451 Research Discovery Report

May 2025

Securing AI at scale

Lessons from industry leaders

S&P Global
Market Intelligence

in partnership with
verizon business

Table of contents

Introduction

While AI adoption continues at a fever pitch, organizations are experiencing a variety of headwinds that prevent them from realizing its full value. According to 451 Research's Voice of the Enterprise: AI & Machine Learning, Infrastructure 2024 survey, security issues present a major technological barrier inhibiting organizations' AI efforts.

Improvements are needed, not only in the tools enterprises use to help secure their environments, but also in organization-wide awareness and training to develop core competencies for AI security.

This report is the second in a three-part series about AI at scale. Specifically, it looks at security as a foundational pillar of successful enterprise-scale AI and includes insights gleaned from an extensive set of interviews with leaders in the space, as identified by 451 Research. The first report in this series discusses networking considerations, and the third examines overarching best practices for implementing AI at scale.

Methodology

The insights presented in this report are derived from comprehensive interviews and peer group discussions with senior executives responsible for AI initiatives at organizations that have successfully implemented AI at scale. These interactions included 15 in-depth interviews with AI and IT decision-makers as well as an executive discussion board with 30 participants. All participants in the study were specifically involved in managing and implementing their organizations' IT infrastructure for AI workloads. Participants were based in the US, UK, Singapore, Australia, Sweden, Germany, Denmark and Japan. This report was developed by S&P Global Market Intelligence and was commissioned by Verizon.

Figure 1: Demographic breakdown of participants

Executive discussion board				One-to-one interviews			
Region	**#**	**Industry**	**#**	**Region**	**#**	**Industry**	**#**
North America	10	Healthcare/life sciences	8	North America	5	Retail/wholesale	3
Europe	10	Financial services	8	Europe	5	Healthcare/life sciences	4
Asia-Pacific	10	Manufacturing	6	Asia-Pacific	5	Financial services	5
		Retail/wholesale	3			Manufacturing	2
		Utilities	1			Utilities	1
		Other	4				

Source: S&P Global Market Intelligence 451 Research AI at Scale study commissioned by Verizon.

Executive summary

The results of this study align with other 451 Research findings in identifying security as a top challenge for AI initiatives. Based on an online executive discussion board with 30 participants and supplemented by 15 live, in-depth interviews, we have extracted four key insights relevant to AI decision-makers, strategists and end users as they create, implement and scale their technical capabilities. These include:

- **Insight 1:** Security for AI should span multiple environments and be designed to scale
- **Insight 2:** Additional security tools and capabilities are sorely needed
- **Insight 3:** Security tools and strategies should support diverse governance requirements
- **Insight 4:** AI models themselves must be secured and monitored, in addition to the environments where they run

We provide further findings and analysis into each of these insights in the sections that follow.

Insight 1: Security for AI should span multiple environments and be designed to scale

Data has long been regarded as a critical organizational asset. As the raw material for rich and unique insights, data provides differentiation and immense business value, and it must be protected accordingly. In fact, respondents to 451 Research's Voice of the Enterprise: Digital Pulse, IT Budgets & Drivers 2025 survey identified data security as the top area for organizational IT investment in 2025.

Data is also the foundation of AI — a fact that further complicates enterprise security efforts. "AI at scale" interview and executive discussion board participants noted significant challenges integrating AI workloads into existing IT and operations technology environments — balancing new AI efforts with ongoing mission- and business-critical objectives, all while keeping the underlying models and data secure.

Organizations generate, prepare, process and store data across diverse locations including on-premises datacenters, colocation facilities, the public cloud and the edge, as well as hybrid combinations of all the above, and each organization has unique preferences for where it undertakes a given activity.

One thing organizations hold in common, however, is the desire to create a single, unified AI security strategy based on a common set of tools and practices. That strategy should be crafted with an eye toward scalability and deployment flexibility as their organizations' AI capabilities progress.

"Many times, security is overlooked. This is a common mistake, not only for AI, but in general."

