Deloitte.

Executive decisions shaping the value of the digital workforce

From value creation, operational impact, token economics, and rethinking work—executives have a bigger role to play than they think in Agentic AI.

Executive Summary

New AI capabilities are being announced at mind-boggling speed. And yet the pace at which enterprises are adopting and scaling these AI advances remains slow. Organizations are struggling to move from experimentation to impact, constrained by the complexity of understanding, implementing, and deploying AI at scale. Ironically, the same transformational potential that makes AI so compelling is also what makes it difficult to adopt.

While agentic capabilities are technically ready, most organizations have yet to activate them beyond experiments or simple use cases. At its core, this is a leadership challenge. In both our research and hands-on experience helping clients implement agentic AI across their organizations, we consistently see the outsized role senior leaders play in determining whether these technologies translate into real business value.

For the C-suite, success requires more than sponsorship or investment approval. Leaders must understand the economics of agentic AI, make informed software and infrastructure decisions, and guide their organizations to fundamentally rethink how work gets done. Those who do are already unlocking stronger business outcomes and building durable competitive advantage.

To that end, each executive has a distinct role to play in scaling agentic AI successfully. This report provides role-by-role guidance for senior leaders, outlining the responsibilities, decisions, and questions required to move agentic AI from experimentation to enterprise impact.

Agentic Acceleration

There's a good reason agentic AI has captured the attention of leaders of all stripes. Who wouldn't want to unleash an army of AI agents that can reason and take action on so many of the tasks creating bottlenecks for human teams—obstacles that can negatively impact customer experience, drive up costs, or limit growth? The potential benefits in terms of both efficiency and new growth opportunities are immense—and the investments are starting to flow as organizations begin to shift from experimentation to implementation at scale.

Over the next three years, the benefits of agentic AI are widely expected to significantly outweigh its costs. Nearly half (45%) of the 500 leaders who participated in Deloitte's 2025 Tech Value Survey[1] anticipate seeing a return on investment from basic AI automation within this period. Successfully deploying agentic AI capabilities to achieve sustainable benefits over the long haul requires starting with a deeper understanding of how the costs work and developing a smart, well-informed plan to manage them—just as with any other initiative launched to increase efficiency and fuel growth. But with AI agents, the cost drivers are different.

AI agents—digital workers—can be far less expensive than their human counterparts, but they do require payment for their work in the form of tokens. All agentic AI models consume tokens—small units of data that the model uses in training and inference. Every token is both a measure of AI work and a unit of cost. This is the foundation of token-based economics—and it represents a new dimension in how organizations need to plan, budget, and execute AI-enabled strategies.

