architectural changes to separate static and dynamic prompt components.
- Batch API: Reduces both input and output costs by 50% for asynchronous workloads. Appropriate for document processing queues, overnight analytics runs, and any use case that does not require real-time response.
- Model routing: Using lightweight routing logic to direct simple queries to Haiku 4.5 or equivalent low-cost models and reserving Sonnet/Opus for tasks that genuinely require higher capability. Can reduce inference costs by 60–80% in mixed-complexity workloads.
- Self-hosted open-weight models: For non-sensitive, high-volume workloads, self-hosting open-weight models (Qwen 2.5, Llama 3.3, Mistral) on enterprise infrastructure can reduce inference costs by 70–90% versus managed API services.

- Quantisation and model compression: Quantised models (INT8, INT4) offer 40–60% cost reduction with modest quality trade-offs acceptable for many enterprise use cases.
- Prompt engineering efficiency: Well-designed prompts that are precise and concise can reduce token consumption by 20–40% versus verbose, exploratory prompts, without any reduction in output quality.

The emergence of lower-cost Chinese AI providers, with DeepSeek V3.2 at \$0.55/million input tokens against Claude Sonnet 4.6 at \$3/million, is fundamentally changing the make-vs-buy calculus for enterprise AI. The price differential means that for cost-sensitive, non-regulated workloads, Chinese open-weight models are increasingly competitive on pure economics. The enterprise AI cost optimisation strategy of 2026–2028 will involve sophisticated multi-model architectures that optimise cost, performance, and compliance across workload types.

**ENTERPRISE  
CFO NOTE**

AI infrastructure costs can scale unexpectedly as adoption grows. Enterprises that did not implement cost governance frameworks early in their AI programmes have faced infrastructure bills 3–5x above initial estimates as usage expanded. Implement AI FinOps capabilities, usage dashboards, budget alerts, and model routing logic, before you scale, not after.

# ENTERPRISE AI IMPLEMENTATION GUIDE

## 10 Phases from Executive Alignment to Continuous Optimisation

A comprehensive step-by-step guide drawing on 80+ Peak State Consulting enterprise programme deliveries

**10**

Implementation  
phases

**Phase 1**

Executive  
alignment

**Phase 6**

Pilot  
deployment

**Phase 10**

Optimise  
& measure

# Enterprise AI Implementation Guide

Most enterprise AI programmes that fail do not fail because the technology did not work. They fail because the organisation was not ready for it. This section is a practical guide to getting that right.

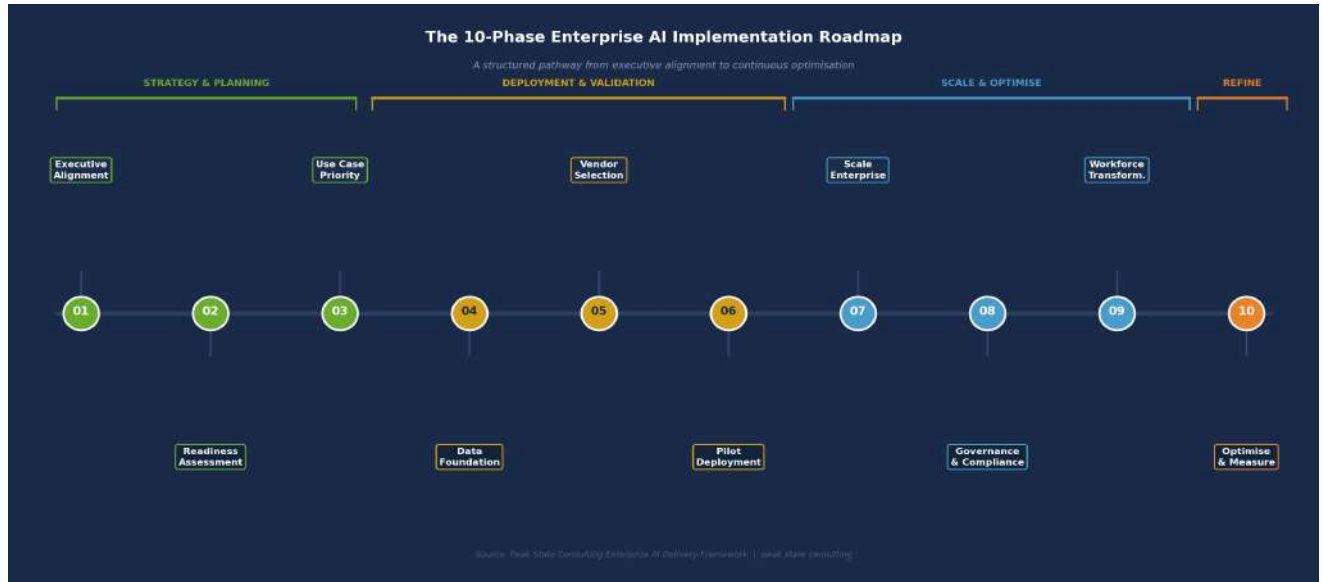


Figure 15: The 10-Phase Enterprise AI Implementation Roadmap | Source: Peak State Consulting Enterprise AI Delivery Framework



Figure 16: Enterprise AI Programme Reference Architecture | Source: Peak State Consulting

## Phase 1, Executive Alignment & Strategy

Successful enterprise AI programmes begin with strategic alignment at board and executive level. Organisations must establish a clear AI vision tied directly to measurable business outcomes, operational efficiency, customer experience improvement, risk reduction, or revenue growth.

Leading enterprises create a cross-functional AI steering committee with representation from the CEO, CIO, CTO, legal, risk, HR, operations, and business-unit leaders.

- Define enterprise-wide AI objectives tied to strategic business outcomes with measurable KPIs.
- Identify priority business challenges where AI can deliver measurable value within 12 months.
- Create an AI Governance Charter establishing principles, accountability structures, and escalation processes.
- Establish executive sponsorship and budget ownership, without named executive ownership, AI programmes consistently underperform.
- Define measurable KPIs and ROI expectations before deployment begins, not after.

## Phase 2, Enterprise AI Readiness Assessment

---

Before deploying AI systems at scale, enterprises must assess organisational readiness across data quality, cloud maturity, security controls, workforce capability, and change management readiness. Best practice organisations allocate substantial effort, typically 4–8 weeks, to readiness assessment and governance framework design before scaling pilots. Peak State Consulting's AI Readiness Framework provides a structured methodology for this assessment.

- Data governance and accessibility assessment, identify gaps between current data quality and AI-readiness requirements.
- Cloud and API architecture review, confirm the integration infrastructure required for AI deployment.
- Cybersecurity and compliance controls assessment, confirm AI-specific security requirements are met.
- Talent and AI capability gap analysis, identify the skills that must be hired, trained, or engaged through partners.
- Organisational culture and change readiness evaluation, assess the change management investment required.

## Phases 3–5, Use Case Prioritisation, Data Foundation & Vendor Selection

---

Enterprises should prioritise a small number of high-value use cases with measurable business impact rather than attempting too many AI initiatives simultaneously. Recommended prioritisation criteria include strategic importance, data availability, implementation complexity, regulatory risk, expected financial return, and time-to-value. Once use cases are prioritised, data foundation investment must be secured, many enterprises underestimate the time required to prepare enterprise data for AI deployment, with data preparation typically consuming 40–60% of total implementation effort.

AI platform selection decisions should be based on long-term scalability rather than short-term demonstrations. Vendor evaluation must assess platform security, compliance certifications, ecosystem maturity, pricing transparency, context window capability, MLOps tooling, and commercial sustainability. For regulated enterprises, EU AI Act readiness and data residency options are mandatory evaluation criteria. See The Vendor Landscape section for the Peak State Vendor Evaluation Framework.

## **Phases 6–8, Pilot Deployment, Scaling & Governance**

---

AI pilots should be tightly scoped with measurable success metrics established before deployment, not after. Critical pilot metrics include accuracy and reliability, user adoption rates, productivity improvement, cost reduction, customer satisfaction impact, and operational risk reduction. Pilot programmes must include human oversight mechanisms and rollback procedures from day one.