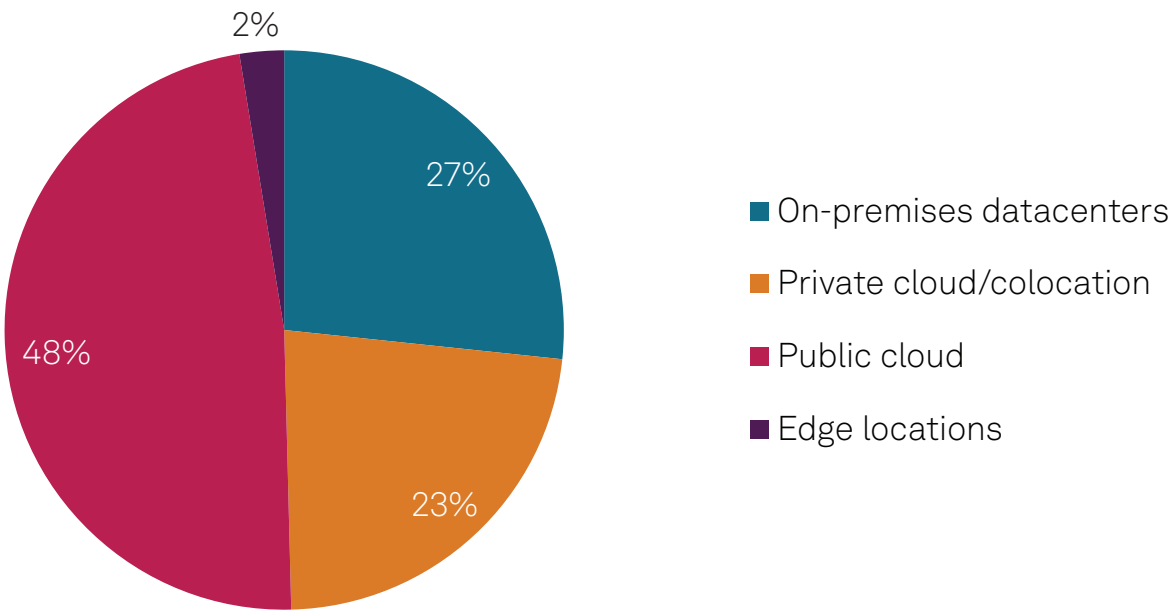
Head of architecture/cybersecurity
Banking, 20,000-50,000 employees, Singapore

"Security and compliance are most important to us. No matter how good the performance [of a model] is, it could be wasted if it's not secure."

Head of IT
Healthcare/life sciences, 5,000-10,000 employees, Japan

On average, study participants reported that 48% of their AI workloads are deployed in the public cloud, 27% in on-premises datacenters, and 23% in private clouds and colocation facilities, while AI workload deployments at the edge remain nascent (see Figure 2). Public cloud remains the most common platform for running AI workloads, with 83% of study participants using it to some extent.

Figure 2: Estimated percentage of AI workloads by execution platform



Q. What percentage of your AI/ML workloads are running in ...
Base: Executive discussion board participants (n=30).
Source: S&P Global Market Intelligence 451 Research AI at Scale study commissioned by Verizon.

There is no single standard or "right" place to deploy AI workloads and applications. In addition to security considerations, an organization may favor a particular venue due to factors such as cost, performance, reliability, flexibility, privacy and governance. For some, the public cloud may offer more expedient access to resources and services that support both AI training and inference workloads, or it may better align to higher-level organizational IT strategies. For others, on-premises and other self-managed or single-tenant environments may appeal based on higher levels of control of data, privacy, management and performance.

Among study participants, smaller organizations tended to deploy a higher percentage of their AI/ML workloads in the public cloud compared to larger enterprises. For example, on average, organizations with 1,000-5,000 employees said 71% of their AI/ML workloads run in the cloud, versus 43% for organizations with 10,000 employees or more. Additionally, EMEA-based organizations tended to deploy higher percentages of AI/ML workloads in the cloud than those in North America and Asia-Pacific.

"We are a global organization that is primarily using cloud-based services (IaaS, SaaS, PaaS), and we need the ability to make our systems available globally."