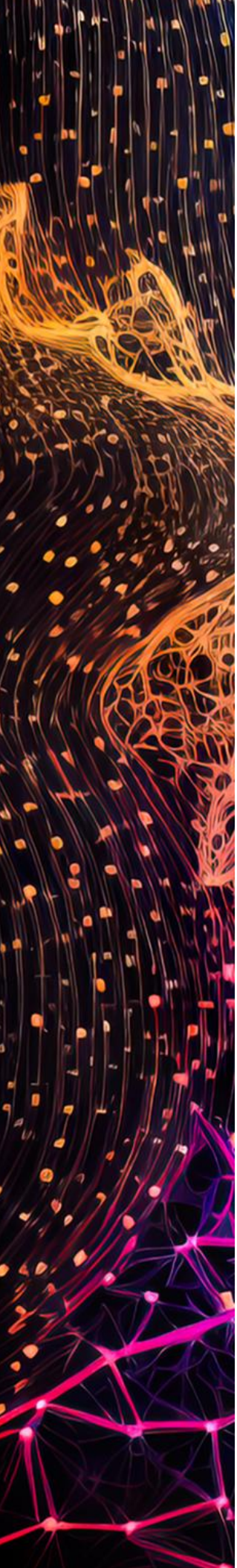
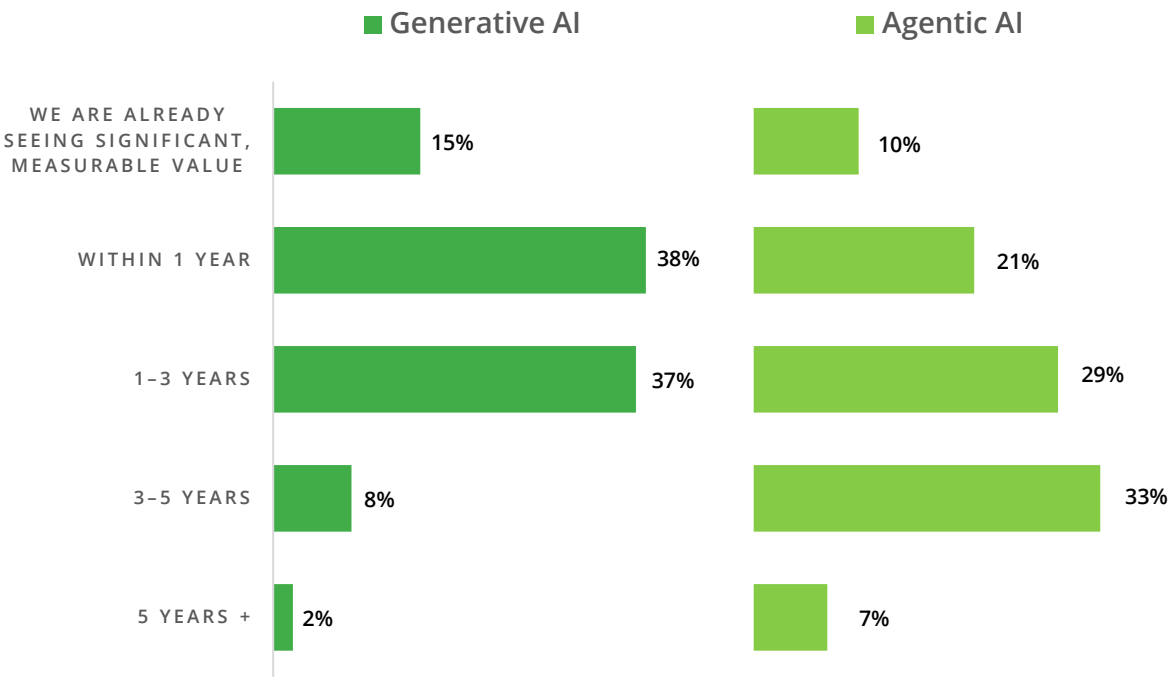
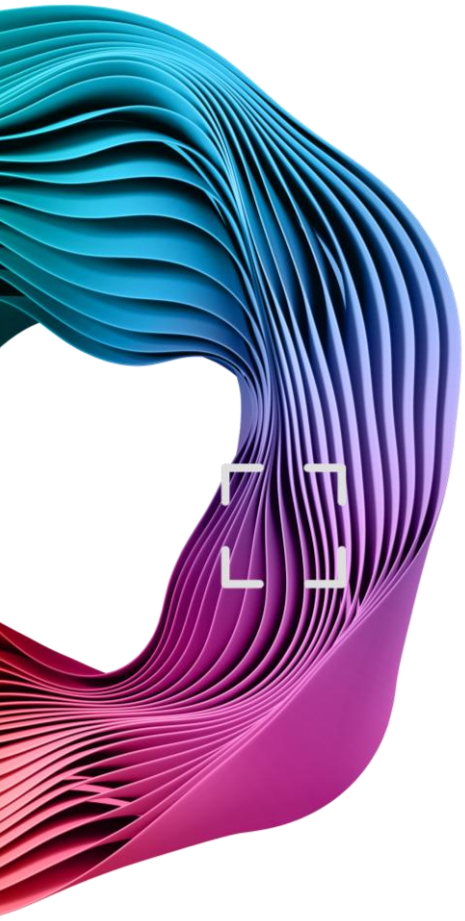


Figure 1: When do you expect to see significant, measurable return on investment from your organization's implementation of AI?

