The C-Suite Guide to GenAI Risk Management

Generative AI (GenAI) is rapidly transforming businesses, offering unprecedented opportunities for innovation and efficiency. From streamlining processes to unlocking new avenues of creativity, the potential is immense. But with great power comes great responsibility. The rapid adoption of GenAI brings with it a new frontier of security challenges that every organization must navigate.

We designed this guide with you—the modern executive—in mind. We'll take a deep dive into the unique risks GenAI introduces to provide you with a clear, actionable framework to protect your enterprise. We'll do this by walking you through:

- **Navigating the landscape of AI tools**: Get practical advice on managing access to sanctioned GenAI tools while keeping unsanctioned tools out of your network.
- **Pinpointing and prioritizing key risk areas**: Understand where your vulnerabilities lie and how to address them before they become threats.
- **Improving risk posture with actionable steps**: Learn how to implement robust security measures tailored to GenAI's unique risks and demands.

At Palo Alto Networks, we work hard every day to be at the forefront of AI security, sharing insights and crafting strategies to help you harness the power of GenAI without compromising on security. By following this guide, you can confidently lead your organization into the AI-driven future, knowing your assets and data are well protected.

The Evolving Landscape of Generative AI Applications

Generative AI is moving fast, with new applications and use cases emerging almost daily. This constant innovation brings inherent complexity to how organizations should manage and secure GenAI. To effectively adopt and control GenAI tools within your enterprise, it's essential to first understand the various ways in which these applications can manifest. Here's a breakdown of the key forms GenAI can take (see figure 1).