Senior director of IT infrastructure
Consumer goods, 10,000-20,000 employees, US

"Resiliency is really, really the key today in terms of infrastructure in parallel with risk and security risk, whether internal or external. We use a mix of cloud and a mix of on-prem to manage them depending on the needs."

Senior data architect
Financial services, 10,000-20,000 employees, Singapore

"We take a hybrid approach: AI training is carried out in the cloud, while real-time inference at the point of service is performed on-premises. This allows us to balance cost and service-level requirements effectively...for processing sensitive data – such as confidential or personal information that cannot be sent directly – we often rely on on-prem infrastructure to ensure data security."

Head of advanced tech lab
Retail, >50,000 employees, Japan

Integrating AI workloads into organizations' IT environments requires careful evaluation and may entail revision of broader data strategies. Exposing data to AI/ML models carries inherent risk, and data security policies should consider the entire data life cycle from creation to disposal, as well as the total ecosystem from datacenters to endpoints. While study participants primarily discussed AI workloads deployed in IT environments, some organizations also mentioned overlap with their OT environments.

"AI solutions pull our OT environment into the process. Historically, the plant systems (OT) weren't accessible to the outside world."

Senior director of ITOps and strategy
Chemicals manufacturing, 1,000-5,000 employees, US

Workload environments should be secure by design, with appropriate controls and management capabilities in place. This will be particularly important as organizational "Ops" landscapes evolve in support of AI and broaden to include IT, developers and data scientists who require access to resources and data assets. Growing AI capabilities and toolsets dramatically expand an organization's attack surface, exposing assets to a greater number of potential external and internal threat actors. This results in a more complex identity management environment. Looking ahead, organizations' security capabilities will need to adapt even further to support the increasing number of non-human actors (e.g., agents) in their environments.

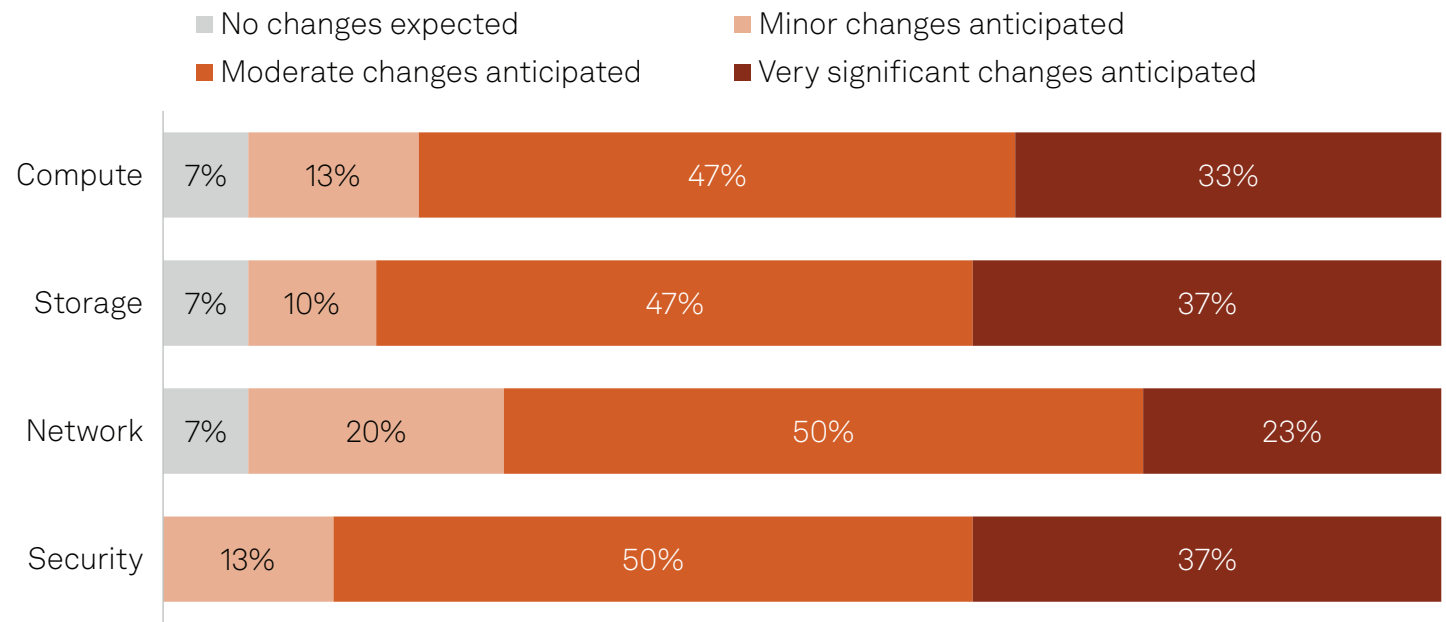
"Balancing performance with governance is critical. Fast experimentation is useful only if it's also secure and compliant."