Agentic AI is only in the early stages of adoption, but respondents to Deloitte's recent survey expect to see significant, measurable returns from their agentic AI investments in the coming years.



Source: Deloitte, AI ROI: The paradox of rising investment and elusive returns, 2025 survey of 1,854 executives across Europe and the Middle East.



As organizations seek to scale their digital workforces, reinvent functions, and rethink how work is done, token-based economics requires the attention of the full executive team, not just the CIO. Choices regarding technology, operational strategy, and budgeting that aren't informed by an understanding of the underlying economics of AI can leave the organization exposed to the risk of quickly escalating costs due to unchecked token consumption. This is because digital workers are not static—they will scale up and down depending on workload, making the cost dynamic based on the organization's needs. This elasticity of token consumption is amplified by other choices such as which large language model (LLM) is used, infrastructure choices, and application architecture—which may drive either higher consumption or higher cost of tokens, making the cost line sensitive to these choices.

What's **driving** token costs?

Tokens are the currency of agentic AI—and their costs are shaped by a range of factors set in motion by business and technology decision-making.

Infrastructure efficiency

Compute
Storage
Networking
Energy

Agent design

Reasoning depth
Prompt structure
Memory

Hosting models

On-premises
Cloud
API access

As AI agents extend beyond isolated tasks into collaborative models with large numbers of them working together to take on complex, end-to-end workflows, their numbers, token consumption, and resulting costs will increase. Even as token costs come down, costs will rise as organizations consume more and more—following Jevon's Paradox[2].

For example, if a single AI agent is used to generate a customer email today, a team of AI agents (under human supervision) will soon facilitate a growing share of the end-to-end customer experience. This type of digital team will scale up and down throughout the day based on the volume of calls and their complexity. This is significantly different from expanding a human workforce, where the costs are known. Digital workforces grow dynamically. As a result, their payment in tokens will be unpredictable, leading to variable cost line items that no CFO would accept.

Tokens are effectively a new cost line item. If not well understood, costs can scale faster than expected and drag down ROI. With careful, thorough planning at the outset based on the understanding of technology, operational, and strategic choices on the cost drivers of AI work, this is avoidable. Understanding and managing the costs of agentic AI is one of the most powerful levers executives can pull to make agentic AI a long-term sustainable success in your human and digital workforce.

How fast is token consumption **growing?**

NVIDIA projects a billion-fold surge in AI computing, and Google now processes 1.3 quadrillion tokens a month. The cost implications of this acceleration in consumption are profound—and since we are early in the lifecycle of AI adoption, this will only grow[3,4].

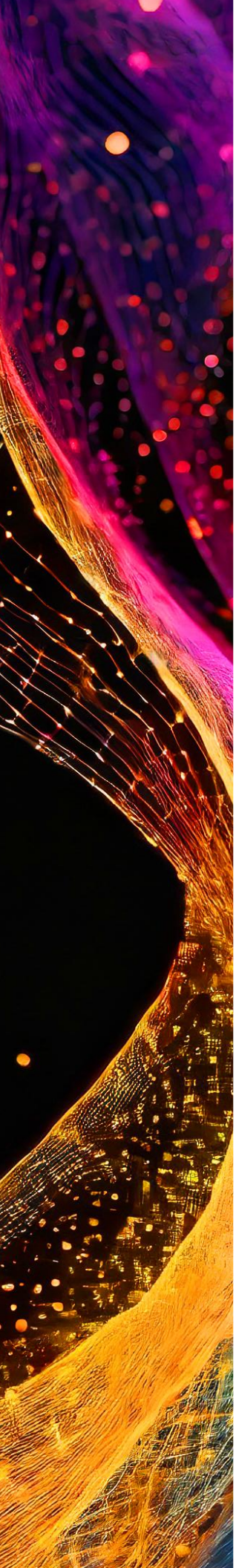
Smart decisions today, **stronger outcomes tomorrow**

Agentic AI requires smart planning and coordination among the C-suite at the outset to ensure that the organization gains the full benefit of these powerful tools. Leaders need to not only think differently about how to use agentic AI capabilities, but they need to also think differently about the benefits they expect to achieve with them as a result.

