Gartner

Gartner Insights

2026 Planning Guide for IT Operations and Cloud Management

Joe Antelmi, Marco Meinardi, D.B. Cummings, Tyler Cheney, Stanton Cole, Lydia Leong, Ross Fomerand, Paul Delory, Steve White, Dylan Roberts, Lucas Albuquerque, Ang Troy, Luis Fernandez

2026 Planning Guide for IT Operations and Cloud Management

13 October 2025 - ID G00834028 - 57 min read

By: Joe Antelmi, Marco Meinardi, D.B. Cummings, Tyler Cheney, Stanton Cole, Lydia Leong, Ross Fomerand, Paul Delory, Steve White, Dylan Roberts, Lucas Albuquerque, Ang Troy, Luis Fernandez

Initiatives: IT Operations and Cloud Management for Technical Professionals; Evolve Service Management and Cloud Operations

IT operations teams must evolve people, platforms and processes to support AI agents and self-service, and reduce the burden of supporting a brittle web of legacy systems and dependencies. This Planning Guide helps I&O technical professionals adapt to this evolution.

Overview

Key Findings

- As AI systems become more complex, it is increasingly difficult to evaluate production-ready versus immature use cases. The value of AI adoption is often limited by immature governance models, security concerns and the weight of unrealistic expectations.
- Generative AI (GenAI) and large language model (LLM)-based AI agent systems are immature from an IT operations and FinOps perspective. By rushing GenAI into production and working in silos, organizations address operational requirements as an afterthought and create technical debt.
- IT consumers demand greater autonomy. Vendors are responding with frictionless SaaS solutions and easy-to-deploy offerings that make it simple to establish platforms and democratize technologies. The challenge of governing across these self-service platforms is expanding, becoming more nuanced and increasingly distributed.
- The self-service and AI-agent-enabled future of I&O will fail if there is a lack of investment in standardization, digitization and modernization. Future IT organizations will have both AI agents and mainframes. Successful IT operations teams are discovering creative methods to thread the needle between AI transformation and traditional legacy modernization.

Recommendations

- Start now with internal LLM-based AI agent capabilities for lower-risk use cases to deliver immediate value. Make a longer-term investment in vendor management, change management and upskilling for operators, engineers and architects to assist with scaled, cost-effective AI solution adoption.
- Adopt cross-functional AI engineering platforms to enable mature, resilient AI in production. Establish operational expertise for both adopting GenAI within IT and supporting other teams' GenAI and AI agent systems.
- Develop and deploy policies to govern autonomous users and AI agents on internal platforms. Require governance across platforms, and where possible, unify governance practices. Drive value by creating governance use cases that intersect cost management, security and compliance.
- Identify legacy systems that force "bolt-on" operational fixes or limit the ability to govern autonomous users. Minimize technical debt by identifying "seams" within the legacy system, enabling modular separation and incremental improvement of individual components.
- Accelerate existing initiatives to standardize, automate and streamline IT service delivery, as strong foundations continue to be a prerequisite to successful AI innovation.

Strategic Planning Assumption

Enterprises deploying industry-specific AI agents in support of critical business objectives will grow to over 80% in 2030, up from less than 10% today.

IT Operations and Cloud Management Trends

The pace of AI capability advancement is accelerating, while organizational change remains stubbornly linear. The latest AI model releases are highly capable, mastering performance benchmarks such as SWE-bench that humans previously dominated.

A shift is coming in IT professional roles. AI agent systems will increasingly be built or bought to augment human task completion. IT operations teams are not well prepared for this change. I&O technical professionals must become augmented I&O operators, engineers and architects, mastering the synergistic adoption of AI capabilities in their roles.

Market and political uncertainty, cost pressures and rapid changes will complicate modernization strategies. In a broader environment of market and political uncertainty, many organizations are hesitant to make large, long-term investments. For most organizations, budgets for legacy modernization or innovation outside of AI will remain flat. Organizations will be asked to do their best with what they have. Cost optimization must be a consistent priority. Sometimes, IT teams will be asked to deliver quick, visible wins in case of a budget windfall.

Operating models continue to shift toward self-service and decentralization. The clash between rapid technological changes and limited IT resources is driving changes in operating and service models. Much of IT operations is shifting from centralized to more decentralized models. Individual teams and business units increasingly want to self-serve, develop and deploy their own technologies. For example:

- Application development teams have been driving demand for autonomy and self-service for years.
- AI teams and AI practitioners are also increasingly demanding autonomy and cloud resource access, and, given the hype and pressure to innovate with AI agents, they are often tempted to sidestep existing I&O and cloud governance structures.
- Business technologists and fusion teams are using low-code platforms and AI-enabled tools to create applications and automations. In fact, the adoption of shadow AI via this approach is likely to surpass enterprise-mediated AI adoption. However, more than hallucination, the problem with shadow AI capabilities is sycophancy; they will build the app or capability that the user prompts, even if it's a bad idea, a risky application or a policy violation.

Legacy approaches to IT are insufficient to deliver modern services and systems at scale. AI disruptions and self-service capabilities introduce many operational challenges and risks. IT operations are being stretched to adapt technology, people and processes to the increased volume, variety and complexity of applications to release, operate and govern.

To avoid getting left behind, focus more on enablement and governance and less on delivery. In the past, I&O teams could solve problems with bolt-on fixes. For example, they could install traditional agents as an afterthought and meet requirements for monitoring. They could react to emerging threats by manually modifying production configurations. These retrospective approaches are no longer sufficient.