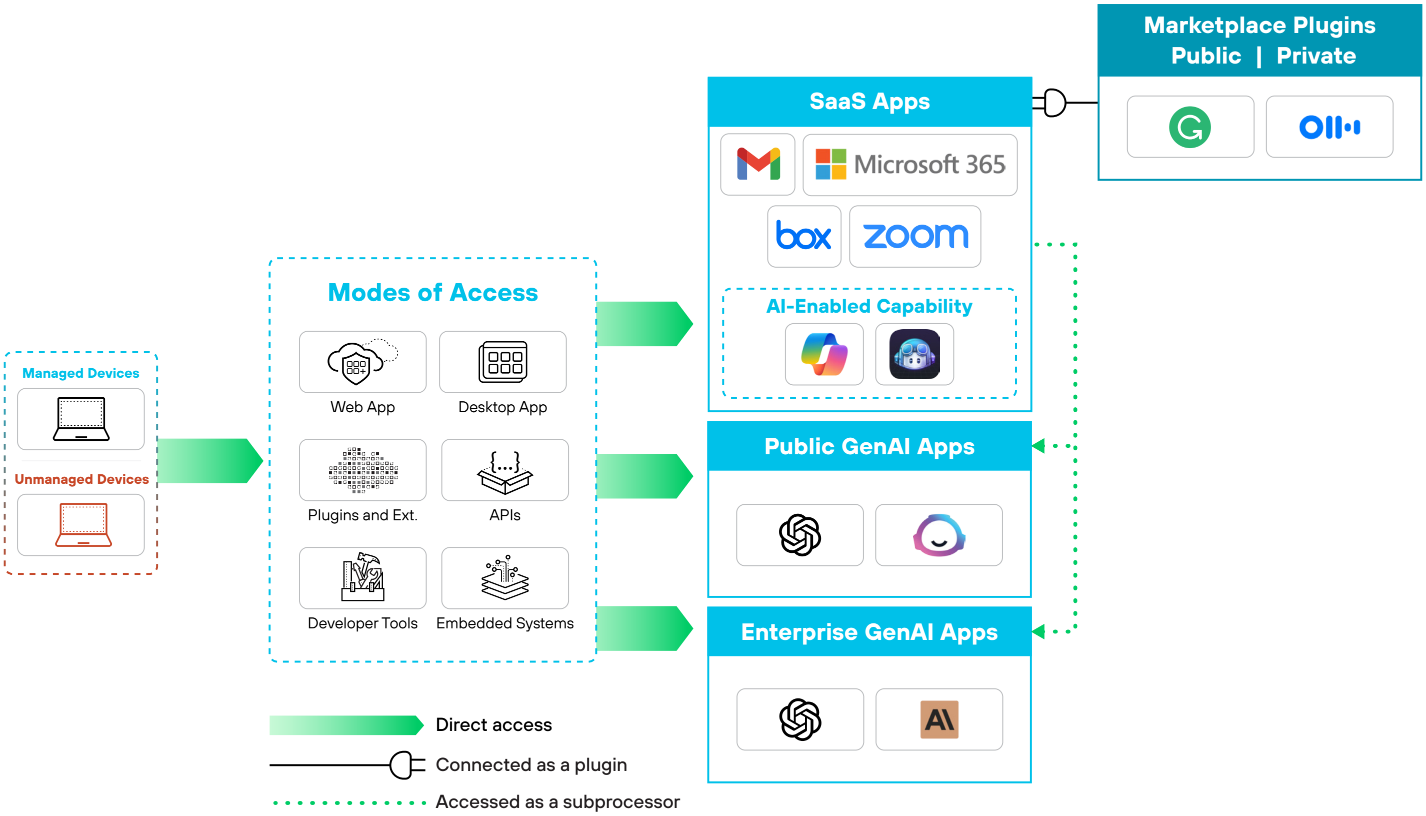


Figure 1: Users can access GenAI apps in many different ways

Publicly Available GenAI Applications

GenAI applications can be directly available to the public, accessible without any intermediary platforms. These applications are typically hosted online, and anyone with access to the internet can use them. They can be either public consumer apps or enterprise apps. For example, OpenAI's ChatGPT is available as a free web application that anyone can access to generate text-based responses.

Marketplace AI Plugins

GenAI applications can also appear as plugins in various marketplaces. Marketplace AI plugins, public and private, are transforming how businesses operate and interact with customers. While they offer substantial benefits in terms of efficiency and personalization, stakeholders must navigate the complex interconnected AI ecosystem. For instance, many AI plugins cut across both corporate and personal browser profiles, creating additional risk.

Public Marketplace Apps

Public marketplace apps are plugins available for installation through the public app marketplaces of major SaaS platforms. They are visible and accessible to all users of those platforms. For example, the AI plugin People.ai (available in the Slack app directory) transforms business activities such as email, meetings, and contacts into account opportunity management within their Slack application.

Private Marketplace Apps

Private marketplace apps are plugins distributed through app marketplaces but only visible or available to specific organizations or user groups, often for security or exclusivity reasons. A custom ServiceNow AI plugin listed in the ServiceNow Store, for example, is visible and accessible only to specific organizations that have received approval from the organization's administrator.

Direct Integrations

GenAI applications can also be integrated directly into enterprise systems with specific authentication and connection methods, including through connected apps and service account integrations.

Connected Apps

Connected apps are plugins that use a client ID and client secret for authentication and integration. They typically involve a more direct, API-based connection between your product and the target SaaS application. For example, an AI-driven analytics tool connects to Salesforce using OAuth 2.0, utilizing a client ID and secret to authenticate and access Salesforce data for generating sales insights.

Service Account Integrations

Service Account Integrations are plugins that utilize service accounts for integration, often providing a higher level of access and control, suitable for more complex or sensitive integrations. This could be a security plugin that uses GitHub apps or service accounts to perform security audits on code repositories and manage access controls.

GenAI Natively Embedded within Existing SaaS Applications

GenAI capabilities can also be embedded natively within existing SaaS applications, enhancing the functionality of these platforms. These can include new AI-enabled services and subprocessors in SaaS applications.

New AI-Enabled Services

SaaS applications introducing entirely new AI-enabled services, such as AI-powered assistants or copilots, augment user productivity and decision-making. Microsoft 365 Copilot, for example, provides AI-driven assistance in writing, data analysis, and task management within Microsoft 365 applications.

Subprocessors in SaaS Applications

GenAI functions as a subprocessor within SaaS applications, handling specific tasks or processes to improve efficiency and performance. For example, an AI-based customer service automation feature within Zendesk helps triage and respond to customer inquiries.

As we look to the future, we envision that eventually, every application will have some GenAI component natively infused. Soon, we can expect GenAI to be seamlessly integrated into nearly every application your organization relies on. This shift will demand more than just passive oversight; it will require a proactive approach to visibility and control. Continuous monitoring and strategic management will be essential to fully harness the power of GenAI while keeping your data secure. By staying ahead of these developments, you can enable your organization to reap the benefits of GenAI without increasing risk.

The Application Taxonomy

In the diverse landscape of GenAI, not all applications are created equal. To effectively manage and secure these tools, it's crucial to understand the different categories they fall into. We can generally group applications into three distinct types: sanctioned, tolerated, and unsanctioned, each with its own unique characteristics and security implications. Understanding this taxonomy will help you tailor your risk management strategies to better protect your organization.