The transition from pilot to enterprise-scale deployment is where many organisations fail. Scaling requires operational governance, MLOps capability, workforce training, and executive alignment. Governance must be embedded into enterprise operating models, regulatory frameworks including the EU AI Act, GDPR, and sector-specific compliance obligations are forcing enterprises to adopt explainability, transparency, and accountability controls.

## **Phases 9–10, Workforce Transformation & Continuous Optimisation**

---

AI transformation is fundamentally a people transformation challenge. Enterprises that invest in employee enablement, communication, and upskilling significantly outperform organisations focused only on technology deployment. Priority workforce investments include AI literacy training, executive education, AI-assisted workflow redesign, internal AI champions, and cross-functional collaboration programmes.

AI transformation is not a one-time project. Organisations must establish continuous measurement frameworks covering financial impact, operational performance, customer outcomes, and governance maturity. Leading organisations treat AI as an evolving enterprise capability rather than a standalone technology initiative, continuously optimising performance, cost, and governance as the technology and regulatory landscape evolves.

# AI GOVERNANCE TEMPLATES & CHECKLISTS

## Five Practical Frameworks for Enterprise AI Risk Management

Ready-to-adapt governance templates aligned with EU AI Act, GDPR, and sector-specific compliance requirements

**5**

Governance  
templates

**10**

Items per  
checklist

**EU AI Act**

Full application  
by Aug 2027

**GDPR**

Data privacy  
by design

- Template 1: AI Governance Charter – board-level principles and accountability
- Template 2: AI Risk Assessment – pre-deployment compliance checklist
- Template 3: AI Deployment Checklist – go-live governance gates
- Template 4: AI Vendor Evaluation – due diligence checklist
- Template 5: AI Monitoring – continuous operations checklist

# AI Governance Templates & Implementation Checklists

Effective AI governance requires structured processes and documented frameworks. The following templates and checklists are designed to be adapted by enterprise teams for their specific regulatory environment, risk appetite, and organisational context. They reflect Peak State Consulting's delivery experience and are aligned with requirements under the EU AI Act, GDPR, and major sector-specific regulatory frameworks.

## 1. AI Governance Charter Template

The AI Governance Charter is the foundational document establishing your organisation's principles, accountability structures, and decision-making processes for AI. It should be board-approved and reviewed annually.

### AI GOVERNANCE CHARTER, KEY COMPONENTS

- ✓ Define enterprise AI principles, transparency, fairness, accountability, safety, and privacy by design.
- ✓ Establish accountability structures, named owners for AI strategy, risk, compliance, and operations.
- ✓ Define AI risk ownership, identify which business functions own risk for each category of AI use.
- ✓ Clarify escalation and review processes, how AI incidents are reported and resolved.
- ✓ Set risk tolerance thresholds, define acceptable levels of AI error rates, bias, and uncertainty.
- ✓ Establish AI ethics committee composition, meeting cadence, and decision authority.
- ✓ Define AI classification framework, distinguish high-risk, medium-risk, and low-risk AI applications.
- ✓ Set policy review cycle, commit to annual governance charter review and update process.

## 2. AI Risk Assessment Checklist

Complete this checklist for every AI system before production deployment. High-risk AI systems (as defined by the EU AI Act) require formal conformity assessment documentation.

### AI RISK ASSESSMENT, PRE-DEPLOYMENT CHECKLIST

- ✓ Data privacy assessment completed and documented, confirm GDPR / sector-specific compliance.
- ✓ Regulatory obligations reviewed, identify applicable sector regulations and AI Act classification.
- ✓ Model explainability requirements assessed, can the system explain its outputs to affected individuals?
- ✓ Human oversight requirements defined, who reviews AI outputs before consequential decisions?
- ✓ Cybersecurity testing completed, prompt injection, adversarial input, and access control testing.

- ✓ Prompt security controls validated, system prompt confidentiality and injection resistance confirmed.
- ✓ Bias and fairness assessment completed, test for demographic and protected characteristic bias.
- ✓ Data lineage documented, confirm data provenance and consent for all training and inference data.
- ✓ Incident response process established, who is notified, what is the rollback procedure?
- ✓ EU AI Act conformity assessment completed (for high-risk AI systems as defined in Annex III).

### 3. AI Deployment Checklist

Complete this checklist before go-live for each AI system entering production. This checklist should be retained as evidence of governance compliance.

#### AI DEPLOYMENT CHECKLIST, GO-LIVE GATES

- ✓ Business owner assigned with named accountability for AI system performance and outcomes.
- ✓ Governance approval obtained from AI ethics committee or equivalent body.
- ✓ Data lineage documented for all data sources used in training and inference.
- ✓ MLOps monitoring enabled, performance metrics, drift detection, and alerting configured.
- ✓ Incident response process established and communicated to all relevant teams.
- ✓ Rollback procedures tested and documented, can the system be disabled rapidly if required?
- ✓ User training completed, all users have received appropriate training on AI capabilities and limitations.
- ✓ Baseline metrics established, quantitative performance benchmarks recorded before go-live.
- ✓ Audit logging enabled, all AI inputs and outputs logged for the required retention period.
- ✓ EU AI Act post-market monitoring plan in place (for high-risk AI systems).

### 4. AI Vendor Evaluation Checklist

#### AI VENDOR EVALUATION, DUE DILIGENCE CHECKLIST

- ✓ Security certifications verified, SOC 2 Type II, ISO 27001, and sector-specific certifications.
- ✓ Data residency confirmed, confirm where enterprise data is processed and stored.
- ✓ No-training commitment confirmed, vendor contractually commits not to train on enterprise data.
- ✓ SLA commitments reviewed, availability, latency, and incident response commitments documented.
- ✓ Commercial pricing validated, full cost modelling including API consumption at anticipated scale.

- ✓ Long-term roadmap assessed, vendor capability roadmap aligns with anticipated future requirements.
- ✓ Independent testing completed, accuracy and reliability independently validated, not vendor-tested.
- ✓ Exit strategy documented, data portability and migration path confirmed before commitment.
- ✓ EU AI Act alignment confirmed, vendor provides conformity documentation for high-risk use cases.
- ✓ Financial stability assessed, vendor funding, revenue trajectory, and longevity confirmed.

## 5. Ongoing AI Monitoring Checklist

---

### AI MONITORING, CONTINUOUS OPERATIONS CHECKLIST

- ✓ Model performance monitoring active, accuracy, confidence scores, and error rates tracked.
- ✓ Performance drift tracking enabled, automated alerts when model performance degrades.
- ✓ Bias monitoring operational, ongoing testing for demographic and protected characteristic bias.
- ✓ Audit logs retained, all AI interactions logged and retained per regulatory requirements.
- ✓ Security monitoring operational, prompt injection detection, anomaly detection on agent actions.
- ✓ Compliance reporting automated, regulatory reporting and evidence collection automated where possible.
- ✓ Quarterly model performance review scheduled, formal review of AI system performance and quality.
- ✓ Annual governance review scheduled, governance charter, risk assessments, and policies reviewed.
- ✓ Incident log maintained, all AI incidents, near-misses, and complaints documented and reviewed.
- ✓ User feedback mechanism in place, systematic collection of user feedback on AI output quality.

## Executive Overview, Claude Case Study

---

Anthropic's Claude has become one of the most widely adopted enterprise AI platforms globally. As of May 2026, Anthropic serves more than 300,000 business customers, with 100,000+ running Claude on Amazon Bedrock alone. Annualised revenue reached \$47 billion as of May 28, 2026 (CNBC), up from \$30 billion in April. In April 2026, Anthropic's share of the enterprise AI market reached 34.4%, surpassing OpenAI's 32.3% for the first time. More than 1,000 customers now spend over \$1 million annually with Anthropic, double the number just two months earlier, and eight of the Fortune 10 are Claude customers. Amazon committed a further \$5 billion investment in April 2026, bringing its total Anthropic investment to \$13 billion.