Existing operating structures that focus on change review, architecture review, or reactive operations processes cannot sustain the scale or pace necessary. Complex processes in IT operations must be mechanized and automated, ideally using the highest-impact, lowest-complexity approach possible.

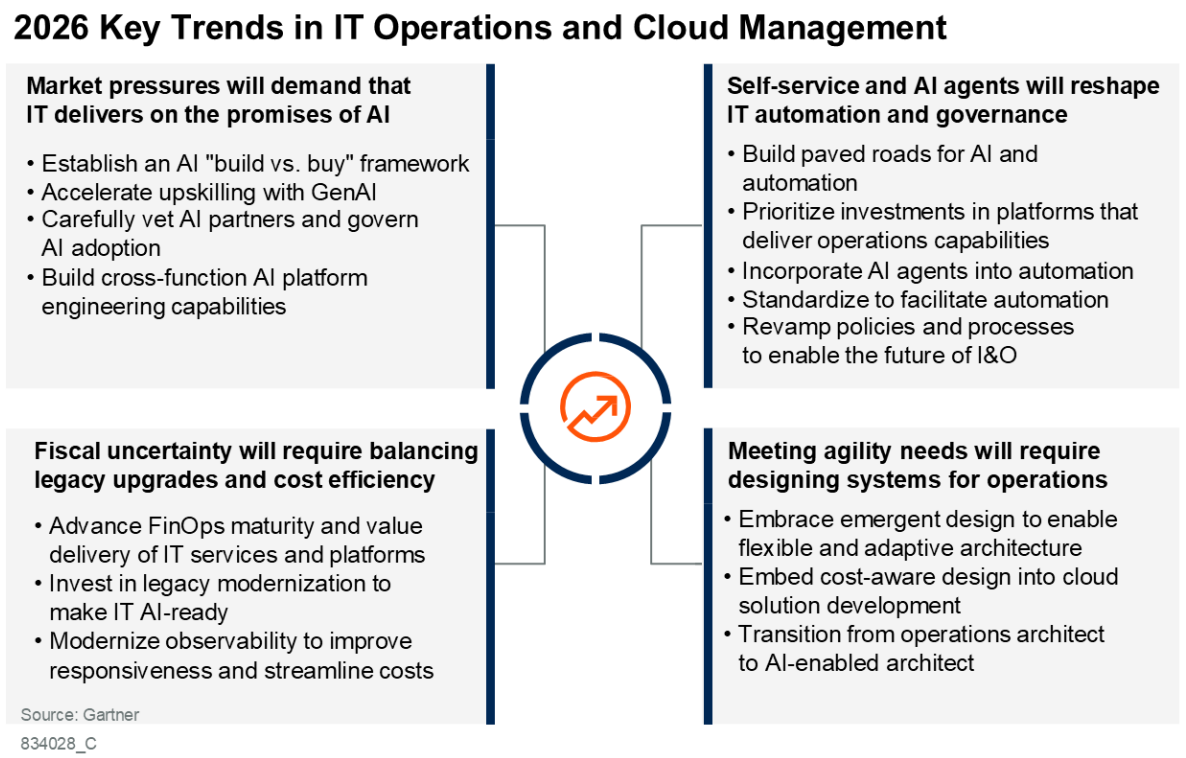
System architecture must become "operations-ready." Solution architects should address operational requirements early to create an environment where agile development becomes the norm and policy enforcement occurs automatically. IT teams should manage everything via policies and as code. Addressing these fundamental shortcomings benefits both:

- **I&O,** by enabling more manageable, efficient and cost-effective infrastructure
- **Consumers,** because they will gain greater autonomy with fewer risks

IT operations teams have the mandate to adapt, build new systems and create new value. If they don't, business leaders will favor expensive outsourced services, shadow IT processes or fragile, error-prone AI solutions. Strong IT operations and infrastructure management is essential. No AI agent or outsourced service will have the same ability to understand, support and protect your organization that you bring in as a seasoned technology professional. You, the I&O technical professional, must lead the change.

Figure 1 outlines the technical trends and associated planning considerations that will be most impactful in 2026.

Figure 1: 2026 Key Trends in IT Operations and Cloud Management



The major trends in IT operations and cloud management for 2026 are (click links to jump to sections):

- Market pressures will demand that IT deliver onthe promises of AI.
- Self-service and AI agents will transform IT automation and governance.
- Fiscal uncertainty will require balancing legacy modernization and cost optimization .
- Meeting agility requirements will require architecting for operations.