Sanctioned Applications

Sanctioned applications are software services that an organization officially approves for use, typically managed by the IT department. Recognized for providing business benefits, these applications have undergone rigorous security vetting. For example, a large financial institution could approve the use of OpenAI's ChatGPT Enterprise for generating text-based content and automating customer service interactions.

Tolerated Applications

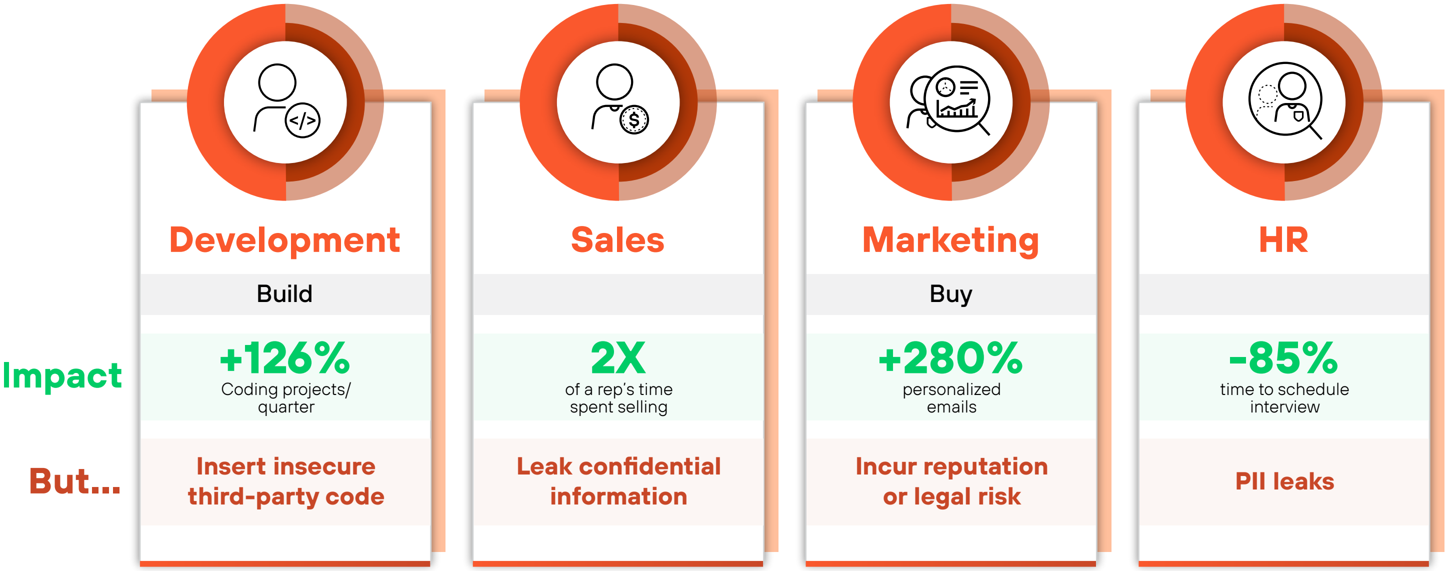
Tolerated SaaS applications are those that an organization allows due to a legitimate business need, even though the IT department doesn't officially provide or manage them. These applications may come with certain restrictions to mitigate risks associated with their use, as they do not meet the same security and compliance standards as sanctioned applications. A high-tech company, for example, might allow its marketing team to use Jasper only for creating marketing material and social media content with strict data usage and handling guidelines.

Unsanctioned Applications

Unsanctioned applications lack formal approval and may pose security risks or compliance issues. Used without official approval, these apps often introduce significant security risks. For example, employees of an automobile company could use free online AI writing assistants for drafting sensitive business documents despite IT not officially sanctioning or allowing the apps.

Key Risk Areas: The Double-Edged Sword of GenAI

Picture this: a bustling modern enterprise, alive with the excitement of GenAI. Across departments, employees are embracing these innovative tools to streamline work, boost productivity, and unlock their creative potential. The energy is infectious, and the benefits are undeniable. But as we've explored, GenAI isn't just a powerful ally—it's a double-edged sword. Beneath this enthusiasm lies a complex web of risks that IT and InfoSec teams must carefully navigate to protect the organization.



