Predicts 2024: Generative AI Will Transform IT Infrastructure and Operations

Published 1 December 2023 - ID G00801167 - 20 min read

By Analyst(s): Cameron Haight, Chris Matchett, Jonathan Forest, Andrew Lerner, Tom Cipolla, Chris Saunderson, Colin Fletcher, Pankaj Prasad

Initiatives: I&O Organizational Strategy

Generative AI will increasingly have a profound impact on IT infrastructure and operations. To prepare, it is crucial for I&O leaders to assess not only the benefits, but to also take stock of the continuing challenges to ensure that its implementation can be performed in a safe and controlled manner.

Overview

Key Findings

- Generative AI (GenAI) accelerates the learning curve for less senior team members and empowers senior team members with advanced insights and automation capabilities, but it is critical to not view GenAI as a replacement for human expertise.
- Most network teams are generally risk averse and do not initially trust recommended actions from AI tools.
- Virtual support agents and other AI-generated advisory features in IT service management (ITSM) platforms need internal enterprise knowledge and documentation to provide nontrivial value.
- GenAI hype places unrealistic expectations for the scope and scale of automation development improvements and will have to be tempered by negative experiences.
- While an assistant-based approach to GenAI adds value today, longer term, the value found from GenAI within I&O products will be based upon how they change the work of I&O.

Recommendations

- Improve digital employee experience (DEX) and increase operational efficiency by aggressively selecting and implementing GenAI-powered digital workplace management tools, preferring SaaS tools over on-premises tools due to their ability for rapid feature implementation by vendors.
- Start small and iterate with GenAI solutions by testing functionality in a proof of concept (POC) to validate capabilities before moving to production. Once in production, utilize recommendations and predictions, and then graduate to unattended automation over time as efficacy and trust improve.
- Adopt a knowledge management documentation policy and format standard (such as KCS) to ensure that AI solutions can leverage informational context through metadata.
- Automation developers should adopt a “fail-safe” approach to their development practices, ensuring that failures in their automation fail in a controlled fashion.
- Evaluate I&O-focused vendors in terms of their GenAI capabilities and their longer term product roadmaps.

Strategic Planning Assumptions

By 2028, GenAI technology will be used for 35% of network configuration and troubleshooting activities, up from near zero in 2023.

By 2027, GenAI will create more IT support and knowledge-based articles than humans will.

By 2027, more than 40% of digital workplace operational activities will be performed using management tools that are enhanced by GenAI, dramatically reducing the labor required.

By 2026, more than 90% of IT operations management vendors will have embedded GenAI capabilities in their products and/or services, up from less than 5% in 2023.

By 2026, 90% of organizations will suffer more than 10 production-impacting events annually due to insufficient GenAI skills and testing investments.