Market Pressures Will Demand That IT Delivers on the Promises of AI

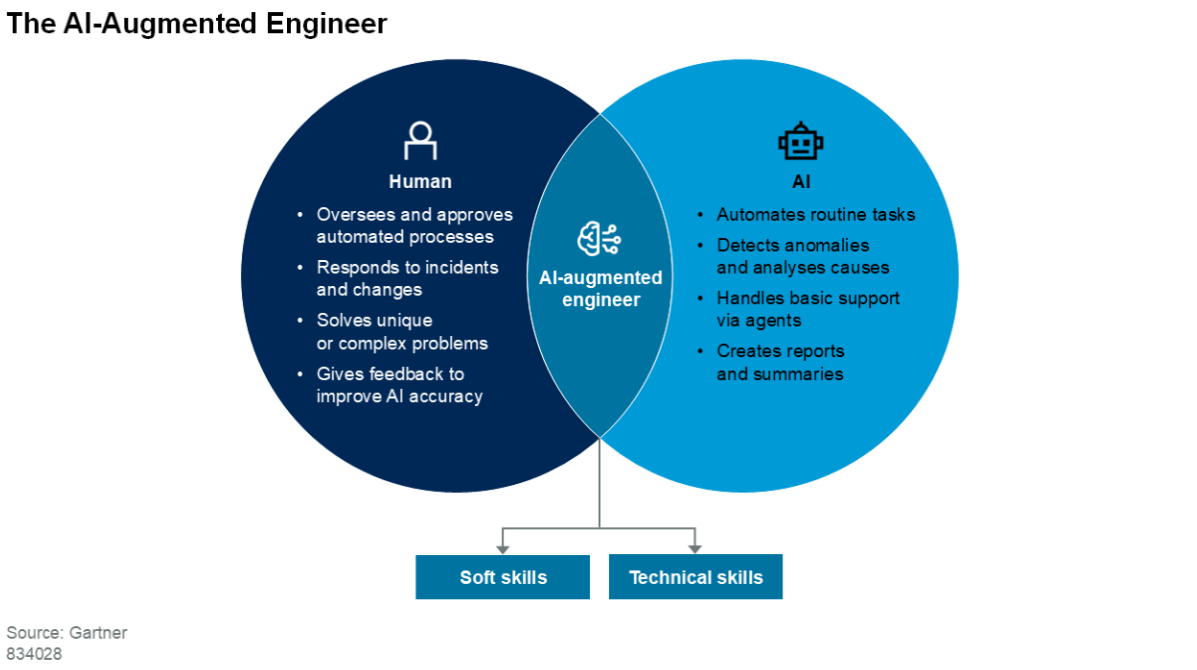
Back to top

In 2026, organizations that move from GenAI pilots to production will be under pressure to demonstrate measurable business outcomes — not just technical achievements. Those delaying GenAI adoption risk falling behind industry leaders who are rapidly gaining efficiency and market differentiation.

As GenAI becomes integral to IT and business processes, the gap between early adopters and laggards will widen at an accelerating pace, with direct implications for competitiveness and the capacity for innovation.

The volume and complexity of work facing organizations far exceed their existing operational capacity. To address this, augment I&O teams and tools with AI and automation. These technologies are critical for improving service quality and efficiency. AI and automation should be seen as strategic enablers, not end goals. Define clear objectives, select production-ready solutions and avoid lengthy implementations that delay ROI. Moreover, start with the end goal in mind and work backward to enable the "augmented operator, engineer and architect" future in I&O. See Figure 2 for an example of the future I&O professional — in this case the augmented engineer.

Figure 2: Trajectory: AI-Augmented Engineer



Source: Gartner
834028

Gartner

The role of the augmented practitioner right now is a practical idea with great potential. Many processes that are currently manual will be replaced by a combination of human skill sets and AI task augmentation. Initially, these kinds of skilled technical professionals will not be easy to find. Demand for AI-skilled talent is projected to outpace supply by two to four times, with significant skills gaps anticipated beyond 2026. [1]

Reinvest a portion of your capacity gains from successful automation and AI initiatives into comprehensive workforce development. Ongoing training is critical — not only to ensure effective use, management and support of AI systems, but also to quickly build capacity and expertise in areas that align with evolving business strategies. To realize GenAI's value, you will need disciplined, business-aligned I&O processes; without them, even advanced AI solutions will not deliver sustainable ROI.

AI and automation are not just operational enhancements; they are foundational pillars for the future of IT operations and digital transformation. By strategically integrating these technologies into operational workflows and reinvesting in human capital, you can achieve sustainable growth, increased agility and a competitive edge in the digital era.

For 2026, I&O technical professionals should (click links to jump to sections):

- Establish an AI "build versus buy" framework for I&O.

- Accelerate upskilling with GenAI.
- Carefully vet AI partners, and govern AI adoption.
- Build cross-function AI platform engineering capabilities.

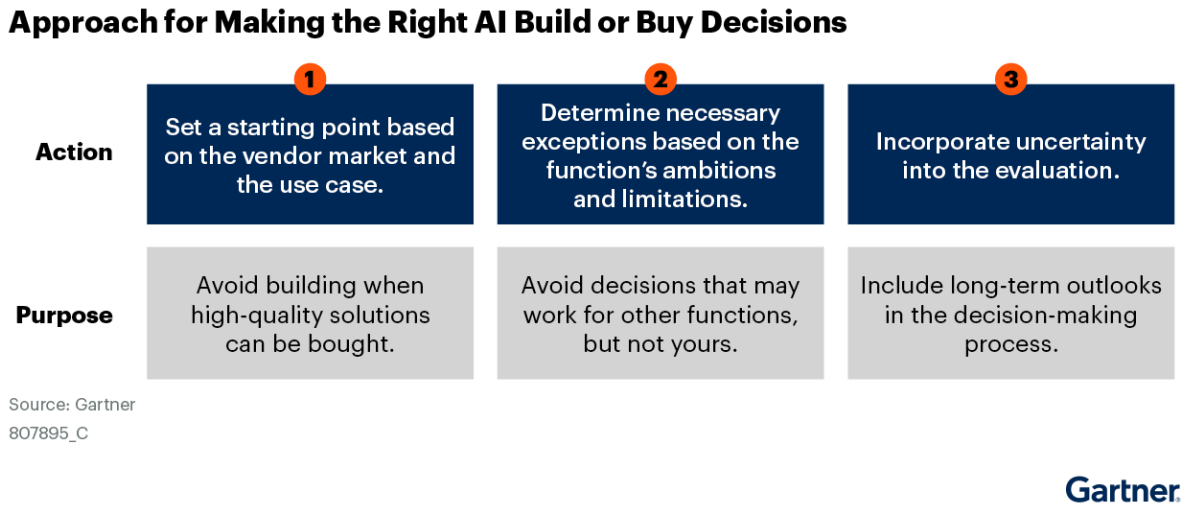
Planning Considerations

Establish an AI "Build Versus Buy" Framework for I&O

Back to top | Back to planning considerations

To facilitate adoption, I&O architects and infrastructure platform engineering teams must develop and enforce a comprehensive framework for build-versus-buy decisions. Define and follow key principles to make AI build and buy decisions, as described in Figure 3.