The Claude Partner Network, launched in March 2026 with a \$100 million initial investment, has established a global ecosystem of consulting firms, system integrators and technology partners. On June 9, 2026, Anthropic introduced a new frontier tier above Opus: Claude Fable 5, the first publicly available Mythos-class model. Fable 5 is state-of-the-art on nearly all tested capability benchmarks and scores more than 10% higher than Opus 4.8 on key coding evaluations. It is available on the Anthropic API, Amazon Bedrock, Google Vertex AI, Microsoft Foundry, and GitHub Copilot. Claude Mythos 5, the unrestricted version, remains restricted to trusted partners through Project Glasswing due to advanced cybersecurity capabilities.

However, Claude Fable 5's public availability proved short-lived. On June 12, 2026, just three days after launch, the United States Commerce Department issued an export control directive ordering Anthropic to suspend access to both Claude Fable 5 and Claude Mythos 5. Citing national security authorities, the directive prohibited access by any foreign national, whether inside or outside the United States, including Anthropic's own foreign national employees. Unable to filter users by nationality in real time, Anthropic complied within hours by disabling both models for all customers worldwide, making this the first government-mandated withdrawal of a publicly deployed frontier AI model. All other Claude models, including Opus 4.8, Sonnet 4.6 and Haiku 4.5, remained fully available throughout. Anthropic stated it was working to restore access to Fable 5 as quickly as possible and committed to issuing refunds to customers who had paid for access to a model withdrawn so abruptly.

The suspension carries significant implications for enterprise AI planning. The government's stated rationale referenced a purported jailbreak technique; Anthropic publicly disputed the severity of the vulnerability, characterising it as narrow and already replicable by existing public models. Regardless, the episode illustrates a risk that enterprise technology leaders must now factor into their AI procurement and continuity frameworks: even a fully commercial, generally available frontier model can be removed from service at short notice by regulatory intervention, with no guaranteed timeline for restoration. Organisations that had begun integrating Fable 5 into production workflows face an unplanned migration back to Opus 4.8, underscoring the importance of vendor-agnostic architecture design, robust fallback provisions and governance frameworks that account for model availability risk as a distinct category alongside performance and cost.

# THE CLAUDE MODEL FAMILY

Opus 4.7 · Sonnet 4.6 · Haiku 4.5

Three-tier model family for every enterprise workload

This section covers:

**200K**

Context window  
(all models)

**90%**

Cost reduction  
with caching

**50%**

Savings via  
Batch API

**500K**

Enterprise plan  
context window

- Three models optimised for cost, speed and reasoning depth
- All models share a 200,000-token context window
- Enterprise plan extends context to 500,000 tokens
- Prompt caching reduces input costs by up to 90%
- Source: Anthropic API documentation, June 2026

### The Claude Model Family

MODEL	POSITIONING	API COST (PER 1M TOKENS)	ENTERPRISE USE CASES
<b>Claude Haiku 4.5</b>	High-volume, cost-optimised	\$1 input / \$5 output	Customer-facing chatbots, real-time query classification, data extraction at scale
<b>Claude Sonnet 4.6</b>	Balanced intelligence and speed	\$3 input / \$15 output	Contract review, regulatory analysis, knowledge management, code review, internal assistants
<b>Claude Opus 4.7</b>	Highest reasoning capability (deprecated June 2026)	\$5 input / \$25 output	Complex legal and financial analysis, strategic research, agentic multi-step workflows
<b>Claude Opus 4.8</b>	Flagship Opus; reliability + speed	\$5 input / \$25 output	Production agentic workflows; browser agents (84% Online-Mind2Web); 1M context; Fast mode at 2.5x speed. 4x fewer undetected code errors than Opus 4.7. Supports zero data retention (ZDR).
<b>Claude Fable 5</b> (Mythos-class, June 2026)	New frontier tier above Opus. State-of-the-art on nearly all benchmarks. 1M context, 128K output, always-on adaptive thinking.	\$10 input / \$50 output. Included on Pro/Max/Team/Enterprise until June 22, 2026; usage credits thereafter.	Large-scale codebase migrations, long-horizon autonomous research, complex multi-stage document processing. Safeguards block offensive cybersecurity and biology requests; falls back to Opus 4.8. Available on API, Amazon Bedrock, Google Vertex AI, Microsoft Foundry, GitHub Copilot.

Prompt caching reduces input token costs by up to 90% for repeated context. The Batch API reduces both input and output costs by 50% for asynchronous workloads. Source: Anthropic API documentation, CloudZero Claude Pricing Guide June 2026.

On 13 June 2026, Anthropic suspended its powerful new AI model after US authorities raised security concerns just days after its public release. In a statement published on its website, external, Anthropic said it was ordered to suspend foreign nationals from using Claude Fable 5, a program that the company self-described as "too powerful".

"The net effect of this order is that we must abruptly disable Fable 5 and Mythos 5 for all our customers to ensure compliance," the company wrote.

# ENTERPRISE SUBSCRIPTION TIERS

Pro · Max · Team · Enterprise

Complete pricing guide from individual developers to enterprise

This section covers:

**\$20**

Pro plan  
per month

**\$100**

Max 5x plan  
per month

**\$25**

Team Standard  
per seat/month

**Custom**

Enterprise  
pricing

- Pro at \$20/month through to Enterprise with custom SLA
- Team Premium adds Claude Code for engineering teams
- Enterprise includes SSO, SCIM, RBAC and full audit logs
- HIPAA-ready infrastructure for regulated industries
- Source: Anthropic.com, mem0.ai Pricing Guide, June 2026

### Enterprise Subscription Tiers

PLAN	PRICE	WHO IT IS FOR	KEY FEATURES
<b>Pro</b>	\$20/month	Individual power users	Claude Code, file creation, code execution, unlimited projects, Google Workspace integration, remote MCP connectors
<b>Max 5x</b>	\$100/month	Heavy individual users	5x Pro usage limits, priority access to new features and models
<b>Max 20x</b>	\$200/month	Intensive developers	20x Pro usage; rate limits stop being a practical concern for full-day development work
<b>Team Standard</b>	\$25/seat/month (\$20 annual)	5+ person teams	All Pro features, admin controls, centralised billing. Min 5 seats
<b>Team Premium</b>	\$125/seat/month (\$100 annual)	Engineering teams	All Team Standard features + Claude Code. Mix and match Standard + Premium seats
<b>Enterprise</b>	Custom (~\$20/seat + API usage)	Large / regulated organisations	500K context window, HIPAA readiness, SSO/SAML, SCIM, RBAC, audit logs, dedicated support

**ENTERPRISE PRICING NOTE** Enterprise seat pricing starts at approximately \$20/seat/month, with API usage billed separately at standard token rates. For high-volume deployments, Anthropic offers reserved capacity contracts. Source: Anthropic.com, mem0.ai Pricing Guide June 2026.

# VERIFIED ENTERPRISE DEPLOYMENTS

Deloitte · TELUS · GitLab · Salesforce · Snowflake  
Five named, publicly verified production deployments

This section covers:

**470K**

Deloitte  
professionals

**57K**

TELUS  
employees

**90%**

Salesforce  
Cursor adoption

**\$200M**

Snowflake  
partnership

- All facts drawn from official Anthropic press releases
- Deloitte: 470,000 professionals across 150 countries
- TELUS: 57,000 employees on Claude-powered Fuel iX
- Salesforce: 90% Cursor adoption across 20,000 engineers
- Source: [anthropic.com/news](https://anthropic.com/news), [claude.com/customers](https://claude.com/customers)

## **Verified Enterprise Deployments**

The following deployments are based exclusively on publicly announced, verified partnerships and named customer stories from Anthropic's official press releases and the [claude.com/customers](https://claude.com/customers) library.

### **Deloitte, Largest Enterprise AI Deployment to Date**

In October 2025, Anthropic and Deloitte announced an expanded alliance representing Anthropic's largest enterprise AI deployment to date. Claude is being made available to more than 470,000 Deloitte professionals across 150 countries. A certification programme will train and certify 15,000 Deloitte professionals on Claude. Source: Anthropic official press release, October 6, 2025.