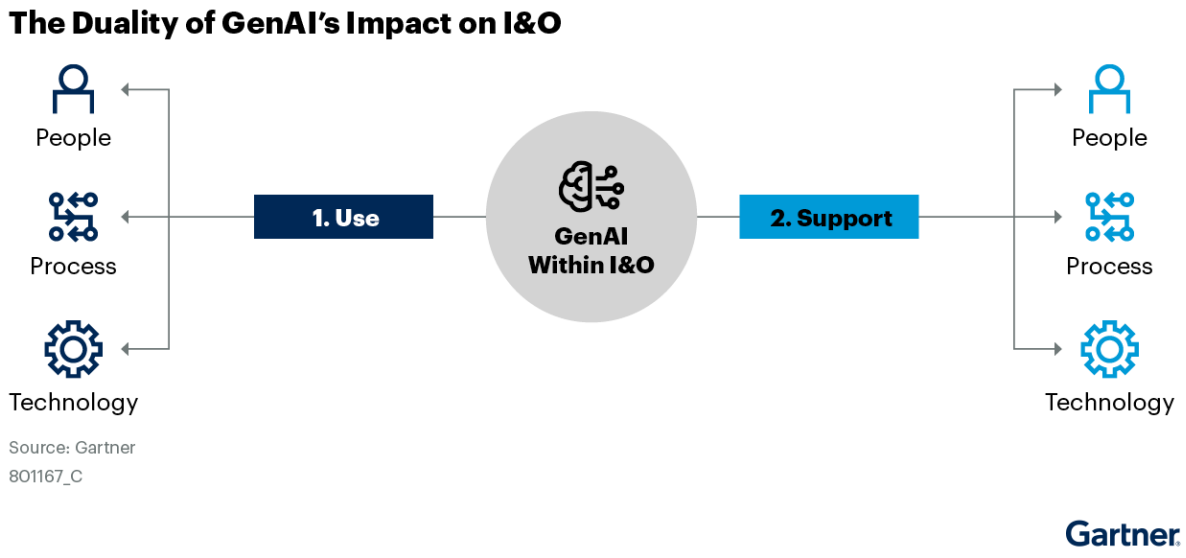
Analysis

What You Need to Know

Almost immediately after the availability of GPT 3.5 from OpenAI, interest in its potential application within IT has been nothing short of monumental. In the area of IT infrastructure and operations, the initial excitement has been tempered as evidenced by a recent survey (see Executive Pulse: GenAI Initiatives Take Shape Across the Enterprise). However, as I&O organizations continue to grapple with skills availability, cost optimization and the need to improve their business alignment, adoption (primarily in the form of augmented vendor tooling) is expected to grow.

In the context of GenAI, I&O occupies an interesting position (see Figure 1). Furthermore, I&O, as with other IT organizations, will be a consumer of GenAI-based capabilities to enhance its own processes and decision-making activities. However, it must also be prepared to support enterprise initiatives involving the development of proprietary large language models (LLMs) from an infrastructure and process perspective, regardless of whether the models are developed (and run) on-premises or within the cloud.

Figure 1: The Duality of GenAI's Impact on I&O



I&O-focused vendors see the opportunity with GenAI, and there has been a steady stream of announcements about assistive or copilotlike capabilities in observability, service desk, experience and configuration management (infrastructure-as-code). The use cases that seem to be of the greatest interest from Gartner client conversations so far include:

- **Conversational UIs:** Provides a natural language "query assistant" to pose questions (i.e., what devices have had the most errors), thereby removing the need to learn a product's domain-specific query language. Virtual assistants (chatbots) are also in this category, and while there is interest, they will require advanced skills to fully exploit the capabilities that GenAI promises.
- **Content generation:** This category covers a wide range of digital asset creation, from summarization of a Slack channel conversation to the creation of infrastructure-as-code artifacts.
- **Knowledge discovery: Provides,** for example, an error message and asking for an interpretation with potential remediation steps.

For I&O leaders, the opportunity that GenAI represents to complement existing roles and drive scale and consistency in its use outweighs the risks. I&O's role in protecting their organization must be balanced with the rapid developments in this space. The key is to work with the other groups within IT (i.e., HR, compliance, security) to collaboratively create the policies and structural precautions that limit the exposure of the organization to the effects from inadequately tested code or insufficiently validated guidance (see Note 1 for additional information on GenAI).

Strategic Planning Assumptions

Strategic Planning Assumption: By 2028, GenAI technology will be used for 35% of network configuration and troubleshooting activities, up from near zero in 2023.

Analysis by: Jonathan Forest and Andrew Lerner

Key Findings:

- Most network teams are generally risk-averse and do not initially trust recommended actions from AI tools.
- Despite a massive amount of hype and marketing associated with AI (including GenAI) networking, adoption of these capabilities has been limited to date.
- GenAI offers great potential to enable improvements to long-standing traditional networking operations practices across Day 0, Day 1 and Day 2.

Market Implications:

AI networking, which is inclusive of GenAI, uses AI and machine learning (ML) to deliver granular and specific actionable network insights to enhance documentation, configuration management and deliver efficient operations. AI networking can be a feature within a network vendor's management platform, a stand-alone multivendor platform (i.e., a solution that supports multiple networking vendors), which is part of an AIOps platform or delivered as part of a managed network service.

Today, AI networking primarily delivers Day 2 network insights and offers recommendations to accelerate incident resolution and prevent outages and trouble tickets. Going forward, we see GenAI further enhancing Day 2 operations and also supporting Day 0 and Day 1 functions in terms of network design and setup as well as dynamic recommendations to optimize network performance.