Figure 3: Approach for Making the Right AI Build or Buy Decisions



These principles are powerful. Many I&O organizations, right now, believe that the path to success is to insert I&O into the full, end-to-end AI solution. This might include deploying AI infrastructure such as GPUs and self-hosted AI services, like a Kubernetes-based AI platform. However, becoming an "AI infrastructure shop" requires advanced AI maturity, capex investments, expensive niche skills, and production-ready AI applications.

In short, much like private cloud, it is unrealistic for the average I&O organization to expect that AI services can be delivered internally with the same efficiency, maturity and scale as major AI and cloud providers.

Instead, look to benefit as much as possible from innovations in cloud-hosted AI services, and avoid building solutions when high-quality solutions can be bought. Carve out exceptions, for instance, pursuing build use cases mainly for staff upskilling and training opportunities rather than business value. And finally, continually revisit build versus buy decisions, and incorporate uncertainty into decision making by piloting before going all-in with huge license purchases or infrastructure purchases for AI capabilities. Many organizations are adopting a hybrid approach where they may, for example, leverage SaaS GenAI tool licenses and cloud-hosted foundation LLMs, and make strategic infrastructure investments to run GenAI on their own AI hardware. However, all approaches have hidden complexities and challenges; in particular, operating AI infrastructure requires skills and software that may be unfamiliar to many I&O teams.

Benefits

- AI adoption is streamlined through governed AI decision making.
- Accounting and documenting the uncertainty inherent in AI investments will help manage stakeholder expectations.
- Long-term planning and use case evaluation will reduce the risk of accidentally buying too many AI tools or making wasted GPU investments, and the associated technical debt and management overheads.

Cautions

- AI can introduce and expose new sources of technical debt, potentially masking serious issues that should be resolved at their source.
- Initial investment in rigorous evaluation processes and comprehensive evaluation frameworks may slow AI adoption and reduce early GenAI benefits, requiring a long-term strategic view.

Related Research

- Integrating GenAI Into Your Application Architecture

Accelerate Upskilling With GenAI

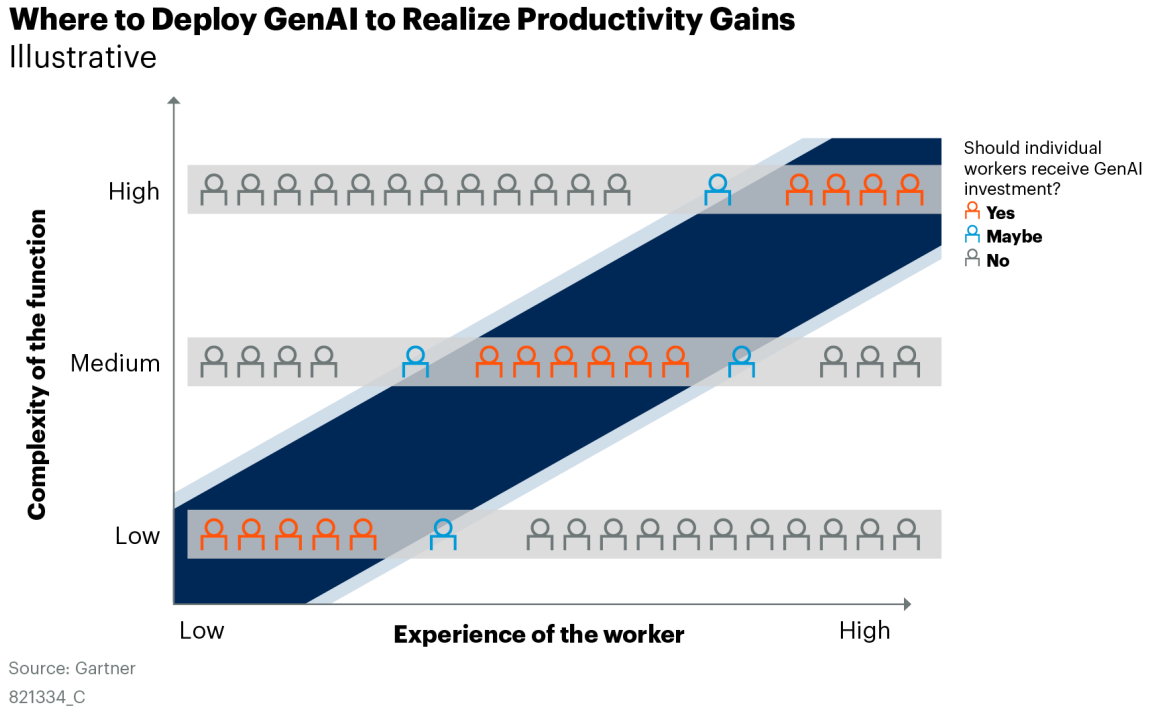
Back to top | Back to planning considerations

The World Economic Forum's Future of Jobs Survey indicates that 63% of employers view skill gaps in the labor market as the primary impediment to business transformation between now and 2030. By 2030, nearly two-thirds of the global workforce will require reskilling or upskilling to maintain their effectiveness in current roles or transition to new positions. [1]

To address these challenges, organizations should view productivity gains from AI and automation not simply as opportunities for cost reduction, but also as catalysts for strategic reinvestment in workforce development. By channeling these capacity gains into upskilling and reskilling initiatives, organizations can strengthen their internal talent pipeline and better meet the growing demand for skilled professionals.