### **TELUS, 57,000 Employees on Claude-Powered AI Platform**

TELUS (106,400 employees, CA\$20 billion revenue) adopted Claude as the core engine for its internal Fuel iX platform, giving 57,000 employees access to advanced AI workflows. Pilot programmes reported 30% faster pull request turnaround times. Source: DataStudios.org, September 2, 2025.

### **GitLab, Internal and External Claude Deployment**

GitLab adopted Claude both via the Anthropic API for its products and through the Claude Enterprise plan for internal processes. "By using the same Claude models internally that we offer externally, we're better able to build AI solutions that we trust and use ourselves.", Taylor McCaslin, GitLab. Source: [claude.com/customers/gitlab-enterprise](https://claude.com/customers/gitlab-enterprise).

### **Salesforce, Claude Models in Einstein and Agentforce**

In September 2024, Salesforce and Anthropic announced a partnership making Claude models available to Salesforce customers via Amazon Bedrock. Salesforce achieved 90% adoption of Cursor (which uses Claude models) across its 20,000-engineer workforce. Source: Anthropic official press release, September 2024.

### **Snowflake, \$200 Million Multi-Year Partnership**

Anthropic and Snowflake expanded their partnership through a \$200 million multi-year agreement to bring Claude models directly into Snowflake's data platform, reaching more than 12,600 Snowflake enterprise customers. Source: Sacra Research, June 2026.

# MODEL CONTEXT PROTOCOL (MCP)

AWS Bedrock · Google Vertex · Snowflake · Salesforce · Slack  
The open standard connecting Claude to enterprise systems

This section covers:

**97M**

Downloads  
March 2026

**5,800+**

Enterprise  
MCP servers

**2024**

Launched by  
Anthropic

**Every**

Major AI provider  
now supports MCP

- Open standard introduced by Anthropic, November 2024
- 97 million downloads in March 2026 alone
- 5,800+ community and enterprise MCP servers available
- Every major AI provider — OpenAI, Google, Microsoft — supports MCP
- Source: Digital Applied Research, April 2026

### Model Context Protocol (MCP)

The Model Context Protocol (MCP) is an open standard introduced by Anthropic in November 2024. In December 2025, Anthropic donated MCP to the Agentic AI Foundation under the Linux Foundation, co-founded with Block and OpenAI, establishing it as a vendor-neutral industry standard. As of March 2026, every major AI provider supports MCP, with 97 million monthly SDK downloads (Digital Applied Research, April 2026). There are now 5,800+ community and enterprise MCP servers.

MCP INTEGRATION	WHAT IT ENABLES	ENTERPRISE EXAMPLE
<b>AWS Bedrock</b>	Claude with VPC isolation, IAM controls and CloudTrail audit logging. Full data residency within chosen AWS region.	Finance teams running Claude for credit analysis with zero data leaving the enterprise perimeter.
<b>Google Vertex AI</b>	Claude via Google Cloud infrastructure, integrated with BigQuery and GCS document storage.	Insurance firm using Claude to analyse policy documents stored in GCS.
<b>Snowflake Cortex</b>	Execute Claude directly against Snowflake data warehouse tables without data egress.	Retail organisation using Claude to analyse customer transaction data at scale.
<b>Salesforce Einstein</b>	Claude models within the Salesforce Platform via Amazon Bedrock for Einstein AI features.	B2B software company using Claude-powered Einstein to automate case summarisation.
<b>Internal APIs &amp; Databases</b>	MCP servers connect Claude to proprietary APIs, databases, file systems and vector stores.	Law firm connecting Claude to its document management system for natural-language search.
<b>Slack &amp; Teams</b>	Claude integrates natively with Slack (2025) and Microsoft Teams via MCP.	Consulting firm enabling staff to invoke Claude knowledge base queries directly from Slack.

# ENTERPRISE DEPLOYMENT ARCHITECTURE

Five-Layer Reference Architecture for Enterprise Claude  
From model access and data sovereignty to agentic AI governance

This section covers:

**Layer 1**  
Model Access &  
Data Sovereignty

**Layer 2**  
RAG & Knowledge  
Management

**Layer 3**  
System Prompts  
& Guardrails

**Layer 5**  
Governance &  
Observability

- Layer 1: Model access with VPC isolation and data residency
- Layer 2: Retrieval-Augmented Generation with vector databases
- Layer 3: System prompts, compliance guardrails, refusal logic
- Layer 4: Agentic AI and tool use with Claude Code
- Layer 5: SSO, RBAC, audit logging, HIPAA, EU AI Act tools

## **Enterprise Deployment Architecture**

### **Layer 1, Model Access and Data Sovereignty**

Claude is available through: (1) Direct Anthropic API, no enterprise data used for training; (2) AWS Bedrock, with full VPC isolation, IAM identity controls and CloudTrail audit logging; (3) Google Vertex AI, within Google Cloud's managed ML infrastructure.

### **Layer 2, Retrieval-Augmented Generation (RAG)**

Most enterprise Claude deployments use RAG to give Claude access to proprietary enterprise knowledge without embedding that knowledge into the model itself. Enterprise documents are chunked and vectorised, stored in a vector database (Amazon OpenSearch, Pinecone, Weaviate, or Snowflake Cortex), and retrieved at query time.

### **Layer 3, System Prompts and Guardrails**

The system prompt is the primary mechanism for configuring Claude's behaviour in enterprise deployments. It defines Claude's role, access rights, output format, compliance guardrails, refusal behaviours and escalation paths. Treat system prompts as code: version-controlled, peer-reviewed, and staged through development, test and production environments.

### **Layer 4, Agentic AI and Tool Use**

Claude Code has 29 million cumulative installs on VS Code and represents approximately 4% of all public GitHub commits (Sacra Research, 2026). Claude's Constitutional AI architecture makes it suitable for agentic deployments by design.

### **Layer 5, Governance, Observability and Cost Management**

Claude Enterprise provides: SSO via SAML/SCIM; role-based access controls (RBAC); full audit logging; HIPAA-ready infrastructure; and EU AI Act compliance tooling. Claude Security (launched April 2026) provides AI-powered vulnerability scanning integrated with CrowdStrike, Palo Alto Networks, SentinelOne and Wiz.

# ENTERPRISE USE CASE PATTERNS

## Five Proven Deployment Patterns

Battle-tested configurations for the most common enterprise use cases

This section covers:

**A**

Knowledge  
Management

**B**

Document  
Analysis

**C**

Software  
Engineering

**D/E**

Customer  
Service / RegIntel

- Pattern A: Knowledge management and internal search (Sonnet 4.6)
- Pattern B: Document analysis and data extraction (Sonnet / Opus)
- Pattern C: Software engineering with Claude Code (Team Premium)
- Pattern D: Customer-facing assistants and service agents (Haiku)
- Pattern E: Regulatory intelligence and compliance monitoring

## **Enterprise Use Case Patterns**

### **Pattern A, Knowledge Management**

Deploy Claude with a RAG pipeline connected to internal knowledge base. Best model: Claude Sonnet 4.6.

### **Pattern B, Document Analysis**

Deploy Claude to read, analyse and extract structured information from high-volume document workflows. Claude's long context window (200K tokens standard, up to 1M on recent models) allows entire documents to be processed in a single API call. Best model: Sonnet 4.6 for standard; Opus 4.8 for complex legal or financial.

### **Pattern C, Software Engineering with Claude Code**

Deploy Claude Code to engineering teams via Team Premium (\$125/seat/month, \$100 annual) or Enterprise plans. Core use cases: multi-file refactoring, automated code review, test generation, legacy code modernisation.

### **Pattern D, Customer-Facing Assistants**

Deploy Claude Haiku 4.5 for high-volume customer-facing applications. Integrate via MCP with CRM systems to give Claude real-time access to customer data and support ticket context.

### **Pattern E, Regulatory Intelligence**

Deploy Claude to monitor regulatory publications and generate structured impact summaries. Connect Claude via MCP to regulatory publication feeds (FCA, EBA, SEC, FINRA, FDA).