GenAI can create detailed configurations and troubleshooting procedures based on human inputs without explicit templates. We believe that GenAI-based AI networking can drive operational management and network setup time savings by up to 25% by driving efficiencies that can't reasonably be achieved by scaling manual resources. It simplifies network setup and operations so that network personnel won't need deep configuration and troubleshooting skills to manage the network.

AI networking will automate tasks without requiring deep network knowledge of network personnel. By leveraging GenAI such as ChatGPT, Bing Chat, Google Bard, etc., organizations will move from automating defined workflows and templating that required manual human setup to automating without requiring humans to do the initial work.

Vendors will first leverage GenAI for documentation purposes, offering enterprises a new mechanism to engage with support documentation and knowledge bases. This will be interactive and conversational. Vendors will also leverage GenAI (and broader AI) to create troubleshooting "digital assistants" or "co-pilots," which will be added to existing network management consoles and which customers can engage with from a conversational and interactive basis. Some of these capabilities exist in a limited fashion today (e.g., Juniper Marvis). However, we believe these will become much wider spread with enhancements to:

- Breadth — These will be standard features in all major vendors' management platforms.
- Depth — These will address more network design, implementation and operate activities than they do today.
- Easier — These tools currently require some level of knowledge to operate (i.e., a device name or protocols), but in the future they will allow for broader, unstructured questions.

Recommendations:

- Start small and iterate with GenAI solutions by testing functionality in a POC to validate capabilities and verify recommendations (e.g., testing in a digital twin environment) before moving to production. Once in production, verify recommendations and predictions, and then graduate to more unattended automation over time as efficacy and trust improves.
- Justify adopting GenAI solutions by quantifying the cost savings, enhanced agility, increased resource efficiency and end-user performance improvement.
- Minimize reliance on service desk for Day 2 operations by leveraging GenAI's conversation interface capabilities to simplify troubleshooting and drive operational efficiencies.
- Accelerate Day 0 network design and Day 1 network configuration by leveraging GenAI through an iterative process to refine the outputs to accelerate bringing networks to production.

Related Research:

Innovation Insight: AI Networking Has the Potential to Revolutionize Network Operations

Hype Cycle for Enterprise Networking, 2023

Strategic Planning Assumption: By 2027, GenAI will create more IT support and knowledge-based articles than humans will.

Analysis by: Chris Matchett

Key Findings:

- Virtual support agents and other AI-generated advisory features in ITSM platforms need internal enterprise knowledge and documentation to provide nontrivial value.
- I&O leaders and IT operations management (ITOM) vendors are investigating approaches to incorporating internal enterprise support knowledge with public LLMs.
- Vendors are starting to offer private LLMs trained for IT support in addition to public LLM access, but newer solutions will use prompt engineering with retrieval augmented generation (RAG).

Market Implications:

Virtual support agents and other AI-generated advisory features in ITSM platforms need internal enterprise knowledge and documentation to provide nontrivial value. The benefits are not limited to the advice and action stages of applied AI in ITSM. IT experts perform better when provided with knowledge and documentation through faster and impactful response and decision making. The problem is that the knowledge base is typically lacking in content because developing the content is very demanding of time and resources. In the past, I&O leaders with the budget to spare recruited technical writers to populate the knowledge base or leveraged external knowledge to fill some common gaps.

Knowledge creation is a use case from the content generation AI use-case family that promises significant value to I&O leaders. It recognizes attempted solutions and successful solutions from sources, including organic conversations and historical cases, and creates associated knowledge articles in a management system (e.g., ITSM platform).

This use case expands the content within a knowledge base while mitigating the time and cost of authoring knowledge articles manually. This improves the employee-facing capabilities of knowledge discovery and powers greater automation as virtual support agents use these generated solutions to provide support. It also provides more contextual data for case clustering to allow ITSM platforms to provide better advice more frequently.

Knowledge creation is an obvious use of existing GenAI technology, but product availability is very limited. Solutions are entering the market now. Otherwise, most offerings focus on rewording knowledge from documents rather than generating new solutions. Automation of knowledge creation is heavily dependent on access to information, including unstructured conversational data, which must be extracted and then processed and summarized using cluster analysis and natural language technologies. Natural language understanding capabilities enabled by GenAI are improving rapidly and are a key focus for vendors across multiple marketplaces, not just within ITOM.