Head of architecture/cybersecurity
Banking, 20,000-50,000 employees, Singapore

Insight 2: Additional security tools and capabilities are sorely needed

The study clearly indicates that organizations will need improved tools and practices to maintain the security of AI workloads across diverse environments. In fact, security is the top area where respondents indicated a need for investment over the next 12-24 months, with all 30 participants projecting some degree of required change (see Figure 3).

Figure 3: Anticipated infrastructure changes to support AI/ML use cases over the next 12-24 months



Q: Looking ahead over the next 12-24 months, what changes do you anticipate in your infrastructure that are being driven by existing or new AI/ML use cases?
Base: Executive discussion board participants (n=30).
Source: S&P Global Market Intelligence 451 Research AI at Scale study commissioned by Verizon.

When asked what they would do if they were to start their AI projects over again, one major commonality among respondents was the desire to adopt zero-trust frameworks for both humans and machines, with identity and access management (IAM) capabilities and role-based access controls (RBAC) to restrict access to model and data assets on an as-needed basis.

> **"IAM/PAM to secure user or device identity within infrastructure and pushing ZTA for both on-prem and public cloud-based solutions."**
>
> **Senior infrastructure/network security architect**
> Medical devices, 1,000-5,000 employees, Denmark

"Enforcing strict role-based access and monitoring to prevent unauthorized usage or insider threats."

Director of infrastructure and cloud
Gaming, 1,000-5,000 employees, Sweden

Additional sought-after capabilities and tools that respondents identified included enhanced privacy controls for data pipelines, end-to-end encryption, and improved monitoring and visibility of workloads throughout their life cycles. Improved threat detection tools were also noted as important, with the caveat that they optimally should not carry significant additional performance overhead.

Further, the adoption of cloud-native infrastructure for AI – particularly public cloud, which 83% of respondent organizations are using – introduces additional security challenges and expense. These challenges extend to the creation, movement and replication of data, and can be amplified when moving AI workloads between diverse and distributed IT environments.

"Data movement and backup costs on the public cloud are far more than on-premise[s]."

VP of AI architecture and infrastructure
Healthcare, 20,000-50,000 employees, US

Other cloud-native technologies — such as containers, APIs and cloud-native databases, which are generally favored by developers — are becoming the standard building blocks of the platforms that support AI training and applications. Securing these cloud-native resources call for a distinct set of vendors, tools and competencies such as extended detection and response (XDR) and cloud-native application protection platforms (CNAPP). Respondents also recommended container runtime isolation and micro-segmentation within networks. Many organizations will need time to build familiarity with these technologies and grow their associated technical capabilities, but our research clearly indicates that change is needed.

Like many other technologies, AI tools contain a certain amount of open-source code, including in the operation of large language models (LLMs). While the benefits of open-source development are well-documented — in many cases, actually leading to improved security outcomes — open source also entails a certain amount of risk and complication. For example, regulatory frameworks increasingly require that organizations prove the security of their software supply chains, which can be more challenging when adopting code from various open-source communities. Such challenges will likely persist — and proliferate — as organizations increasingly develop and deploy AI applications that use LLMs.

To help address this challenge, one respondent noted a practice of scanning organizational data for risk factors such as personally identifiable information before using that data with open-source models.