Strategic technology investments tend to follow a predictable, linear pattern on their way to delivering benefits. For example, when a company invests in a new supply chain solution, its leaders can reasonably anticipate its benefits and financial impact once implemented. The time horizon is similarly clear: If we start the implementation on X date, we can plan on seeing the benefits of these new capabilities by Y date.

Agentic AI is different for two main reasons. First, an investment in agentic AI today is rooted in the (correct) assumption that its capabilities and the benefits they deliver will be significantly more advanced in the not-too-distant future, given the sustained rapid pace of AI innovation. Organizations investing in AI today should plan for the value potential of AI to increase. This means setting up programs to be iterative as well as ensuring technology choices are flexible.

For a clearer idea of the new opportunities agentic AI will likely be able to unlock in the near future, consider this list of anticipated capabilities and their projected evolution in a three-year timespan.



Transformative agentic AI capabilities projected to impact business value by 2028

Capability	2025	2028 (projected)
Autonomy and decision-making	Primarily human-in-the-loop, rule-bound choices	Significantly more autonomous and capable of making proactive decisions; human role shifts to reviewer and monitoring agents emerge
Reasoning and planning	Can handle structured problems with short-term planning	Strategic, real-time adaptive planning in ambiguous situations
Natural language understanding	Context-aware chat and commands	Near-human level of understanding across nuances and multiple languages
Task complexity and execution	Automates repetitive, rule-based tasks	Orchestrates complex, multistep workflows with dependencies based on a defined goal or outcome
Multi-agent system	Early, small scale collaborations within structured frameworks (e.g., LangGraph)	Widespread adoption of sophisticated multi-agent systems with specialized skills, mirroring a human team
Integration and adoption	Point solutions (e.g., customer service or IT help desk); tech-heavy rollout	Enterprise-wide applications driven by low/no-code integration capabilities

Source: Deloitte, Agentic Enterprise 2028, 2025.

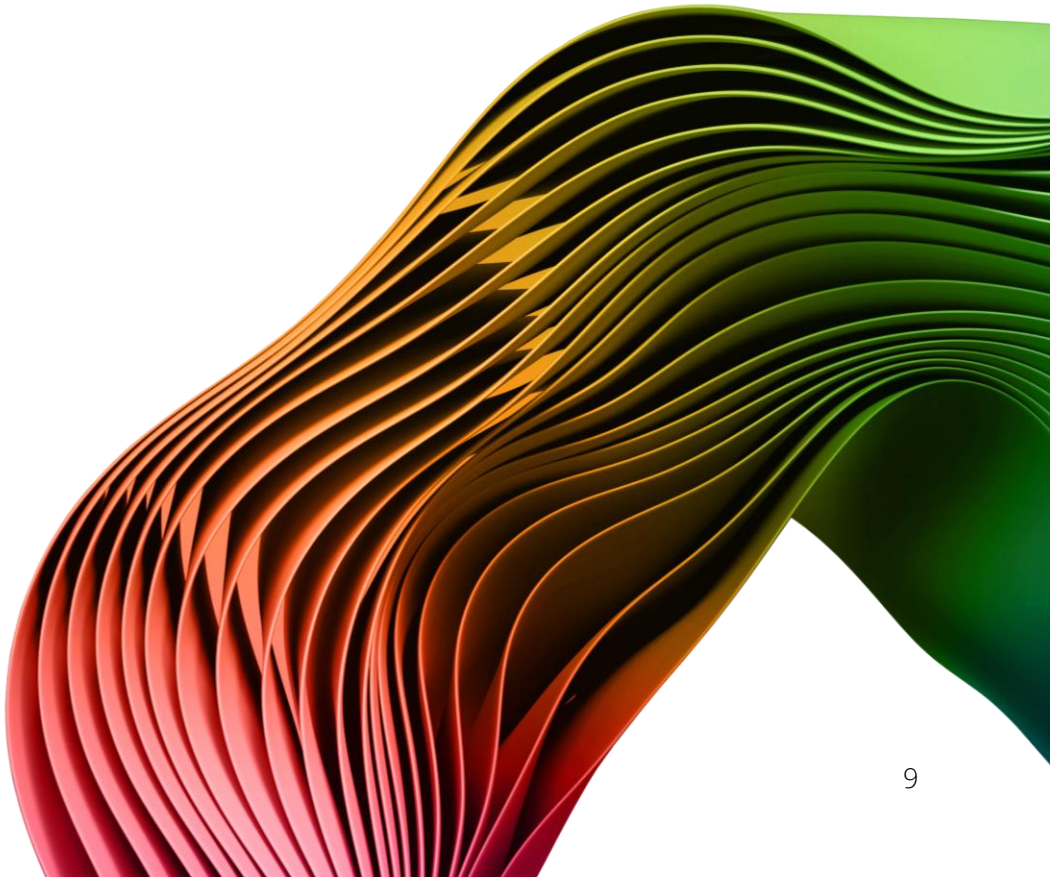
Second, the token-based economics of agentic AI can create unpredictable cost structures that differ fundamentally from traditional technology investments. Because deployment scale and scope remain uncertain, today's choices (on issues including GPU infrastructure, model selection, and architecture) can dramatically affect future expenses. For example, an organization that selects a premium model with verbose prompting on generic cloud infrastructure may face significantly higher token costs as deployment scales from pilot cases to company-wide use. Meanwhile, pursuing an option that optimizes model choice, dedicated GPU infrastructure, and prompt efficiency upfront can maintain predictable pricing at scale.