The clear requirement for AI-actioned knowledge generation to relieve resource-starved I&O leaders, coupled with the keen interest of IT vendors to address this demand, will lead to many new products being offered over the next two years. I&O leaders will pilot and fully implement these solutions as a priority, despite the steep learning curve that these new technologies impose. It is necessary to learn the jargon and fundamentals of how GenAI works, even when outsourcing the design. The technology is evolving rapidly, requiring constant attention to market developments.

Most IT support knowledge is generic and common to most organizations. External knowledge discovery will quickly replace the need to manually write knowledge-based articles for common business applications and platforms. Furthermore, RAG (a prompt engineering technique that incorporates private and proprietary enterprise information into public LLM request sessions) and knowledge creation will generate knowledge that is tailored to the enterprise's internal systems. By 2026, GenAI will write more IT support and knowledge-based articles than humans will.

Recommendations:

- Adopt a knowledge management documentation policy and format standard (such as knowledge-centered service [KCS]) to ensure that AI solutions can leverage informational context through metadata.
- I&O leaders will need to develop their knowledge of prompt engineering to effectively manage vendors and service providers of LLMs.

- Guard against hallucinations by labeling AI-generated documentation as unverified and train employees within and outside of IT to identify and respond to inaccuracies.

Related Research:

Use-Case Prism: Artificial Intelligence for IT Service Desk

Quick Answer: How to Use Virtual Support Agents With Peer IT Support

AI Design Patterns for Knowledge Graphs and Generative AI

Strategic Planning Assumption: By 2027, more than 40% of digital workplace operational activities will be performed using management tools that are enhanced by GenAI, dramatically reducing the labor required.

Analysis by: Tom Cipolla

Key Findings:

- Most digital workplace technology management tools — UEM, DEX, ITSM, ITAM, VDI and UCC — are quickly gaining GenAI capabilities to simplify product usage, generate powerful insights and increase operational efficiency.
- GenAI capabilities can shorten the learning curve and upskill digital workplace engineers and administrators by enhancing their abilities. GenAI will not replace human experience, intuition, creativity, empathy and problem solving.
- Digital workplace teams will also support the deployment of GenAI capabilities and applications to employee endpoints, which will require the use of rapid release cycles and the latest and greatest versions. This will challenge existing update and patching cycles and capabilities.

Market Implications:

Digital workplace technology management tool vendors are quickly capitalizing on the excitement surrounding GenAI by adding features and establishing new offerings. Motivated by the opportunity for increased revenue, many are adding new licensing tiers to their existing products. GenAI promises to simplify the operations, improve intelligence-driven insights, accelerate automation, reduce costs and significantly improve the DEX. It is not a question of if, but when GenAI features will be added to nearly all products; however, that will come at a cost.

Although the potential benefits of GenAI are numerous, it cannot replace human expertise. Instead, GenAI should be viewed as a copilot that enhances digital workplace staff who will train the models and confirm that the answer, action or created artifacts are accurate and that they do not introduce any unintended consequences.

For example, a DEX tool may promote significant performance improvements by disabling a service on the endpoint that AI does not realize is required for security or regulatory purposes. An inexperienced or hasty engineer who approves or implements the recommendation will introduce significant risk. Other common uses of GenAI in the digital workplace include analyzing and synthesizing data, creating a script or drafting communications to employees about the new features of an upcoming software upgrade. All of this requires validation and refinement before use.

GenAI features addition within various productivity applications requires much faster release cycles than what is traditionally used in enterprises. For example, Microsoft 365 Copilot requires that all endpoints subscribe to monthly updates, but most organizations subscribe to semiannual updates to minimize change fatigue and surprises.

Gartner expects the pace to increase as vendors adopt agile development and release approaches so they can rapidly introduce new capabilities, fixes and value. Most organizations struggle to keep pace with annual operating system updates and monthly web browser updates, so the required acceleration will exacerbate the problem and magnify outdated endpoint management approaches built for much slower cycles.

