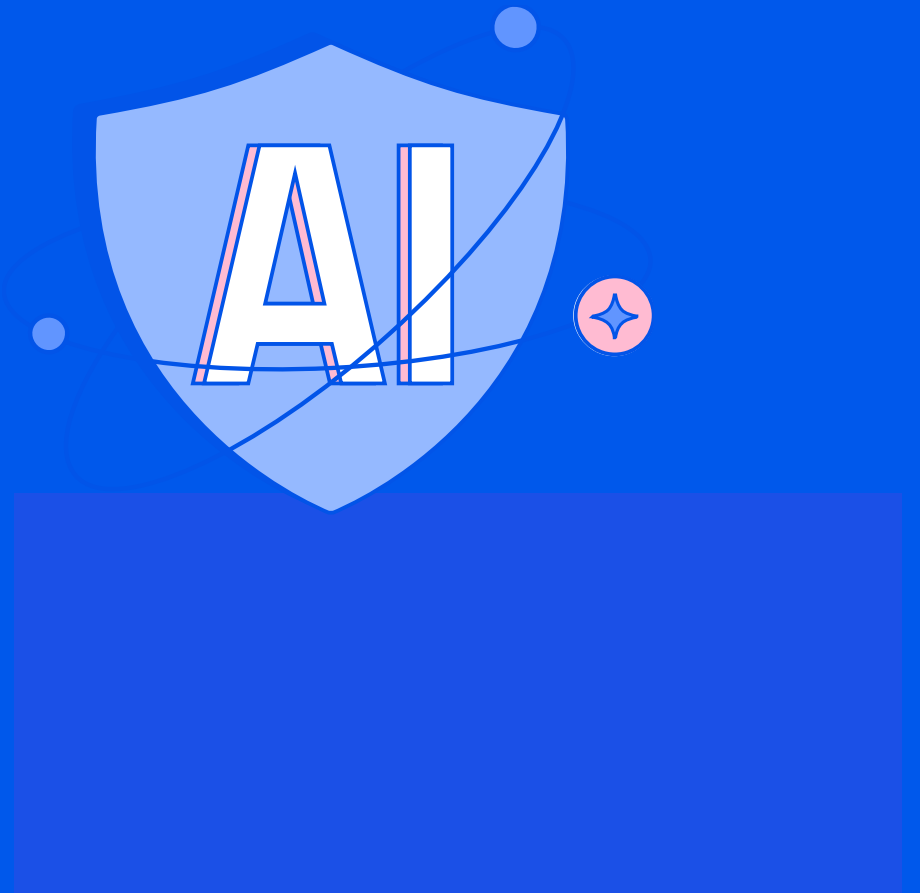


Balancing Innovation and Security: Navigating the AI Landscape



Cloud changed everything



New
environment

How do I get visibility
into my environment?