	Development	Sales	Marketing	HR
	Build		Buy	
Impact	+126% Coding projects/ quarter	2X of a rep's time spent selling	+280% personalized emails	-85% time to schedule interview
But...	Insert insecure third-party code	Leak confidential information	Incur reputation or legal risk	PII leaks

Figure 2: GenAI enables significant improvements in productivity—but comes with risk

The Invisible Threat: Lack of Visibility

In the marketing department, a creative intern discovers a free online AI writing assistant. Thrilled by its ability to generate catchy slogans in seconds, the intern starts using it daily. However, they are unaware that they are inadvertently uploading sensitive company data to an unsanctioned third-party application. The IT team, lacking visibility into this "shadow AI" usage, remains oblivious to the data leak.

The Access Dilemma: Weak Controls

Meanwhile, in R&D, a team is leveraging a powerful AI analysis tool to process years of experimental data. The insights are groundbreaking, but unrestricted access across the department means even junior members can access and mishandle the confidential research findings. The absence of granular role-based access controls turns this innovative tool into a security liability.

The Lack of Training: Malicious Content

Customer service representatives are delighted with their new AI-powered chatbot, which dramatically improves response times. However, the IT team discovers that the chatbot occasionally includes potentially malicious links in its responses, putting both employees and customers at risk. The challenge of inspecting AI responses for malware, phishing links, and malicious content, and training users to be vigilant when clicking on links, becomes evident.

The Plugin Predicament: Hidden Risks in Plain Sight

The sales team raves about a new AI plugin for their CRM system, which predicts customer behavior with uncanny accuracy. As its popularity grows, the IT team struggles to maintain visibility and control over its data access across various SaaS marketplaces. This blind spot in plugin management creates potential vulnerabilities.

Data at Rest: The Invisible Accumulation

As GenAI applications proliferate throughout the enterprise, sensitive data begins to accumulate within these tools, perhaps even being used to train their AI models. The IT team realizes they have limited visibility into what information these apps are storing and where, creating potential compliance nightmares and data exposure risks. Meanwhile, cybercriminals seek to access the leaked code or sensitive data since it's now a part of the LLM training dataset.

The Productivity Paradox

In an effort to lock down this newfound security threat, the IT team rolls out strict controls across the organization. But in their zeal, they unintentionally block access to tolerated GenAI apps. What starts as a well-intentioned move quickly backfires, leaving employees frustrated and productivity stifled.

As the enterprise leadership teams grapple with these challenges, they realize that harnessing the power of GenAI requires more than just adopting new tools. It demands a comprehensive risk management strategy that addresses visibility, access control, data protection, and continuous monitoring. Only by understanding and mitigating these risks can organizations unlock the full transformative potential of GenAI while safeguarding their most valuable assets: their data.

Risk Posture Assessment

If the scenarios we've explored feel all too familiar, you're not alone. Many organizations find themselves navigating similar challenges as they integrate GenAI into their operations. To effectively manage these risks, a thorough assessment of your organization's overall risk posture is crucial. This assessment should consider several key factors to ensure a strong security strategy with consideration toward the evolving threat landscape.

Security Best Practices

Evaluate whether your GenAI application security controls align with industry best practices. This includes implementing robust access controls, data loss prevention, marketplace and plugin visibility, SaaS security posture management, and data-at-rest protections.

Active Traffic Analysis

Understand the nature and profile of active traffic to GenAI apps within your organization. Conduct a thorough risk analysis of the GenAI applications users access and the types of data processed or stored within these environments.

Recommendations for Safe GenAI Adoption

To help your organization navigate this complex landscape, we've outlined a series of actionable recommendations. These steps are designed to provide a balanced approach, ensuring that your enterprise can fully leverage GenAI capabilities while maintaining a robust security posture.

Monitor GenAI Applications, Usage, and Data Flows

Set up a Shadow AI discovery service to capture and analyze all GenAI application usage and data flows throughout your network:

- Conduct a comprehensive inventory of GenAI usage across your organization and inspect usage analytics and firewall logs.
- Evaluate each GenAI application's risk profile based on AI-specific attributes, such as data sensitivity, user base, input/output modes, and compliance-based attributes.

Classify Applications

Implement a classification system for GenAI apps, categorizing them into three distinct groups:

- **Sanctioned**: Officially approved for enterprise-wide use
- **Tolerated**: Allowed with restrictions
- **Unsanctioned**: Not approved for use

This classification enables tailored visibility and control measures for each category, ensuring a balanced approach to GenAI adoption.

Implement Granular Access Controls

Establish detailed access controls specifically for GenAI applications to enhance security:

- Develop distinct policies for sanctioned, tolerated, and unsanctioned GenAI apps.
- Block unsanctioned applications in alignment with Zero Trust principles.
- Restrict access to tolerated apps to specific employee groups based on business needs.
- Regularly review and update access policies to reflect changing business requirements and risk assessments.