To avoid impeding the organization's GenAI ambitions, it is critical to optimize application update and release processes through automation and by adopting best practices. These include ring-based approaches, third-party-patching tools and streamlined or automated testing strategies (see Accelerate Windows and Third-Party Application Patching).

Recommendations:

- Improve DEX and increase operational efficiency by leveraging GenAI capabilities that are added to existing digital workplace technology management tools or migrating to GenAI-powered ones.
- Prioritize SaaS over on-premises offerings, which are less likely to deliver GenAI capabilities and will not have sufficient aggregated data to use to build their LLM.
- Avoid the operational disruptions by using GenAI to augment existing staff and embracing a human-in-the-loop approach to validate generated answers, recommendations, actions and artifacts.
- Prepare for the accelerated pace and volume of software updates by embracing modern update and patching approaches.

Related Research:

How will Generative AI Impact Digital Workplace I&O?

Accelerate Windows and Third-Party Application Patching

Strategic Planning Assumption: By 2026, more than 90% of IT operations management vendors will have embedded GenAI capabilities in their products and/or services, up from less than 5% in 2023.

Analysis by: Cameron Haight and Colin Fletcher

Key Findings:

- While I&O-focused vendors rush GenAI capabilities to market, I&O professional acceptance of the technology has lagged behind that of other IT areas. However, this is expected to change within the next year as the technology matures, common regulatory of security certifications are achieved and privacy concerns are addressed.
- GenAI can be applied to a large number of potential use cases within I&O products. However, most implementations today are in the form of an “assistant” designed to augment human capabilities in areas such as searching, summarization and content development.

- A potential stumbling block to accelerated adoption of I&O-centric GenAI is the pricing for this functionality as some vendors already bundle these capabilities in existing products. More currently, these capabilities are offered for free while in preview mode, and most of I&O has not yet assessed a value to these capabilities.
- While an assistant-based approach to GenAI adds value today, longer term, the value found from GenAI within I&O products will be based upon how they change the work of I&O.

Market Implications:

The I&O technology industry is changing very rapidly due to the growing impact of GenAI. While enterprise I&O organizational uptake to date has been modest, expectations are that GenAI will become a part of the I&O professional's future toolbox, whether in a stand-alone manner or embedded within an technology provider's product. I&O vendors see the same future and want in on the potential "gold rush" because there are nearly daily occurrences of the announcement of GenAI capabilities being offered across the I&O technology stack.

Today, I&O vendor capabilities focus primarily upon "assistant" use cases involving conversational UIs, summarization and content development while more advanced features involving automation and even autonomous functionality will have to earn the trust and confidence of I&O professionals and leaders.

In the longer term, I&O organizations will benefit from a functionality and a cost standpoint as I&O vendors race to stay ahead of an influx of commoditization in which assistant-style technology will be found in almost all technology and service provider offerings. I&O organizations should expect the development of domain- and vendor-specific LLMs by I&O vendors and look for opportunities to leverage GenAI beyond mere role or process augmentation.

Recommendations:

- Work with other organizations within IT to establish governance, valuation and other methodologies to optimize future GenAI technological implementations.
- Pilot GenAI-powered assistants with an eye toward incremental rollouts that address current I&O functional and cognitive demands.

- Inventory GenAI capabilities in current and planned I&O tooling to help maximize the use of already-included and free preview capabilities in service of developing skills and future investment plans.
- Evaluate I&O-focused vendors in terms of their GenAI capabilities, costs and longer term product roadmaps.

Related Research:

How to Pilot Generative AI

Assess the Value and Cost of Generative AI With New Investment Criteria

How to Use Anomaly Detection to Implement AI-Driven Use Cases

Strategic Planning Assumption: By 2026, 90% of organizations will suffer more than 10 production impacting events annually due to insufficient GenAI skills and testing investments.

Analysis by: Chris Saunderson and Pankaj Prasad

Key Findings:

- Automation engineering using GenAI is seeing broad experimentation within client organizations and rapid product feature adoption by vendors.
- GenAI hype will place unrealistic expectations for the scope and scale of automation development improvements, and it will be tempered by negative experiences.
- Without sufficient review and validation by the I&O teams, production impacts will be felt due to incorrect, misinformed or biased results in GenAI output.
- Experimentation and pilot activities occurring today will expose existing deficits in skills, testing and release mechanisms that result in outages that are magnified by GenAI.