GenAI can facilitate upskilling and reskilling initiatives while promoting a culture of continuous learning and experimentation. By adopting a comprehensive approach to workforce development — reinforced by GenAI — you can accelerate learning, enhance individual and team value, foster adaptability, drive innovation, and support both immediate and long-term business objectives. However, it is not always clear where to deploy GenAI. A useful guide can be seen in Figure 4.

Figure 4: Where to Deploy GenAI to Realize Productivity Gains



Gartner

Consider places to implement AI capabilities to help both inexperienced and experienced employees with the right function complexity. In this analysis, function complexity is a degree of both the difficulty of the tasks as well as the skill set required to achieve them. As you can see in the above figure, there is a “sweet spot” for AI adoption for each employee category. Giving sophisticated AI tool licenses to architects and automation engineers, and enabling AI product features to tier-one help desk employees, is a great starting point to improve individual productivity.

Benefits

- Future-proofing the workforce by investing productivity gains from AI and automation into upskilling and reskilling initiatives enables organizations to organically grow talent to fill their internal talent pipeline.
- An organization culture that is linked to an individual attitude of continuous learning and supported by GenAI enables you to rapidly acquire new skills and adapt to emerging technologies. Applying these skills to strategic initiatives helps organizations close critical skill gaps, driving innovation and accelerating business transformation.
- Comprehensive workforce development initiatives demonstrate a commitment to employee growth and career progression. This not only attracts top talent but also improves retention and engagement, as teams feel valued and equipped to contribute in a meaningful way.

Cautions

- Viewing AI-driven productivity gains solely as cost savings may cause teams to underinvest in workforce development, and instead try to justify staffing cuts with AI. In an environment of limited budgets and economic uncertainty, these types of cuts may be imposed on your organization.
- You should try to make the case not to cut the workforce, as this approach comes with costs. This short-sighted approach can exacerbate skill gaps, hinder transformation efforts and leave the organization ill-prepared for future demands.
- Without aligning I&O workforce development strategy to business objectives, upskilling and reskilling efforts may become fragmented and fail to deliver meaningful business impact.

- Efforts to promote continuous learning and experimentation may face resistance from those unaccustomed to rapid change.
- Part of future-proofing AI workforce development may involve prioritizing the most promising prospective employees. If done poorly, this may introduce additional reputational or regulatory or legal risks for organizations.

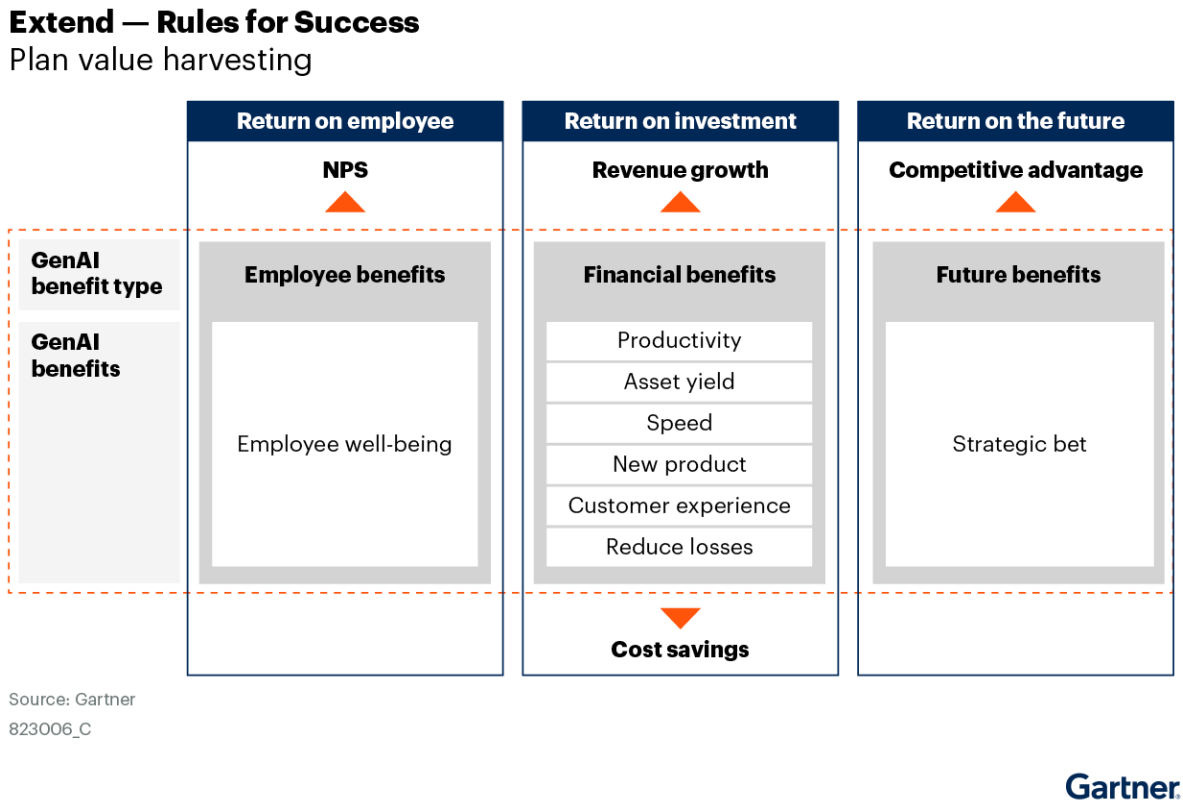
Carefully Vet AI Partners and Govern AI Adoption