"We have a robust security layer on top, which skims for PII information, [and] any security risks when we use some of these open-source models … When it comes to generative AI, I think right now, we have like two to three open-source models that we use … We are using them out of the box, things like Llama and things like that, which is, again, on-prem and the same infrastructure that our classical ML models are hosted on."

Senior director of AI products and platforms
Financial services, 20,000-50,000 employees, US

Respondents also expressed concerns about the safety of prompts themselves and ensuring LLMs do not inadvertently expose sensitive or proprietary data. For many, this becomes more worrisome as deployments scale up, with systems characterized by increasing autonomy and accessed by an ever-growing number of users — potentially including AI agents.

"Sensitive data being exposed through prompts, because once this is controlled by AI, there are some unknown factors. I feel a need to strengthen security there, things AI is exposing through its prompts. In the same sense, sensitive data being exposed by AI implementation through prompts. Data uploads would be second. The worry is, with AI, I don't know where or what the nature of threats could be. Still some mysteries on what the potential for sensitive data being exposed is."

CIO
Wholesale/retail, 5,000-10,000 employees, Japan

Insight 3: Security tools and strategies should support diverse governance requirements

Any organization that handles data is subject to some level of data governance. Particular requirements are typically determined by where data is located, what type of information it contains and what industry an organization operates in. These policies often include mandates regarding factors such as access, accuracy, privacy and transparency. Notably, the number of data governance regulations has increased in recent years, and requirements have become more stringent, largely in response to ongoing technical advancements. In line with this trend, AI tools and technologies are increasingly included in governance policies and legal regulations.

Understandably, many organizations are struggling to strike a balance between fostering innovation and ensuring compliance with complex data governance frameworks. Study results suggest this is even more challenging for those that operate in multiple countries and for those with more distributed technical and infrastructural landscapes.

Governments and industry groups continue to introduce data-protection regulations to address evolving risks. Examples include:

- Australia's Privacy Act
- California Consumer Privacy Act (CCPA)
- EU Data Act
- EU Digital Operational Resilience Act (DORA)
- EU General Data Protection Regulation (GDPR)
- Japan's Act on the Protection of Personal Information (APPI)
- Payment Card Industry Data Security Standard (PCI-DSS)
- US Health Insurance Portability and Accountability Act (HIPAA)
- US Sarbanes-Oxley Act (SOX)

"The biggest challenge is securing sensitive data across hybrid environments while maintaining compliance with regulatory body requirements across multiple countries and regions."

Head of architecture/cybersecurity
Banking, 20,000-50,000 employees, Singapore

"Security requirements are extremely high as we work with confidential patient data, and there are significant costs involved in working with the regulatory agencies and maintaining private cloud environments."

Assistant CTO
Healthcare, 5,000-10,000 employees, Singapore

By prioritizing strong security principles and investing in security-supporting toolsets, organizations can improve their levels of compliance. Employees and contractors should receive regular training on contemporary security threats and best practices. Organizations would be wise to broadly implement data loss prevention (DLP) and zero-trust principles, with specific focus on IAM and RBAC, as well as enforcing the principle of least privilege regarding access to models and data. Similarly, models should combine DLP with appropriate levels of monitoring and observability to ensure their integrity and help verify compliance. Study participants also recommended improving API security and encrypting data, ideally both at rest and in flight.

"Robust encryption for regulation, zero-trust access controls, and API security continues improvement to industry standards."

Director of infrastructure and cloud
Gaming, 1,000-5,000 employees, Sweden

To improve compliance and protect sensitive enterprise data, respondents frequently recommended implementing stronger organization-wide data governance frameworks and improving knowledge-sharing regarding data ownership. This also mitigates the risk of data leaks from "shadow" AI, whether intentional or unintentional.

"Usually, what takes most time is if you want to onboard new data sources is to find someone in the company who wants to act as an owner of it, so we have this governance straight … [Getting new data owners] can take a little bit of while. And here, we are now putting a data governance community in place with the goal to get more data stewards and data owners on board."

Global head of data and AI architecture
Healthcare/life sciences, 50,000-100,000 employees, Germany

"So if I were to start now, I'd have to make sure that my data is fair at birth with a proper data governance framework … And also, it's got to be a proper access framework, meaning that data will have — within the governance framework — we'll have roles and responsibilities of who owns the data and who enables the data access … It's linked to roles and responsibilities."