# COST MANAGEMENT AT ENTERPRISE SCALE

Prompt Caching · Batch API · Model Routing · Benchmarks  
Practical guide to controlling AI inference costs at scale

This section covers:

**90%**

Prompt caching  
cost reduction

**50%**

Batch API  
cost reduction

**80%**

Model routing  
max savings

**\$30**

Max daily cost  
per developer

- Prompt caching: up to 90% input cost reduction on repeated context
- Batch API: 50% cost reduction for async workloads
- Model routing: 60–80% savings routing to Haiku for simple tasks
- 90% of Claude Code users stay below \$30 per developer day
- Source: Anthropic, CloudZero Claude Pricing Guide, June 2026

## **Cost Management at Enterprise Scale**

### **Prompt Caching**

Reduces input token costs by up to 90%. Implement by separating static and dynamic prompt components.

### **Batch API**

Reduces both input and output costs by 50% for asynchronous workloads including bulk document processing and overnight analytics runs.

### **Model Selection and Routing**

A disciplined model selection policy routing simple queries to Haiku 4.5 and reserving Sonnet/Opus for tasks requiring higher capability can reduce inference costs by 60–80%.

### **Real-World Cost Benchmarks**

Anthropic's own enterprise data shows that 90% of Claude Code users stay below \$30 per active developer day. Processing a 50-page contract costs approximately \$0.11 with Sonnet 4.6. With prompt caching, costs fall to approximately \$40–50 per 1,000 contracts per month. Source: CloudZero Claude Pricing Guide, June 2026.

# IMPLEMENTATION ROADMAP

Five Phases: Procurement ✕ Pilot ✕ Scale ✕ Agents

A structured pathway from readiness assessment to production agentic AI

This section covers:

**1-2**

Months:  
Readiness

**2-5**

Months:  
Pilot

**6-12**

Months:  
Scaled Rollout

**10+**

Months:  
Agentic AI

- Phase 1: AI Readiness Assessment and procurement (Months 1-2)
- Phase 2: Pilot deployment to 20-50 users (Months 2-5)
- Phase 3: ROI measurement and validation (Months 4-7)
- Phase 4: Scaled enterprise rollout (Months 6-12)
- Phase 5: Agentic AI expansion with Claude Code (Months 10-15+)

## Implementation Roadmap

PHASE	DURATION	KEY ACTIVITIES	EXIT CRITERIA
<b>1. Readiness</b>	Months 1–2	AI Readiness Assessment; use case prioritisation; governance framework design; SSO/SAML configuration.	Contract executed; 2–3 priority use cases defined with pre-deployment baseline metrics.
<b>2. Pilot</b>	Months 2–5	Build initial RAG pipeline or Claude Code integration; develop system prompt library; deploy to 20–50 users.	User satisfaction >3.5/5; first measurable business impact documented.
<b>3. Measurement</b>	Months 4–7	Blind accuracy testing vs manual benchmarks; ROI calculation against baselines; compliance audit.	Documented ROI evidence; board presentation ready; compliance review passed.
<b>4. Scaled Rollout</b>	Months 6–12	Expand to full target user population; add additional use cases; configure MCP integrations.	Target user population onboarded; >4 MCP integrations active.
<b>5. Agentic Expansion</b>	Months 10–15+	Deploy Claude Code to engineering teams; evaluate agentic workflows; implement multi-agent architectures.	At least one autonomous agentic workflow in production.

<b>WHERE TO START</b>	<p>The Claude Partner Network, launched in March 2026 with a \$100 million initial Anthropic investment, provides enterprises with access to certified implementation partners including Accenture, BCG, Deloitte, Infosys and PwC. Contact Anthropic enterprise sales at <a href="https://anthropic.com/contact-sales">anthropic.com/contact-sales</a>, or engage an experienced independent implementation adviser such as Peak State Consulting for implementation support.</p>
-----------------------	--

# ABOUT THE AUTHOR



## Anuj Khanna

Founder & CEO  
Peak State Consulting

MBA · Univ. of Sheffield  
BA Econ · Univ. of Mumbai

## Anuj Khanna

Founder & CEO · Peak State Consulting

Board Advisor · Non-Executive Director · Strategic M&A · AI Expert · Deep Tech · Renewable Energy · PE & Sovereign Funds · Family Office Advisor  
Shipping & Infrastructure · FinTech / ESG

### About

Anuj Khanna is a globally recognised AI industry expert and enterprise technology strategist with over 25 years of international experience. He is the author of this Enterprise AI Research Report, read by C-suite executives and board directors across Europe, North America and Asia. Anuj advises high-growth companies, multinational enterprises, private equity firms, family offices and sovereign investors on AI strategy, enterprise technology transformation, investment appraisal and digital innovation.

### Areas of Expertise

#### AI Strategy & Enterprise Technology

AI readiness assessments, governance frameworks, vendor selection, agentic AI deployment

#### Investment, Finance & M&A

Leveraged finance, VC/PE, ESG, renewable energy investment, due diligence, M&A advisory

#### Business Strategy & Market Expansion

Business development, strategic partnerships, corporate sales, new market entry

#### Board Advisory & Governance

Non-Executive Director roles, leadership development, investor relations, AI governance

#### Product Development & Corporate Expansion

Product design, lifecycle management, outsourced technology delivery, revenue growth

## 25+

Years International Experience

## 80+

Client Projects Delivered

## 3

Continents — Europe · Americas · Asia

- Board Advisor & Non-Executive Director to technology companies
- Advisor to Family Offices, Sovereign Wealth & Private Equity Funds
- AI, Renewable Energy, Shipping, Infrastructure & Deep Tech Expert
- Enterprise AI Strategy · Digital Transformation · Corporate Finance
- Investor in Renewable Energy Projects across Europe and Asia



### Anuj Khanna

CEO, Peak State Consulting

anuj@peakstate.com · peakstate.com · +44 7916 056916

June 2026

Research Report

# PEAK STATE CONSULTING

Independent AI Strategy · Enterprise Technology Advisory · Board Advisory

## Enterprise AI Strategy

AI readiness assessments,  
platform selection and  
governance frameworks

## Board Advisory & NED

C-suite and board-level AI,  
technology and strategy  
advisory roles

## Investment & M&A

Due diligence, VC/PE,  
renewable energy and  
sovereign fund advisory

## Digital Transformation

Technology consulting,  
product development and  
programme delivery

## AI Governance & Compliance

EU AI Act, GDPR and  
sector-specific regulatory  
readiness frameworks

## Business Strategy & Expansion

Market entry, strategic  
partnerships and corporate  
development programmes

Established 2014 · London, United Kingdom · Operating across Europe, North America & Asia

Trusted by high-growth companies, global enterprises,  
private equity funds, family offices and sovereign investors.

**2014**

Established in London

**Europe · Americas · Asia**

Global Project Delivery

**100%**

Independent & Vendor-Neutral

- Enterprise AI Strategy · AI Readiness Assessments · Governance Frameworks
- Board Advisory & Non-Executive Directorship · Leadership Development
- AI, Renewable Energy, Shipping, Infrastructure & Deep Tech Advisory
- Investment Advisory · M&A & Due Diligence · VC / PE · Family Offices
- Digital Transformation · Product Development · New Market Entry
- EU AI Act Compliance · AI Governance · Regulatory Strategy



**Anuj Khanna**

CEO, Peak State Consulting

[anuj@peakstate.com](mailto:anuj@peakstate.com) · [peakstate.com](http://peakstate.com) · +44 7916 056916

**June 2026**

Research Report

## About Peak State Consulting

Peak State Consulting is a London-based management consultancy founded in 2014 by Anuj Khanna, a recognised AI industry expert and enterprise technology strategist. The firm works with high-growth companies and global enterprises across Europe, North America and Asia, spanning ambitious technology startups, multinational corporations, leading Private Equity firms, Family Offices and Venture Capital investors. Over more than a decade and 80-plus client projects, Peak State has built a reputation for independent, rigorous counsel and for staying ahead of the curve on the technologies that matter most to enterprise leadership teams.