Back to top | Back to planning considerations list

Maintain a critical mindset when adopting AI capabilities. When the decision to buy is made, evaluate vendor GenAI tools against actual business problems, rather than relying solely on vendor marketing promises.

Demand concrete proof-of-concept implementations using your real-world use cases, meticulously assessing integration complexity and quantifying potential operational overheads. Establish clear evaluation criteria focused on measurable workflow improvements. Pilot tools for a defined period of 30 to 60 days with explicit success metrics, and crucially, always maintain fallback procedures for situations where GenAI features may fail or underperform. For an example of how to consider these business cases, see The 3 Business Cases of Generative AI Value and Figure 5.

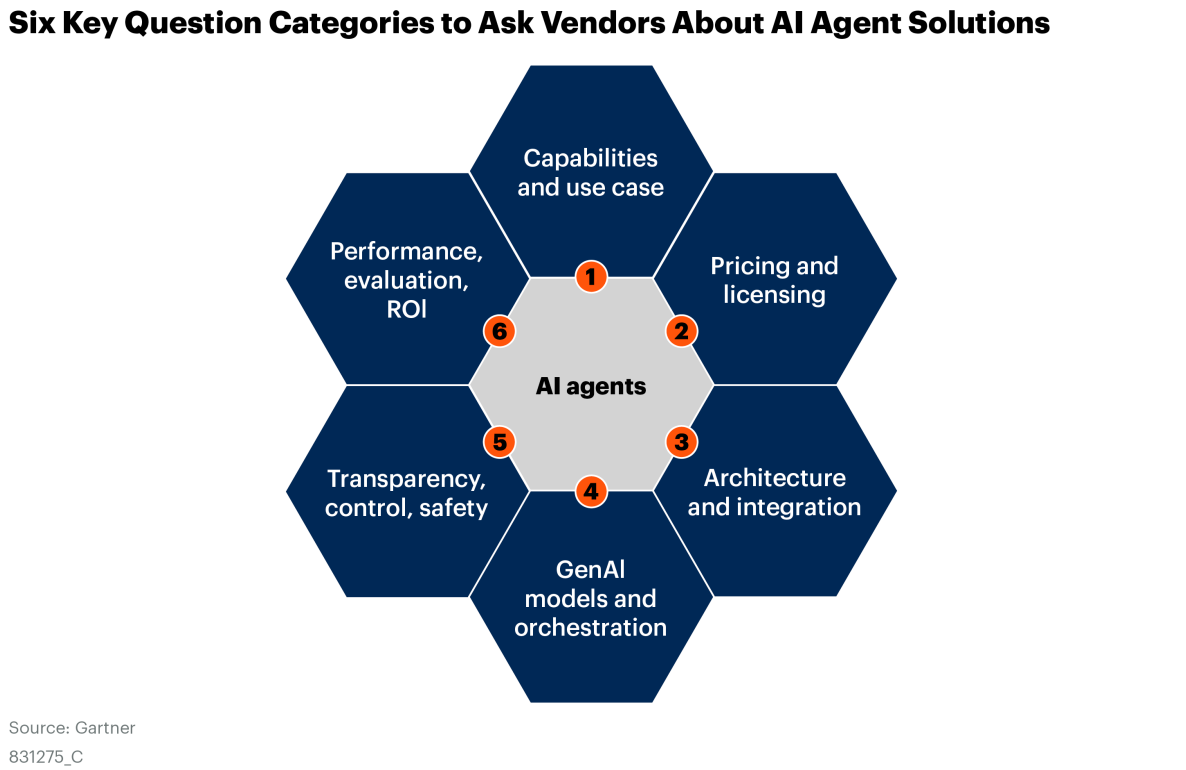
Figure 5: Extend — Rules for Success



In this figure, you can see an example plan to harvest value from GenAI adoption. This includes benefits related to individual employees, benefits for the organization's current operations, and future benefits.

You may not be required to put together the entire business case for GenAI adoption; however, you certainly may be asked to vet AI vendors. The right question domains can help teams carefully vet AI vendor offerings, in accordance with the criteria outlined in Figure 6.

Figure 6: Six Key Question Categories to Ask Vendors About AI Agent Solutions



Source: Gartner
831275_C

Gartner

Overall, the key to managing the complexity of AI solution implementation is orchestration. Automation and AI operate within ecosystems, not in isolation. Creating a unified framework to enforce governance and guide integrations across AI use cases and technical domains is a superior approach to purchasing fragmented, tool-specific AI solutions that are not designed to interoperate.

Along those lines, assume responsibility for learning core GenAI integration patterns, implementing continuous evaluation, modularity, loose coupling (keeping GenAI capabilities loosely coupled from core application logic), and defining robust operational processes to manage costs, nondeterminism, rapid evolution, privacy, intellectual property (IP) and security.