These two main factors combined mean that the choices executives make about agentic AI today can have a dramatic impact on its value in the future.

Your strategy to scale agentic AI and your digital workforce needs to account for the likelihood that AI capabilities will significantly extend and expand. It also needs to be undertaken with token-based economics at the forefront. Without anticipating and preparing for these two issues, the business case for agentic AI is at risk, putting these benefits in jeopardy:

- **Cost takeout**
- **Productivity**
- **Growth**
- **Workforce transformation**
- **Quality**
- **Speed to value**
- **Customer satisfaction**

That's why it's important to make smart decisions today that leave these possibilities open and financially feasible over time.



The responsibility of **the C-suite**

Making the promise of agentic AI a practical reality that delivers sustainable value, safely and securely, requires new thinking and coordination between all the members of the C-suite. Each executive has unique responsibilities for making agentic AI work at scale in a way that is financially responsible, aligns with human roles and expectations, and delivers stronger outcomes.

Here's a role-by-role view of how each member of the executive team should be thinking about their responsibilities as they move to put agentic AI to work.

What **CEOs** need to know

Agentic AI is not solely a technology choice to be handled primarily by CIOs and their teams. AI and AI agents are transformative—and CEOs need to prioritize and drive the rethinking of work, functions, and adoption across the enterprise, without losing sight of the impacts on their organization.

Brand risk, trust, employee satisfaction, and client satisfaction all hang in the balance together with sustained value, and the CEO must ensure that all downstream decisions regarding agentic AI are made in ways that don't compromise them.

What to do:

Set and steer the company's AI and agentic transformation agenda—defining strategy, overseeing major technology and organizational decisions, and mobilizing the board, employees, and market stakeholders around the vision.