Our AI advisory practice is built for the current moment in enterprise AI: a period of rapid capability advances, material implementation risk and intensifying board-level scrutiny. We guide organisations through every stage of the AI lifecycle, from initial readiness assessment and strategic prioritisation, through vendor selection, architecture and governance design, to programme delivery, change management, ROI measurement and long-term optimisation. We operate as a trusted independent adviser, free from any vendor affiliation or technology bias. Our clients benefit from the same depth of analysis that shaped this report, applied directly to their own organisations, competitive contexts and regulatory environments.

We bring particular depth in **key industries**, including enterprise software, financial services, professional services, renewable energy, mission critical infrastructure, telecommunications, deep tech and Shipping. Our clients range from fast-scaling technology companies needing commercial AI strategy to large enterprises managing complex, multi-country AI governance and compliance obligations. Whatever your starting point, whether you are evaluating AI for the first time or scaling an existing programme, we bring the strategic clarity, technical depth and implementation experience to move you forward with confidence.

In addition to its enterprise advisory work, Peak State Consulting serves as a trusted strategy consultant and board adviser to a select group of investors, including venture capital firms, family offices and private equity funds. In this capacity, the firm provides rigorous, independent analysis of AI market dynamics, emerging technology investment theses and portfolio company AI strategy, enabling investment decision-makers to assess opportunities and risks with the same analytical depth applied to Peak State's enterprise engagements. Peak State's founder brings direct board-level experience alongside its advisory mandates, offering investors both strategic counsel and a practitioner's perspective on the enterprise AI landscape.

### Areas of Expertise

---

- Enterprise AI strategy formulation, roadmap design and board-level advisory
- AI readiness assessment using the Peak State proprietary six-pillar framework, with prioritised investment recommendations
- Vendor evaluation and independent AI platform selection, including model, infrastructure and tooling decisions
- AI governance, risk management and regulatory compliance, including EU AI Act readiness and responsible AI frameworks

- AI programme delivery oversight, implementation assurance and quality management across multi-vendor environments
- Digital Transformation, Data strategy, AI-ready data architecture and data governance design
- AI talent strategy, operating model design, organisational change management and workforce capability building
- Post-deployment optimisation, ROI measurement, AI value scorecards and board reporting on AI investment performance

## Why Clients Choose Peak State

---

- **Independent and vendor-neutral.** We hold no financial affiliation with any AI platform, cloud provider or software vendor. Our recommendations are driven entirely by what is right for your organisation.
- **Deep technical and commercial expertise.** Our team combines senior management experience at high-growth technology companies with hands-on delivery of complex AI and digital transformation programmes. We speak the language of both the boardroom and the engineering team.
- **Focused on outcomes, not outputs.** We measure our success by the business results our clients achieve, not by the volume of slide decks produced. Every engagement is scoped around measurable outcomes and designed for practical implementation.
- **A track record built on long-term relationships.** We believe the best indicator of consulting quality is whether clients come back. Our portfolio includes engagements with clients who have worked with us repeatedly over many years, and we actively support ongoing introductions through our client network.
- **International reach from a London base.** With successfully delivered projects across Europe, North America and Asia, we are equipped to support global programmes while remaining accessible and responsive to clients based in the UK and across the continent.

## Contact Us

---

### Anuj Khanna

Founder & CEO, Peak State Consulting

Email [anuj@peakstate.com](mailto:anuj@peakstate.com)

Phone [+44 7916 056916](tel:+447916056916)

Web [peakstate.com](http://peakstate.com)

### Engage With Us

*Request an AI Readiness Diagnostic, a focused, no-obligation executive discussion tailored to your organisation's AI agenda. We will assess where you stand, identify your highest-priority opportunities and risks, and outline a practical path forward. Most executives leave with three to five actionable insights they can act on immediately.*

London, United Kingdom

[peakstate.com](http://peakstate.com)



© 2026 Peak State Ltd. · All rights reserved.

**Disclaimer:** This report has been prepared by Peak State Consulting Limited for general informational and educational purposes only. It is intended to provide a broad overview of market trends, technology developments and strategic considerations in the field of enterprise artificial intelligence as understood at the time of publication, June 2026. Nothing in this report should be construed as a definitive statement of fact, a guarantee of future performance, or a comprehensive treatment of any topic addressed herein. The information, opinions, forecasts, estimates, projections and strategic commentary contained in this publication are based on publicly available information, third-party research sources, industry analyst reports, publicly announced company data, and internal assessments believed to be reliable at the time of writing. However, no representation or warranty of any kind, express or implied, is made as to the accuracy, completeness, currentness, reliability, suitability or timeliness of the information contained herein, and Peak State Consulting Limited expressly disclaims any such representation or warranty.

**Artificial Intelligence Disclosure:** This report has been researched, drafted and edited with the assistance of artificial intelligence tools, including large language models, in addition to conventional analyst research, primary interviews and desk research. AI assistance was used for tasks including, but not limited to, content drafting, data synthesis, structural organisation, language editing and formatting. All AI-generated content has been reviewed, edited and verified by human analysts at Peak State Consulting. Notwithstanding such review, AI language models may produce outputs that contain errors, omissions, hallucinations, outdated information or unintended inaccuracies. The use of AI tools does not diminish the editorial responsibility of Peak State Consulting for the content of this report, but readers are advised that AI-assisted research and writing carries inherent limitations that may affect the completeness or precision of certain statements. Peak State Consulting does not represent that any particular statement, statistic or characterisation generated with AI assistance has been independently verified beyond the review process described above. Readers are encouraged to independently verify specific data points before relying on them for material business or investment decisions.

**Not Professional Advice:** This report does not constitute, and should not be relied upon as, legal, financial, investment, tax, accounting, regulatory, cybersecurity, procurement or other professional advice of any kind. The views, analysis and strategic recommendations expressed in this report represent the general opinions of Peak State Consulting analysts based on the information available to them at the time of writing and do not take into account the specific circumstances, objectives, risk tolerance, financial position, regulatory obligations or operational context of any individual reader or organisation. Readers should obtain independent professional advice, appropriately qualified and tailored to their specific circumstances, before making any business, technology, investment, procurement or operational decision of any kind.

**Forward-Looking Statements:** Artificial intelligence markets, technologies, regulations and vendor capabilities evolve with exceptional speed. This report contains forward-looking statements, including market size projections, adoption forecasts, valuation estimates, regulatory predictions, competitive assessments and strategic recommendations. Such statements are inherently speculative and involve assumptions, risks and uncertainties that may cause actual outcomes to differ materially from those expressed or implied herein. Past performance and historical trends are not a reliable indicator of future results. Any projections, estimates or predictions contained in this report are based on data and assumptions available at the time of publication and may quickly become outdated as markets, technologies and regulatory frameworks evolve. Peak State Consulting undertakes no obligation to update, revise or correct any forward-looking statement contained in this report following publication.

**Third-Party References and Sources:** This report references and cites data from third-party research organisations, analyst firms, public company announcements, news sources and industry publications. Such references are provided solely for informational and comparative analysis purposes and do not constitute an endorsement, recommendation, partnership or commercial affiliation between Peak State Consulting and any referenced organisation, unless explicitly stated. Peak State Consulting has not independently verified all third-party data cited in this report and cannot guarantee the accuracy, completeness or currency of information derived from third-party sources. Third-party organisations referenced in this report have not reviewed, approved or endorsed the content of this publication. All

company names, product names, brand names and trademarks referenced in this report are the property of their respective owners and are used herein solely for descriptive and analytical purposes.

**Limitation of Liability:** To the fullest extent permitted by applicable law, Peak State Consulting Limited, its directors, officers, employees, consultants, agents and affiliated entities shall not be liable for any direct, indirect, incidental, special, consequential, punitive or commercial loss, damage, cost, expense or liability of any kind arising from or in connection with: (i) any reliance placed on this report or any part of its contents; (ii) any errors, omissions, inaccuracies or misstatements contained herein, including those arising from AI-assisted research and writing; (iii) any action taken or not taken on the basis of this report; (iv) any interruption, delay or failure in the transmission or delivery of this report; or (v) any changes in market conditions, technology developments or regulatory requirements occurring after publication. This limitation applies regardless of whether Peak State Consulting has been advised of the possibility of such loss or damage, and regardless of the form of action, whether in contract, tort, negligence, strict liability or otherwise.