Once an AI platform is built or bought, target use cases that can operate within a defined "error budget" (a concept from site reliability engineering [SRE]) and implement rigorous validation processes. A critical first step in AI governance is to catalog candidate tasks. For example, low-risk, low-complexity and high-volume tasks are great candidates for AI and automation. In the GenAI world, organizations are prioritizing AI use cases that deliver development environment provisioning, documentation generation and noncritical automation scripts. These use cases all have high error tolerance, i.e., failures have minimal immediate impact on production. This strategic approach allows teams to build essential expertise and confidence through low-risk experimentation before scaling to more critical systems.

Benefits

- Careful use case vetting and value plan harvesting reduces the risk of AI failures related to costs, security, and lack of alignment with meaningful outcomes. This is important because trust in AI, once lost, is not easily regained
- Introducing value, error tolerance, task volume, task complexity, costs and feasibility as criteria for AI use case selection aligns with the organization's mission of risk reduction and resilience. I&O will not, and should not, become a high-risk innovation laboratory for AI tools, given its mission to scale, secure and manage costs for IT platforms for the entire organization.

Cautions

- Overambitious spending on AI platforms can hurt ROI if too many AI projects are being implemented. This is both due to implementation and development costs as well as overlapping functionality and inflated expectations.
- Lack of alignment with business objectives often leads to isolated, low-value AI systems or automations that fail to scale effectively.
- Overreliance on AI introduces unnecessary complexity and costs when simpler automation could achieve the same result with less effort and oversight. It also introduces potential errors and security issues, if not well governed.
- AI creates the risk of skill degradation, as humans may default to GenAI solutions when they could try to solve problems themselves.

- Vendor marketing often significantly overstates capabilities while profoundly understating integration complexity and inherent limitations, leading to misaligned expectations and costly misallocations of resources.

Related Research

- The State of AI in Observability
- 4 Agentic AI DevOps Use Cases for Infrastructure Platform Engineers

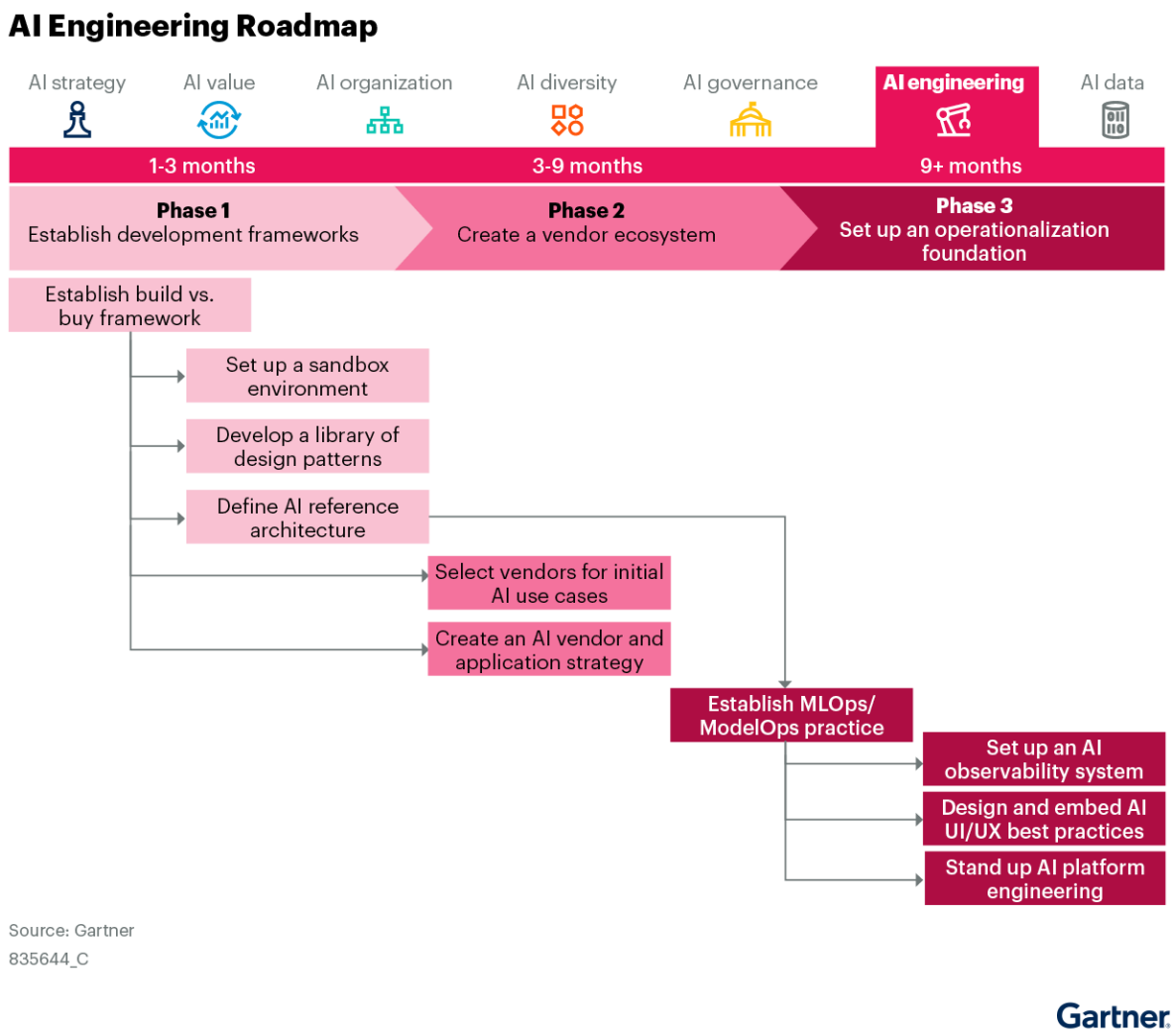
Build Cross-Function AI Platform Engineering Capabilities