- **Set and communicate the "AI Ambition" for the organization.**
- **Set success criteria.**
- **Lead the charge.**

What to ask:

- **What broader ambitions are we seeking to serve with our AI and agentic AI strategy?**
- **How do we maximize the impact of AI while minimizing the risk?**
- **How will we manage our combined human and digital workforce?**
- **How should we communicate to the market what we are doing with AI and the value it is delivering in our business—and how will we measure progress over time?**
- **Which AI technology advances are real versus short-term hype, and which of them are truly applicable to our business and strategy?**
- **Which AI decisions must be made once, centrally, to avoid duplicated spend and inconsistent risk across the enterprise?**

What **COOs** need to know

As agentic AI deployments gain momentum and scale, Chief Operations Officers (COOs) should focus on driving deployments towards the transformation agenda and tracking value to ensure investments are delivering against success criteria—and achieving measurable results. COOs should pay special attention to issues such as flexibility, orchestration, value delivery, and determining where these capabilities should be applied in the organization.

What to do:

Align technology with enterprise strategy and build the workforce, process, and organizational capabilities required to enable and scale digital and agentic transformation.

- **Understand the cost drivers of agentic AI.**
- **Prioritize agent deployments against value and track this.**
- **Implement monitoring systems for token usage and operational KPIs.**

What to ask:

- **What is the projected impact of AI capabilities in our organization?**
- **Are we applying AI agents to today's processes—or are we rethinking our operations around agents?**
- **Which processes present the greatest opportunities to make meaningful ROI improvements rather than incremental productivity gains?**
- **How do we make the variable cost of agents more predictable and manageable?**
- **What is the balance of build cost vs. run costs over time?**

What **CIOs** & **CTOs** need to know

Chief Information Officers (CIOs) and Chief Technology Officers (CTOs) have full-stack responsibility for AI deployment—from infrastructure to agent build and maintenance. That means they need to determine where AI agents will execute: for example, should agents operate within enterprise software such as enterprise resource planning (ERP) and customer relationship management (CRM) systems, or in custom-built agent systems? How will the organization orchestrate, monitor, and control its multi-agent systems? These leaders need to identify the sources of compute capacity needed to power all these capabilities, which is fundamentally a question of GPU/infrastructure strategy. Infrastructure choices should support predictable performance and steady cost control.

CIOs and CTOs need to be prepared to address a host of other issues that influence the functional outcomes, costs, and ROI of agentic AI. How will the organization manage high-volume token, low-complexity workloads on smaller, less expensive models, compared to lower-volume, high-complexity workloads—where high accuracy is required and larger, more expensive models are needed? How will they ensure flexibility and interoperability along the way?

Plus, they need to make critical decisions that directly influence functional outcomes, cost, and ROI—all in a technology environment characterized by constant, rapid-fire change. They can't afford to wait—yet moving too quickly in the wrong direction risks lock-in or costly rework. Maintaining flexibility is essential as CIOs and CTOs guide the organization toward choices that maximize speed to value and long-term return.

What to do:

Work across business and technology silos to align architecture decisions and technology investments with enterprise-wide strategy.

- **Avoid analysis paralysis—standing still is not an option. Progress is more important than perfection.**
- **Never stop learning—without a clear, up-to-date understanding of the AI landscape, CIOs can't effectively deliver on the organization's agentic AI strategy.**
- **Demand transparency in vendor pricing, especially regarding consumption and what is included within bundled licenses.**
- **Prioritize flexibility to avoid vendor lock-in as the organization scales its use of agentic AI and other AI capabilities. Create a roadmap that allows for a rapidly evolving tech landscape.**

What **CIOs** & **CTOs** need to know

What to ask:

- Do we have the capabilities and capacity in our organization to deliver the AI agenda?
- How should our partnership and alliance strategy evolve?
- What can we build rather than buy? How much of our software development life cycle (SDLC) can we automate?
- What is our short-, medium- and long-term GPU infrastructure plan (e.g., compute capacity, model access, cost-per-work outcomes)?
- Do we have the right balance of centralized (or reusable) capabilities and business unit-specific (or bespoke) tools?
- How will we design our platform so teams reuse shared tools instead of building duplicated agents that each incur new run costs?