**Intellectual Property and Reproduction Rights:** This publication and all of its contents, including but not limited to text, analysis, frameworks, infographics, charts, data visualisations, methodologies and strategic models, are the intellectual property of Peak State Consulting Limited and are protected by applicable copyright, intellectual property and database rights. This report may not be reproduced, redistributed, modified, republished, retransmitted, stored in a retrieval system, or used to train, fine-tune or develop any artificial intelligence model or system, in whole or in part, in any form or by any means, whether electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of Peak State Consulting Limited. Brief quotations of up to fifty words may be used in commentary, criticism or review with appropriate attribution to Peak State Consulting as the source. Any unauthorised reproduction, redistribution or use of this report, including use as AI training data, constitutes an infringement of Peak State Consulting's intellectual property rights and may result in legal action.

**Securities, IPO and Investment Tracker Disclaimer:**

Any references in this report to initial public offerings, proposed public listings, private financings, venture capital investments, mergers, acquisitions, fundraising activity, valuations, share prices, market capitalisation, investor participation, funding rounds, securities offerings or other investment-related activity are provided solely for general market intelligence, industry mapping and informational purposes. Such references do not constitute and should not be interpreted as an offer to sell, a solicitation of an offer to buy, a recommendation to purchase, hold or sell, or an invitation to participate in any security, investment product, financing transaction, fund, company, token, digital asset, IPO, private placement or other investment opportunity.

Peak State Consulting Limited is not acting as a broker-dealer, investment adviser, financial adviser, placement agent, underwriter, securities promoter, sponsor, finder, exchange, trading platform or fiduciary in connection with this report or any company, issuer, investor, transaction or security referenced herein. Peak State Consulting Limited does not provide investment recommendations, investment ratings, price targets, valuation opinions, trading strategies, suitability assessments or personalised investment advice. No reader should make any investment decision based on this report, and readers should consult their own licensed financial, legal, tax and other professional advisers before considering any investment activity.

The IPO and investment tracker sections may include information regarding announced, reported, rumoured, pending, completed or withdrawn transactions. Such information may be derived from public filings, company announcements, press reports, third-party databases, investor disclosures and other sources believed to be reliable at the time of publication, but Peak State Consulting Limited has not independently verified all such information. Transaction status, valuation, timing, investor participation, deal size, pricing terms, listing plans and other investment-related data may be incomplete, estimated, inaccurate, stale, withdrawn, superseded or subject to change without notice. Inclusion of a company, investor, IPO, financing round or transaction in this report does not imply that any investment opportunity is available, suitable, lawful, approved, endorsed, successful or likely to complete.

Investment in public or private securities, including IPOs, pre-IPO companies, growth companies and artificial intelligence businesses, involves substantial risk, including the possible loss of the entire investment, illiquidity, valuation volatility, dilution, limited operating history, lack of public information, regulatory risk, market risk and execution risk. Past performance, prior fundraising success, reported valuation, investor backing, customer adoption, media coverage or market interest is not indicative of future results.

Unless expressly stated otherwise, Peak State Consulting Limited has not been retained, compensated, sponsored or authorised by any issuer, investor, underwriter, placement agent, broker-dealer or other transaction participant to promote, market or recommend any security, company, transaction or investment opportunity referenced in this report.

No company, issuer, investor or transaction participant has reviewed, approved or endorsed this report unless expressly disclosed. If any material commercial relationship, sponsorship, compensation arrangement, advisory engagement, ownership interest or other potential conflict of interest exists with respect to a company or transaction specifically discussed in this report, Peak State Consulting Limited will seek to disclose such relationship where legally and commercially appropriate.

Independence and Governance: Peak State Consulting Limited is an independent research and management advisory firm. This report has not been commissioned by, sponsored by, paid for by, or prepared at the direction of any of the companies, vendors, platforms or organisations referenced herein. Peak State Consulting has received no financial or other consideration from any third party in connection with the preparation or publication of this report. The analytical views and vendor assessments expressed herein represent the independent judgement of Peak State Consulting analysts and are not influenced by any commercial relationship. Any advisory, consulting or implementation services that Peak State Consulting may offer are separate from and independent of the research and analysis contained in this report. This disclaimer is governed by and construed in accordance with the laws of England and Wales. Any disputes arising in connection with this report shall be subject to the exclusive jurisdiction of the courts of England and Wales. For permissions, corrections, feedback or enquiries regarding this report, please contact: Anuj Khanna, CEO, Peak State Consulting, [anuj@peakstate.com](mailto:anuj@peakstate.com).

## References & Sources

All data, statistics, forecasts and company information cited in this report are drawn from the sources listed below, categorised by type. All sources accessed June 2026.

### Market Research & Forecasts

- IDC. "Worldwide AI Spending Guide Q1 2026." International Data Corporation (IDC), 2026. [idc.com](https://www.idc.com)
- IDC. "AI Infrastructure Spending Caps Historic Year at ~\$90 Billion in Q4 2025; 2029 Spending to Eclipse \$1 Trillion." April 16, 2026. [idc.com](https://www.idc.com)
- Gartner. "Gartner Forecasts Worldwide AI Spending to Grow 47% in 2026, Totalling \$2.59 Trillion." May 19, 2026. [gartner.com](https://www.gartner.com)
- Goldman Sachs. "AI Infrastructure CapEx Outlook: Hyperscaler Spending Model." Goldman Sachs Global Investment Research, May 2026. [goldmansachs.com](https://www.goldmansachs.com)
- Fortune Business Insights. "Artificial Intelligence Market Size, Share & Industry Analysis 2025–2034." 2026. [fortunebusinessinsights.com](https://www.fortunebusinessinsights.com)
- Mordor Intelligence. "Artificial Intelligence in Enterprise Applications Market Report." 2026. [mordorintelligence.com](https://www.mordorintelligence.com)
- Verdantix. "Enterprise AI Platform Market Forecast 2024–2030." Verdantix Group, 2026. [verdantix.com](https://www.verdantix.com)
- Sacra Research. "Anthropic Revenue, Valuation & Funding." June 2026. [sacra.com](https://www.sacra.com)

### Industry Surveys & Analyst Research

- McKinsey & Company. "The State of AI in 2025." QuantumBlack / AI by McKinsey, November 2025. [mckinsey.com](https://www.mckinsey.com)
- McKinsey Global Institute. "The Economic Potential of Generative AI: The Next Productivity Frontier." McKinsey & Company, 2023. [mckinsey.com](https://www.mckinsey.com)
- McKinsey & Company. "The Value of Getting Personalisation Right, and Wrong." McKinsey & Company, November 2021. [mckinsey.com](https://www.mckinsey.com)
- Gartner. "Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026." August 2025. [gartner.com](https://www.gartner.com)
- Gartner. "Lack of AI-Ready Data Puts AI Projects at Risk." February 2025. n=248. [gartner.com](https://www.gartner.com)
- Gartner. "Gartner Survey Finds 45% of Organizations With High AI Maturity Keep AI Projects 3+ Years." 2025. [gartner.com](https://www.gartner.com)
- Gartner. "Infrastructure & Operations Survey: AI Adoption Barriers 2025." Gartner, 2025. [gartner.com](https://www.gartner.com)
- Forrester Research. "Agentic AI Comes of Age: The Forrester Wave Q1 2026." Forrester Research. [forrester.com](https://www.forrester.com)
- S&P Global 451 Research. "Enterprise AI Adoption Survey Panel 2026." S&P Global Market Intelligence. [spglobal.com](https://www.spglobal.com)
- KPMG. "KPMG 2025 Global CEO Outlook." n=1,325 CEOs, October 2025. [kpmg.com](https://www.kpmg.com)
- KPMG. "Venture Pulse Q1 2026: Global Trends in Venture Investment." Q1 2026. [kpmg.com](https://www.kpmg.com)
- PwC. "MoneyTree Report Q1 2026." PricewaterhouseCoopers, 2026. [pwc.com](https://www.pwc.com)
- PwC. "AI reshapes global labour market into two distinct paths, rewarding human skills: PwC 2026 Global AI Jobs Barometer." PricewaterhouseCoopers, 15 June 2026. [pwc.com](https://www.pwc.com)
- IBM. "Global AI Adoption Index 2024." n=8,584 IT professionals, January 2024. [ibm.com](https://www.ibm.com)
- IBM. "IBM Cybersecurity Intelligence Report 2025." IBM Security, 2025. [ibm.com/security](https://www.ibm.com/security)
- PEX Network. "PEX Report 2025/26: AI Adoption Barriers." n=200+ professionals, 2025. [aiandanalytics.network](https://www.aiandanalytics.network)
- ACA Group / NSCP. "2025 AI Benchmarking Survey." October 2025. [financialcontent.com](https://www.financialcontent.com)
- Deloitte / McKinsey. "Industrial AI Benchmarking: Predictive Maintenance and Supply Chain AI." 2024.