Market Implications:

Automation development capacity has long been a bottleneck for I&O teams because the cycle of scoping/ramping skills and technology; development and testing; introduction of automation; and then ongoing maintenance of automated capabilities has elongated time to value.

The promise of GenAI accelerating automation outcomes is tempting as a solution to the constraints organizations feel. The equilibrium between accelerated development and being able to trust what results are returned is a delicate balance to maintain. Accordingly, we expect that adopters that complement their automation development activities with GenAI will suffer some missteps, resulting in production-impacting events.

The reaction to these events will take a usual path of requiring additional checks, balances and gates for introduction of automation that is sourced from GenAI, but these will ultimately be unproductive. The necessary investments are:

- Upskilling automation developers to take advantage of the efficiency that GenAI offers and enabling them to build the needed knowledge and understanding to make the best use of this opportunity, such as the ability to leverage prompt engineering.
- Adopting peer and code reviews into the development pipeline and focusing those reviews on the entire target task or process, rather than the structure, syntax and format of the automation.
- Implementing an “automation bill of materials” that identifies the sources of the automation to distinguish between human and GenAI heritage to aid in root cause analysis of problems and improvements in the training of private or public models.
- Standardizing processes and identifying where GenAI can offer the most value in future standardization and automation efforts.

For vendors that are integrating GenAI into their automation tools, these actions are also applicable.

IT organizations should account for a human-in-the-loop for validating the correct sequence of actions created by GenAI for addressing any automation task. GenAI can also be leveraged to identify distinct stages in an automated process that lack embedded telemetry (in the form of traps and specific error codes) to help assess the health of the automation and identify the point of failure.

Recommendations:

- Focus GenAI introduction on complementing the development of automation capabilities, rather than replacing developers.
- Protect the use of GenAI in automation development by implementing peer and code reviews that are focused beyond the functional execution of the task or process that is being automated.
- Automation developers should adopt a “fail-safe” approach to their development practices, ensuring that failures in their automation are handled in a controlled fashion during development and reduced in production.
- Leveraging GenAI for automated test case development automation should be carefully adopted, taking advantage of the guard-rails implemented around the automation development process.
- Reinvest freed-up time in helping staff to build mastery of the skills and tools they need to drive more robust automation, rather than increasing the scope of automation without keeping foundational skills refreshed.

Related Research:

Quick Answer: What 3 Actions Should I&O Leaders Take Now On ChatGPT?

How to Pilot Generative AI

How to Create an Ambitious, Achievable Generative AI Strategy

A Look Back

In response to your requests, we are taking a look back at some key predictions from previous years. We have intentionally selected predictions from opposite ends of the scale — one where we were wholly or largely on target, as well as one we missed.

This report is too new to have on-target or missed predictions.

Note 1: Additional Information

For additional information on GenAI, see:

- C Egersdoerfer, D Zhang, D Dai, " Early Exploration of Using ChatGPT for Log-based Anomaly Detection on Parallel File Systems Logs," HPDC '23: Proceedings of the 32nd International Symposium on High-Performance Parallel and Distributed Computing, 2023.
- GG González, P Casas and A Fernandez, " Fake it till you Detect it: Continual Anomaly Detection in Multivariate Time-Series using Generative AI" (link downloads PDF), 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), 2023.
- Y Chen, H Xie, M Ma, Y Kang, X Gao, L Shi, Y Gao, H Fan, M Wen, J Zeng, S Ghosh, X Zhang, Q Lin, S Rajmohan, D Zhang, T Xu, " Automatic Root Cause Analysis via Large Language Models for Cloud Incidents," 2023.
- Transforming Customer Support with Generative AI, DevRev.
- S Pujar, L Buratti, X Guo, N Dupuis, B Lewis, S Suneja, A Sood, G Nalawade, M Jones, A Morari and R Puri, " Automated Code Generation for Information Technology Tasks in YAML Through Large Language Models," 2023.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Executive Pulse: GenAI Initiatives Take Shape Across the Enterprise

© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner's Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity." Gartner research may not be used as input into or for the training or development of generative artificial intelligence, machine learning, algorithms, software, or related technologies.