Enhance Data Inspection and DLP Capabilities

Implement comprehensive data loss prevention (DLP) measures:

- Monitor and inspect data outflows to GenAI applications to prevent data exfiltration.
- Set up policies to decrypt and inspect data flowing to GenAI applications.
- Enforce controls based on data sensitivity levels.
- Regularly update DLP rules to address emerging threats and new data types.

Implement Continuous Risk Monitoring

Establish ongoing security and compliance processes:

- Implement continuous monitoring of GenAI application usage and data flows.
- Conduct regular risk assessments to identify new threats or vulnerabilities.
- Use automated tools to alert on potential security breaches or policy violations.

Provide Employee Training

Develop a training program to educate employees on safe and responsible GenAI usage:

- Deploy end-user coaching or notification alerts through emails, text messages, or chat messages to educate employees or direct them to sanctioned GenAI applications.
- Educate employees on the safe and productive use of GenAI applications.
- Direct users to guidelines on data sensitivity, acceptable use policies, and potential risks.
- Provide regular updates and refresher courses to address new GenAI applications and inform employees on emerging AI-based threats.

Ensure Visibility and Control of GenAI Plugins

Extend security measures to include GenAI marketplaces, plugins, and AI bots:

- Enable visibility into plugins across multiple marketplaces.
- Detect GenAI applications used as plugins, which may access data even if you've blocked direct access to parent applications.
- Implement a vetting process to manage plugin usage and data access.
- Identify, monitor, and remediate unauthorized AI bots.

Establish Comprehensive Data-at-Rest Scanning

Conduct thorough data-at-rest discovery scans of GenAI applications:

- Detect any sensitive data residing at-rest within GenAI apps.
- Identify potential external sensitive data exposure risks.
- Implement remediation measures for any unauthorized data storage or exposure.

Measuring Success for GenAI Security

Successful GenAI security requires mitigating risks alongside enabling innovation and productivity. To ensure your GenAI security strategy is hitting the mark, keep an eye on the key metrics that follow.

Productivity Gains and Innovation

GenAI applications can enhance productivity and foster innovation within enterprise organizations. The adoption of these applications can help streamline operations, automate repetitive tasks, and augment human creativity, leading to substantial gains across various business functions. Start to measure its productivity gains, whether that's employee productivity, reduction in service tickets, time savings, or improved customer service.

Adoption Rate of Sanctioned GenAI Applications

By embracing GenAI, organizations not only boost productivity but also create a more dynamic and responsible workforce, ultimately driving growth and success. Start to track the percentage of employees actively using approved GenAI tools. A high adoption rate indicates successful enablement and user satisfaction with secure GenAI tools.

Employee Satisfaction and Empowerment

Conduct surveys to assess employee satisfaction with available GenAI tools and gauge their sense of empowerment. GenAI applications can personalize interactions and provide instant support for employees, enhancing their overall experience. Positive feedback here will lead to higher job satisfaction and retention rates as they feel supported and engaged in their roles.

Reduction in Shadow AI Usage

Monitor the decrease in unsanctioned GenAI app usage across your organization as employees transition to approved tools. Reducing shadow AI and unsanctioned apps is essential to minimize the likelihood of data breaches and sensitive data exposure. Organizations can face legal penalties, remedial action, and reputational damage if the use of unsanctioned AI apps results in noncompliance with regulatory requirements. This metric demonstrates the effectiveness of your GenAI enablement strategy.

Data Protection Effectiveness

Monitor the reduction in data exposure incidents related to GenAI usage. Employees may inadvertently input confidential data into GenAI tools, increasing the risk of data leaks and unauthorized access to proprietary information. Consistent declines in data security incidents highlight the strength of your data security protocols and their proper implementation.

Securing the Future of AI, Today

As GenAI continues to reshape the business landscape, the importance of a well-crafted risk management framework cannot be overstated. By following the strategies outlined in this guide, your organization can harness the full potential of GenAI while safeguarding against its inherent risks. Success in this journey hinges on continuous vigilance, adaptability, and a proactive commitment to safe and responsible AI usage. With the right approach, you can lead your organization confidently into the AI-driven future.

Learn more about Palo Alto Networks AI Access Security, the purpose-built solution designed to enable safe adoption and use of GenAI applications.

paloalto NETWORKS

3000 Tannery Way
Santa Clara, CA 95054

Main: +1.408.753.4000
Sales: +1.866.320.4788
Support: +1.866.898.9087

www.paloaltonetworks.com