Back to top | Back to planning considerations list

I&O teams commonly abandon data and AI teams to “You Build It, You Run It” (YBI YRI) operations. However, this approach rarely addresses the total operational needs of AI or AI agents in mission-critical production applications. Collaborate with software engineering, data and AI teams to ensure that AI-related operations are secure, reliable and scalable.

This collaboration is critical and should focus on developing a nuanced and shared understanding of the operational and logical constraints of the system. Figure 7 illustrates various steps in introducing AI platform engineering as one of the key end-goal functions that support AI maturity. For more on this, see the Gartner AI Maturity Model.

Figure 7: AI Engineering Roadmap



Gartner 2025

AI platform engineering will be crucial to scale AI implementation within the organization. These functions will support diverse AI product and development teams, including data scientists. Building shared services to support diverse development efforts aligns well with a platform engineering (PE) approach. Although many argue about where to place such a PE team, the team's reporting structure is less important than its mission — improving the experience of AI engineers and broader teams adopting AI, and enabling them to innovate rapidly. Focus on this goal and continuously adapt your platform to keep pace with evolving AI capabilities and use cases.

Verizon implemented an AI platform called VEGAS, and was able to multiply its GenAI impact (see Case Study: Enable and Scale GenAI Experiments With Verizon's Platform Strategy). In terms of the impact of the shared platform, around half of the potential users for the VEGAS platform have been trained and have used about 70% of its features, demonstrating its applicability at scale.

AI will require infrastructure-as-code (IaC) and policy-as-code (PaC) approaches, because manual methods of managing state and configuration will not scale to the speed and capability of AI systems. If you have skills gaps, you might find that GenAI and AI assistants allow you to create and collaborate on as-code platforms. Moreover, complex AI agent systems at scale will likely be built on event-based, policy-based and self-contained infrastructure capabilities. These new self-service consumers — AI agents — will require well-described, easy-to-programmatically-call IT operations services, for example, via an Model Context Protocol (MCP) server connected to an existing IaC, observability or IT service management (ITSM platform) .

Define and formally document your role and responsibilities to contribute to AI platform engineering. Build capability investments that support AI resilience through observability, IaC and PaC as part of a strategy to create I&O specific AI platform investments. Complete the life cycle of requirements with formal practices for managing AI incidents before they happen.

Benefits

- An organization can derive tangible value by implementing an AI policy, maintaining an easy-to-use catalog of approved tools and datasets, and providing targeted training — even before investing in comprehensive AI platform engineering development.
- AI platform engineering improves the AI engineer and data engineer experience. It reduces cognitive load by facilitating the delivery of AI-enabled applications at scale, with appropriate security and resilience.
- AI platform engineering is a key element of operationalizing AI. It improves the success rate of AI initiatives by providing a curated set of technologies and tools, along with custom supporting documentation and engineers that offer guidance and support.

- A specialized AI platform engineering team can develop deeper and broader experience with AI-related services in the public cloud than the organization's general cloud platform operations team (or other cloud team). This specialization allows a better understanding of AI scientists' needs and enables more tailored capabilities.

Cautions

- Ensure that your efforts add value to the AI engineer's experience. Data and AI teams often consider themselves self-sufficient, especially when using public cloud services. They likely have greater expertise with AI cloud services than I&O engineers, including cloud engineers.
- AI platform engineering capabilities often tend to neglect operational requirements. An AI platform engineering function may already be underway, developed by MLOps, software engineering and data teams. However, if I&O is left out, you can expect your organization to underinvest in resilience, operations, observability, automation and FinOps capabilities, or lack a good plan to scale these parts of the function.
- You may lack colleagues with the skills needed to staff the I&O domains of an AI platform engineering function. Moreover, AI engineers may need training and guidance in DevOps processes and tools, especially if they do not have prior experience in agile software engineering.

Related Research

Actionable, objective insights

Position your IT organization for success. Explore these additional, complementary resources and tools for I&O leaders and their teams:

Planning Guide

Adapt IT Governance to Meet the Challenges of Cloud Computing

Explore five-step solution path for IT governance in the cloud computing era.

Download Now

Insights

Infrastructure and Operations Solutions for Technical Professionals

Access action plans for the key trends for cloud, data center and edge infrastructure.

Learn More

Webinar

How to Identify and Implement Security Automation Use Cases

Understand how security automation can revolutionize your cybersecurity operations.

Watch Now

Tool

Gartner Infrastructure and IT Operations Initiative Accelerators

Achieve your strategic and innovation goals faster.

Learn More

Already a client?
Get access to even more resources in your client portal. Log In

Connect with us

Get actionable, objective business and technology insights that drive smarter decisions and stronger performance on your mission-critical priorities. Contact us to become a client:

U.S.: 1 855 322 5484

International: +44 (0) 3300 296 946

Become a Client

Learn more about Gartner for Technical Professionals
gartner.com/en/gartner-for-technical-professionals

Stay connected to the latest insights

Attend a Gartner conference
View Conference

Gartner®