What **CHROs** need to know

As organizations scale their digital workforce, it will have an impact on the combined human and digital workforce dynamics, work design, roles, and responsibilities—and will create new capabilities and skill requirements as well as introduce new work culture considerations. Chief Human Resources Officers (CHROs) will need to ensure their organization designs and talent strategies are flexible enough to adjust quickly to this new dynamic. The CHRO is key to redesigning for the work and the workforces we see in our organizations in the future—now is the time for them to think deeply about the implications of the emergence of AI agents. CHROs should be prepared now to offer informed, expert opinions to other executives weighing their options for deploying AI agents.

What to do:

Lead the organizational and talent strategy to enable work and workforce transformation, including redesigning roles, reskilling the workforce, and ensuring leaders and employees are equipped and supported to thrive alongside AI.

- **Develop a human-agentic workforce strategy.**
- **Develop and deploy AI literacy and adoption programming.**
- **Embed governance to prevent "shadow" AI adoption—the unauthorized use of AI solutions within the organization.**
- **Develop leaders who know how to lead and manage human-agentic workforces.**

What to ask:

- **What does our combined human and agentic workforce look like? What is its size, shape, and composition?**
- **How does our operating model and organization design need to evolve in the next 3-5 years as our organization integrates agentic into enterprise workflows?**
- **How do we ensure our uniquely human team is retained, engaged, and high-performing while we scale our use of agents?**
- **How do we evolve our culture?**
- **How should we adjust our internal metrics and incentive structures?**
- **How do we drive adoption of AI with our teams?**

What **CFOs** need to know

Traditional IT metrics don't apply to agentic AI, so Chief Financial Officers (CFOs) need to hone their understanding of token-based economics, in close consultation with their CIO counterparts. For example, although token prices appear to be steadily decreasing, if agentic AI usage goes unchecked, the total spending amounts can still grow. The CFO's role is not simply to set token-based spending limits, but to understand the potential value and tradeoffs that come with increased spending on agentic AI. Developing a smart framework that allows agentic AI to thrive while avoiding unwanted surprises for the finance team is vital.

What to do:

Work closely with the CEO, CIO, and CTO to align funding, governance, and financial oversight of the agentic transformation, ensuring disciplined investment and measurable value realization.

- **Add "token consumption" as cost line item.**
- **Set a token unit cost target and limit.**
- **Value-track projects including tokens.**
- **Implement new capabilities for tracking token consumption and forecasting demand.**
- **Develop and enforce ROI thresholds to ensure sustainable AI investments and prevent budget overruns.**

What to ask:

- **Are our AI investments distributed appropriately across ambition levels?**
- **What is the projected impact of AI capabilities in our organization? Are our investment plans aligned with these goals?**
- **At what point should we consider investing in our own GPU infrastructure?**
- **How do we track the cost and efficiency of our digital workers?**
- **How are we forecasting token consumption and aligning it with ROI expectations?**
- **Which costs scale with usage (tokens/ compute) vs. fixed, and what are the caps on spending?**
- **What is our target unit cost for an agent to complete a standard task? What architecture decisions most influence that number?**
- **What financial controls should we have in place to prevent runaway AI costs as adoption scales?**

What **CROs** need to know

AI has already opened the door to a new world of risks that Chief Risk Officers (CROs) must understand, monitor, and help the organization manage effectively—and agentic AI simply adds another layer of complexity for them to address. While different members of the executive team must understand and manage the risks that are unique to their domains, the CRO should be prepared to offer oversight and guidance to ensure an overall risk-aware approach aligned with the organization's risk tolerance.

Their role is to make sure models are reliable, prevent bias, and guarantee that AI outcomes are fair and understandable. These responsibilities become even more important as AI adoption accelerates. CROs can rely on well-established safeguards such as clear accountability, thorough documentation, ongoing change management, rigorous monitoring, and effective incident response processes to prevent AI systems from become opaque "black boxes," and to ensure that governance matches both regulatory and economic requirements.

What to do:

Define the enterprise risk appetite for AI and agentic systems, ensuring that governance, controls, and third-party arrangements enable business objectives while meeting regulatory and privacy requirements.