Director of data strategy, data science and AI
Life sciences, 50,000-100,000 employees, UK

Data breaches entail significant costs. From a regulatory perspective, this can include substantial fines as well as administrative penalties. Longer-term consequences of a data breach may include increased oversight and additional audits that can slow innovation and production. Further external consequences can include lasting reputational damage among current and potential customers, which may translate to lost commercial opportunities.

"Due to a significant incident in the past where our company faced heavy criticism from the media and society, we are much stricter about security compared to other companies. Even when considering new and innovative technologies, we often face opposition regarding security and compliance concerns, which imposes significant constraints on the use of AI. As a result, there are many instances where we miss business opportunities."

IT strategy director
Electronics manufacturing, >50,000 employees, Japan

When participants were asked what they would do differently if they could start their AI initiatives over, many said they would sharpen their focus on data governance and protection. Many said they underestimated — or in some cases, failed to consider — the work required to ensure compliance as they grew their AI projects to scale, especially across multiple geographies. Regarding the balance between expectations for innovation and requirements for security and compliance, one respondent put it succinctly:

"Expectations are high, but regulatory requirements are higher."

Head of IT strategy and transformation
Banking, 500-1,000 employees, Germany

Insight 4: AI models themselves must be secured and monitored, in addition to the environments where they run

For many organizations, AI initiatives represent forays into uncharted territory fraught with challenges. Skillsets are scarce, initiatives are time-consuming to develop and implement, and projects are costly to maintain. Organizations need to protect the value of their investments while also protecting their data and other intellectual property that may be used to build, train and augment models. These proprietary inputs translate to competitive differentiation in AI outcomes.

Notably, these outcomes are not only the product of significant AI-related investments — they represent the distillation of critical organizational data. Accordingly, both models and data sources are valuable targets for attackers, and they come with a new set of risks not encountered with traditional data repositories such as databases. For AI models to be broadly useful, they must be distributed to the various venues where inference occurs, including datacenters, cloud servers and edge devices. This complicates the process of securing models across their life cycle. They must also have safeguards in place to protect against attacks and minimize the chance of unwanted behavior. To protect company assets, study participants recommended adopting best practices such as zero-trust principles combined with extensive staff education regarding models and data sources.

"All model-related IP is stored within the data center, and access is only granted for specific use cases."

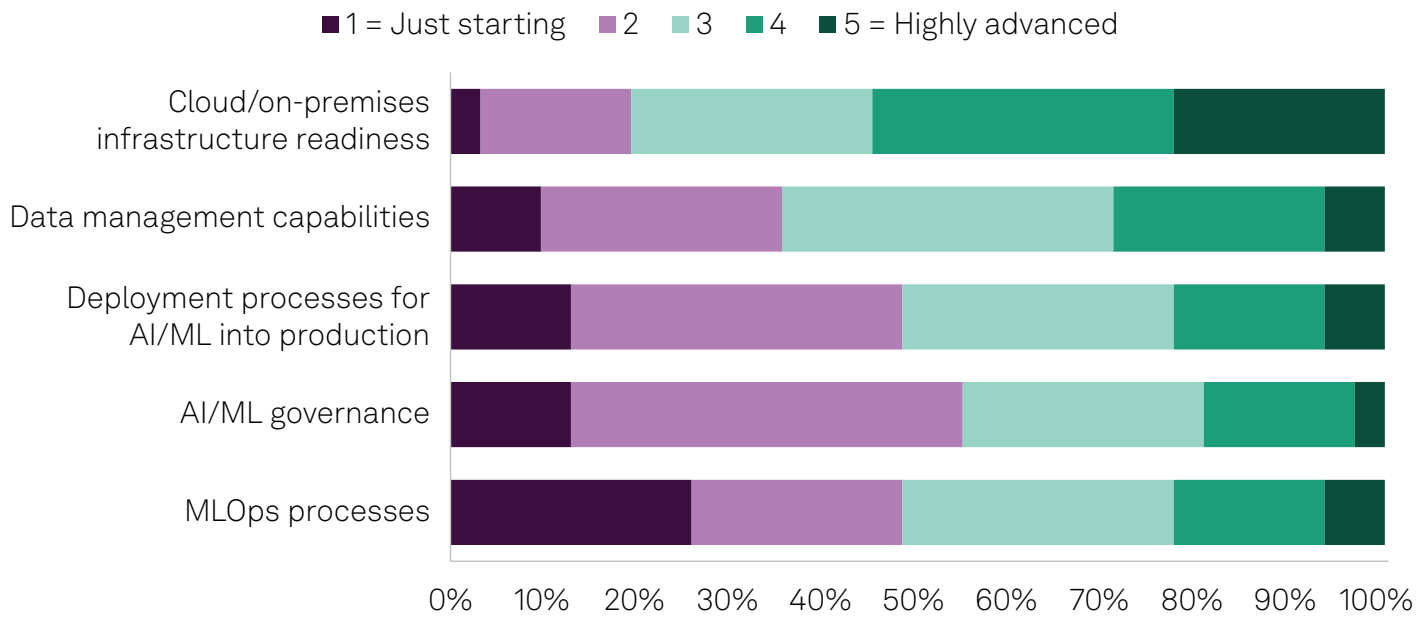
Director of engineering infrastructure
Financial services, >50,000 employees, US