Crunchbase. "AI Funding Tracker: Global Venture Rounds 2026." June 2026. [crunchbase.com](https://crunchbase.com)  
LinkedIn / Glassdoor. "AI Specialist Salary and Hiring Trends 2025–2026." LinkedIn Talent Insights and Glassdoor Salary Reports, 2025–2026. [linkedin.com](https://linkedin.com) / [glassdoor.com](https://glassdoor.com)

### **Company Announcements, Press Releases & News Sources**

Anthropic. "Anthropic and Deloitte Announce Expanded Enterprise Alliance." October 6, 2025. [anthropic.com/news/deloitte-anthropic-partnership](https://anthropic.com/news/deloitte-anthropic-partnership)

Anthropic. "Anthropic and Salesforce Announce Strategic Partnership." September 2024. [anthropic.com/news/salesforce-partnership](https://anthropic.com/news/salesforce-partnership)

Anthropic. "Anthropic and Snowflake Expand Partnership." 2026. [anthropic.com](https://anthropic.com)

Anthropic. "Introducing the Claude Partner Network." March 2026. [anthropic.com](https://anthropic.com)

Anthropic. "Claude Enterprise: Product Overview and Pricing Documentation." June 2026. [anthropic.com](https://anthropic.com)

Anthropic. "Claude API Documentation and Pricing Guide." June 2026. [docs.anthropic.com](https://docs.anthropic.com)

Anthropic. "GitLab Enterprise Customer Story." 2025. [claude.com/customers/gitlab-enterprise](https://claude.com/customers/gitlab-enterprise)

Anthropic. "Model Context Protocol." November 2024. [github.com/modelcontextprotocol](https://github.com/modelcontextprotocol)

Anthropic. "Claude's Character and Constitutional AI: Safety Documentation." 2025. [anthropic.com/research/constitutional-ai](https://anthropic.com/research/constitutional-ai)

NVIDIA. "NVIDIA Announces Financial Results for Fourth Quarter and Fiscal 2026." February 2026. [investor.nvidia.com](https://investor.nvidia.com)

Bloomberg. "Anthropic Is Exploring IPO as Soon as October 2026." Bloomberg, May–June 2026. [bloomberg.com](https://bloomberg.com)

Bloomberg. "Cerebras Systems Prices IPO at \$185, Surges on First Trading Day." Bloomberg, May 13–14, 2026. [bloomberg.com](https://bloomberg.com)

Bloomberg; CNBC. "OpenAI Confidentially Files Draft S-1 with Goldman Sachs and Morgan Stanley." June 2026. [bloomberg.com](https://bloomberg.com)

CNBC. "Anthropic Annualised Revenue Reaches \$47 Billion." CNBC, May 28, 2026. [cnbc.com](https://cnbc.com)

Reuters. "SpaceX Completes Record IPO, Trades Above \$2 Trillion in Nasdaq Debut." Reuters, June 2026. [reuters.com](https://reuters.com)

SEC EDGAR. "S-1 and Registration Statement Filings." U.S. Securities and Exchange Commission, 2026. [sec.gov/edgar](https://sec.gov/edgar)

AI Funding Tracker. "AI IPO Pipeline and Private Market Valuations." June 2026. [aifundingtracker.com](https://aifundingtracker.com)

DevOps.com. "Anthropic Launches Claude Security for Enterprise Cybersecurity." April 2026. [devops.com](https://devops.com)

OpenAI. "Accelerating the Next Phase of AI: \$122 Billion Funding Round at \$852 Billion Post-Money Valuation." Q1 2026. [openai.com](https://openai.com)

Anthropic. "Anthropic Raises \$65 Billion in Series H Funding at \$965 Billion Post-Money Valuation." May 28, 2026. [anthropic.com](https://anthropic.com)

Databricks. "Databricks Is Raising a Series K Investment at a >\$100 Billion Valuation." 2026. [databricks.com](https://databricks.com)

xAI. "xAI Raises \$20 Billion Series E." Q1 2026. [x.ai](https://x.ai)

Harvey. "Harvey Raises at \$11 Billion Valuation to Scale Agents Across Law Firms and Enterprises." Q1 2026. [harvey.ai](https://harvey.ai)

CoreWeave. "CoreWeave Announces Pricing of Initial Public Offering at \$40 per Share." March 2025. [investors.coreweave.com](https://investors.coreweave.com)

Anduril, Mistral AI, Scale AI, Cohere, Glean, Reflection AI, Cursor / Anysphere and Safe Superintelligence. Company funding-round announcements and investor-relations releases, 2025–2026. Compiled via Crunchbase and TechCrunch.

### **AI Coding & Developer Tools Research**

DataStudios.org. "Claude in the Enterprise." September 2, 2025. [datastudios.org](https://datastudios.org)

GitHub. "Octoverse 2025: State of Open Source and AI." GitHub Inc., 2025. [github.com/octoverse](https://github.com/octoverse)

Bloomberg / TechCrunch. "Cursor's \$2B ARR and \$50B Valuation Talks." April 2026. [bloomberg.com](https://bloomberg.com) / [techcrunch.com](https://techcrunch.com)

Sacra Research. "Claude Code: 29 Million Installs and GitHub Commit Share." Sacra Research, 2026. [sacra.com](https://sacra.com)

Digital Applied Research. "MCP Download Metrics and Ecosystem Report Q1 2026." April 2026.

Digital Applied Research. "Chinese AI Model Share of Global Coding Tokens Q2 2026." Digital Applied Research, 2026.

SitePoint / Kanerika / NxCode / Metacto. "AI Coding Platform Comparison Reports 2026." Various authors, 2026. [sitepoint.com](https://sitepoint.com) / [kanerika.com](https://kanerika.com) / [nxcode.ai](https://nxcode.ai) / [metacto.com](https://metacto.com)

### **AI Pricing & Commercial Analysis**

CloudZero. "Claude AI Pricing Guide: Full Cost Breakdown 2026." June 2026. [cloudzero.com](https://cloudzero.com)

mem0.ai. "Anthropic Claude Pricing Guide: API, Pro, Max, Team and Enterprise." June 2026. [mem0.ai](https://mem0.ai)

Tygart Media. "Claude AI Pricing: Complete Plan Comparison June 2026." June 2026. [tygartmedia.com](https://tygartmedia.com)

### **Regulatory & Governance**

European Union. "Regulation (EU) 2024/1689, Artificial Intelligence Act." Official Journal of the European Union, 2024. [eur-lex.europa.eu](https://eur-lex.europa.eu)

National Institute of Standards and Technology. "Artificial Intelligence Risk Management Framework (AI RMF 1.0)." NIST, January 2023. [nist.gov/system/files/documents/2023/01/26/AI%20RMF%201.0.pdf](https://nist.gov/system/files/documents/2023/01/26/AI%20RMF%201.0.pdf)

OpenRouter. "Global AI API Token Traffic Share by Provider." OpenRouter Platform Analytics, Q2 2026. [openrouter.ai](https://openrouter.ai)