- **Set clear governance—which should include guidance on token consumption.**
- **Develop risk frameworks for AI adoption.**
- **Monitor volatility in token economics.**
- **Ensure regulatory guidance and track regulatory updates as they unfold.**
- **Embed security by design.**

What to ask:

- **Are our contracts, governance, and privacy safeguards aligned with current and emerging AI regulations? How are we ensuring compliance as we scale AI workloads under token-based pricing models?**
- **What level of autonomy are we willing to grant agents in each domain?**
- **What approvals and controls will we require before allowing agents to make decisions?**
- **Do our vendor agreements include provisions for exit strategies to mitigate lock-in risk as token economics evolve?**

Executive teams that look at both sides of the equation will scale sustainable human and agentic teams.

It's a mistake to approach agentic AI as an exercise in cost management. Agentic AI is a powerful capability for generating enterprise value, and leaders should start every conversation about how to scale AI agents with a focus on value creation. At the same time, those conversations should also be grounded in a clear understanding of the mechanics of token consumption and their bottom-line impact. How can the organization best position itself to convert token consumption directly into enterprise value, in ways that both front-line employees can see for themselves, and the Finance team can appreciate, too?

Strategic management of token economics, organizational responsibilities, and technology choices is fundamental to scaling agentic AI in a way that delivers sustainable enterprise value. As token consumption accelerates and costs become increasingly dynamic, leadership must coordinate across the C-suite to ensure every decision—whether in model selection, infrastructure, or operational deployment—converts AI capability directly into measurable business impact. Proactive planning, ongoing governance, and transparent tracking of token usage are essential levers for long-term success, enabling organizations to avoid runaway costs and maximize ROI even as agentic AI evolves. Ultimately, it is informed, decisive leadership—aligning technology investments with strategic ambitions and clearly defined executive responsibilities—that will unlock the full potential of agentic AI and deliver enduring value.

Authors

Stuart Scotis

Deloitte Global Agentic AI Leader
sscotis@deloitte.com.au

Karen Pastakia

Deloitte Global Human Capital Leader
kapastakia@deloitte.ca

Nitin Mittal

Deloitte Global AI Leader
nmittal@deloitte.com

Matt Dalton

Engineering, AI & Data Lead Partner
Deloitte New Zealand
mdalton@deloitte.co.nz

Aravind Subramanian

Partner - AI & Data
Deloitte New Zealand
aravsubramanian@deloitte.co.nz

Amanda Williamson

Director - AI Institute
Deloitte New Zealand
amawilliamson@deloitte.co.nz

End notes

1. Deloitte Insights, "AI is capturing the digital dollar. What's left for the rest of the tech estate?" 16 October 2025. https://www.deloitte.com/us/en/insights/topics/digitaltransformation/aitechinvestmentroi.html
2. Jevons, William Stanley. *The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our CoalMines. Macmillan and Co.,* 1865.
3. Sundar Pinchai, CEO of Google and Alphabet, "Q3 earnings call: Remarks from our CEO", 29 October 2025. https://blog.google/companynews/insidegoogle/messageceo/alphabetearningsq32025/#introduction
4. Caleb Naysmith, "'It's about to go up by a billion times': NVIDIA CEO Jensen Huang says the next 'Industrial Revolution' is already here", Yahoo Finance, 15 October 2025. https://finance.yahoo.com/news/billiontimesnvidiaceojensen135926062.html?

Additional resources on agentic AI

- Deloitte, *AI is capturing the digital dollar. What's left for the rest of the tech estate?* 2025.
- Deloitte, *Agentic Enterprise 2028,* 2025.
- Deloitte, *AI ROI: The Paradox of Rising Investment and Elusive Returns,* 2025.
- Deloitte, *Navigate the Economics of AI,* 2026.
- Deloitte, "Agentic AI and the Future of IT Managed Services," *The Wall Street Journal*, 2025.
- Scotis, Stu and Alonso, David, "Is your agentic AI strategy increasing your cost?" *The Australian*, 2025.
- Scotis, Stu, "AI's new economic reality: Why tokens are reshaping enterprise technology," *The Australian, 2026.*