"[Recommended practice is] training our colleagues to use internal AI models."

IT director
Life sciences, >50,000 employees, US

Organizational maturity lags in areas that correlate with securing models, such as MLOps and AI/ML governance, as illustrated in Figure 4. Roughly half of respondents self-evaluate as having early-stage capabilities (rated 1 or 2 on a 5-point scale) in each of these categories.

Figure 4: Organizational AI maturity



Q. Overall, how would you rate your organization's AI/ML maturity across the following dimensions?
Base: Executive discussion board participants (n=30).
Source: S&P Global Market Intelligence 451 Research AI at Scale study commissioned by Verizon.

To better address these practices, participants recommended the practice of "shift left," which dictates earlier incorporation of security testing, risk management and governance efforts in the AI model and application development cycles. Respondents often wished for better version-control mechanisms with deeper logging and monitoring capabilities throughout the model life cycle. Many regretted not having proper backup and snapshotting capabilities. They also saw potential touchpoints between these capabilities and best practices and other AI infrastructure elements, such as retrieval-augmented generation (RAG) and vector databases.

"[Recommended practice is] continuous monitoring to detect outputs that are not expected."

Director of digital health infrastructure
Healthcare, 5,000-10,000 employees, Australia

These recommendations will become increasingly relevant as organizational AI capabilities progress. Enterprises will need to leverage increasing quantities of proprietary data from various sources to create value. This often involves using RAG to enhance models with additional enterprise data. While there is clear value in doing so, study respondents acknowledge the accompanying risk of exposing highly valuable data assets to AI models.

"Generative AI capabilities are fantastic, especially if you combine the power of the LLM with your own data. This is what's called retrieval augmented generation … And it can be extremely important. You can connect your emails, all your documentation, technical project documentation, all your customer databases, whatever you want, because it could be useful in the right model for you to be more efficient. But it also generates additional cybersecurity risks."

VP of IT
Utilities, 10,000-20,000 employees, France

While the outlook for AI is bright, the variety of new tools and competencies that it requires can be overwhelming for organizations, particularly for those just beginning the journey. Study participants frequently recommended starting small and gaining experience and confidence before moving to larger projects with greater unpredictability.

"All these questions that you have around security, compliance, AI governance. These are best learned and experienced by just doing it in smaller environments, that are in controllable environments."

Director of IT for EMEA
Manufacturing, 10,000-20,000 employees, Germany

Implications and next steps

451 Research's Voice of the Enterprise: AI & Machine Learning, Use Cases 2025 survey shows clear organizational momentum behind AI technologies. Nearly all respondents (95%) said generative AI is in use in some capacity in their organization, with 27% indicating it is fully deployed in production scenarios. Organizations at all stages of adoption need to secure their data and AI workloads, and these efforts can benefit from the four major insights highlighted in this report:

Insight 1: Security for AI should span multiple environments and be designed to scale

While participants reflected a preference for public cloud as an AI workload execution venue, there is no one-size-fits-all strategy. Organizations run workloads in various locations to suit the needs of the business. Security strategies and capabilities must encompass all locations in use and should be thoughtfully crafted with future scalability and flexibility in mind.

Insight 2: Additional security tools and capabilities are sorely needed

The data is clear: Every participant in our study indicated their organization will need to upgrade security toolsets over the next 12-24 months to support AI workloads. In many cases, these upgrades will allow organizations to implement zero-trust frameworks with robust identity management and the visibility and management capabilities needed to secure their environments at scale. Respondent organizations are also seeking enhanced threat detection tools that support cloud-native technologies and are designed for AI workloads, as well as infrastructure and tooling to support end-to-end encryption.

Insight 3: Security tools and strategies should support diverse governance requirements

Enterprise adoption of AI further compounds the challenges posed by evolving regulatory environments. Not only must organizations continue to secure traditional workloads, but they must also supervise increasingly autonomous models that can access — and may expose — valuable enterprise datasets. Regulations vary by geography and industry, so organizations must craft their security strategies carefully to satisfy all applicable frameworks. The stakes for doing so are high, as security incidents and related regulatory failures can lead to downtime, fines and even long-lasting reputational damage.

Insight 4: AI models themselves must be secured and monitored, in addition to the environments where they run

Organizations must protect their investments in AI infrastructure and related assets. Models and the datasets they leverage are highly valuable to enterprise stakeholders. They're also prime targets for bad actors seeking unauthorized access or control. Proper levels of security and visibility should be in place to ensure models are functioning as expected and are insulated from potential threats. These safeguards also provide additional benefits such as reduced AI application downtime and improved compliance.

About Verizon Business

Verizon Business is all about helping organizations like yours succeed in today's dynamic digital environment. We provide essential network solutions that support and enhance business operations and empower how millions of people live, work, and play every day. This research was commissioned to provide insights into the challenges of AI deployment, share peer experiences, and clarify the strategic choices facing IT leaders. Our aim is to offer a deeper understanding of these intricate dynamics, especially the critical importance of network and security infrastructure. We believe these insights can help your organization to scale AI confidently, maximize its potential, and consistently stay ahead in this rapidly evolving landscape.

About the author

Greg Macatee

Senior Analyst

Greg Macatee is a senior analyst leading coverage of AI infrastructure within the Data, AI & Analytics channel of S&P Global Market Intelligence 451 Research. In his research, Greg seeks to identify and highlight the impacts of current and future enterprise AI and machine learning technology trends on the IT infrastructure that underpins them. His areas of expertise include AI infrastructure (compute and storage), software-defined compute (virtualization and containers), storage software, software-defined storage and data protection; and cloud-native and open-source technologies.

About this report

A Discovery report is a study based on primary research survey data that assesses the market dynamics of a key enterprise technology segment through the lens of the "on the ground" experience and opinions of real practitioners — what they are doing, and why they are doing it.

About S&P Global Market Intelligence

At S&P Global Market Intelligence, we understand the importance of accurate, deep and insightful information. Our team of experts delivers unrivaled insights and leading data and technology solutions, partnering with customers to expand their perspective, operate with confidence, and make decisions with conviction.

S&P Global Market Intelligence is a division of S&P Global (NYSE: SPGI). S&P Global is the world's foremost provider of credit ratings, benchmarks, analytics and workflow solutions in the global capital, commodity and automotive markets. With every one of our offerings, we help many of the world's leading organizations navigate the economic landscape so they can plan for tomorrow, today. For more information, visit www.spglobal.com/marketintelligence.

CONTACTS

Americas: +1 800 447 2273
Japan: +81 3 6262 1887
Asia-Pacific: +60 4 291 3600
Europe, Middle East, Africa: +44 (0) 134 432 8300

www.spglobal.com/marketintelligence
www.spglobal.com/en/enterprise/about/contact-us.html