New risks

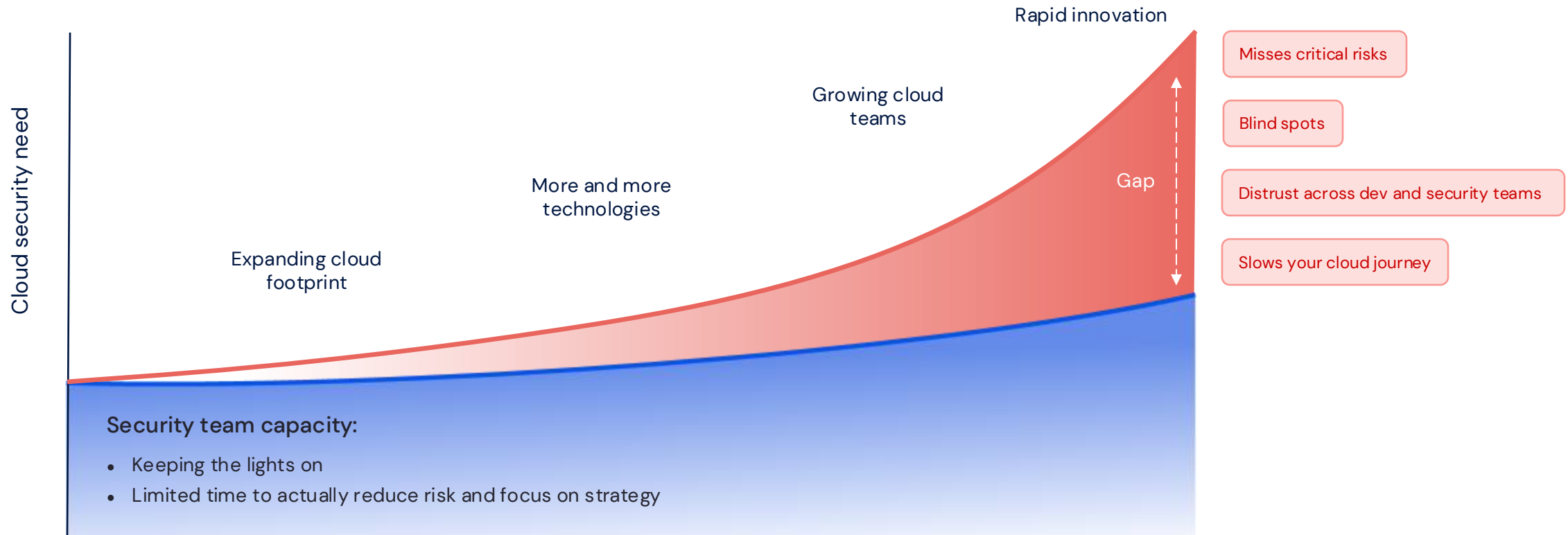
How do I prioritize the real
risks and eliminate the noise?



New ownership model

How do I ingrain
security into our teams?

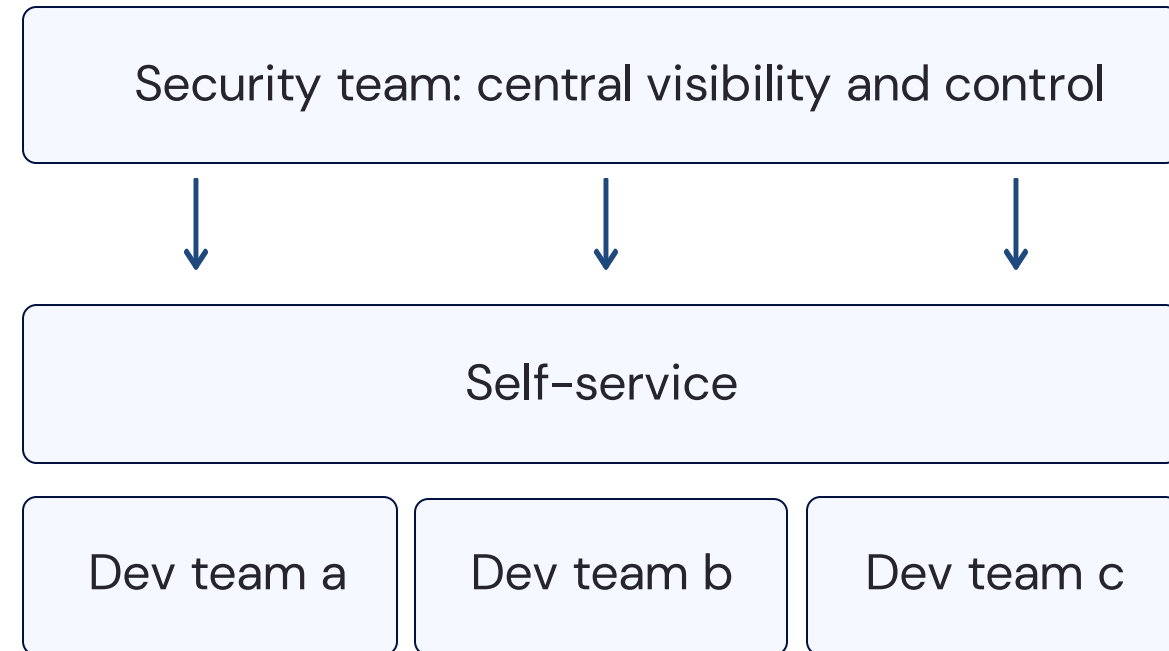
The time to scale your cloud security model is **now**



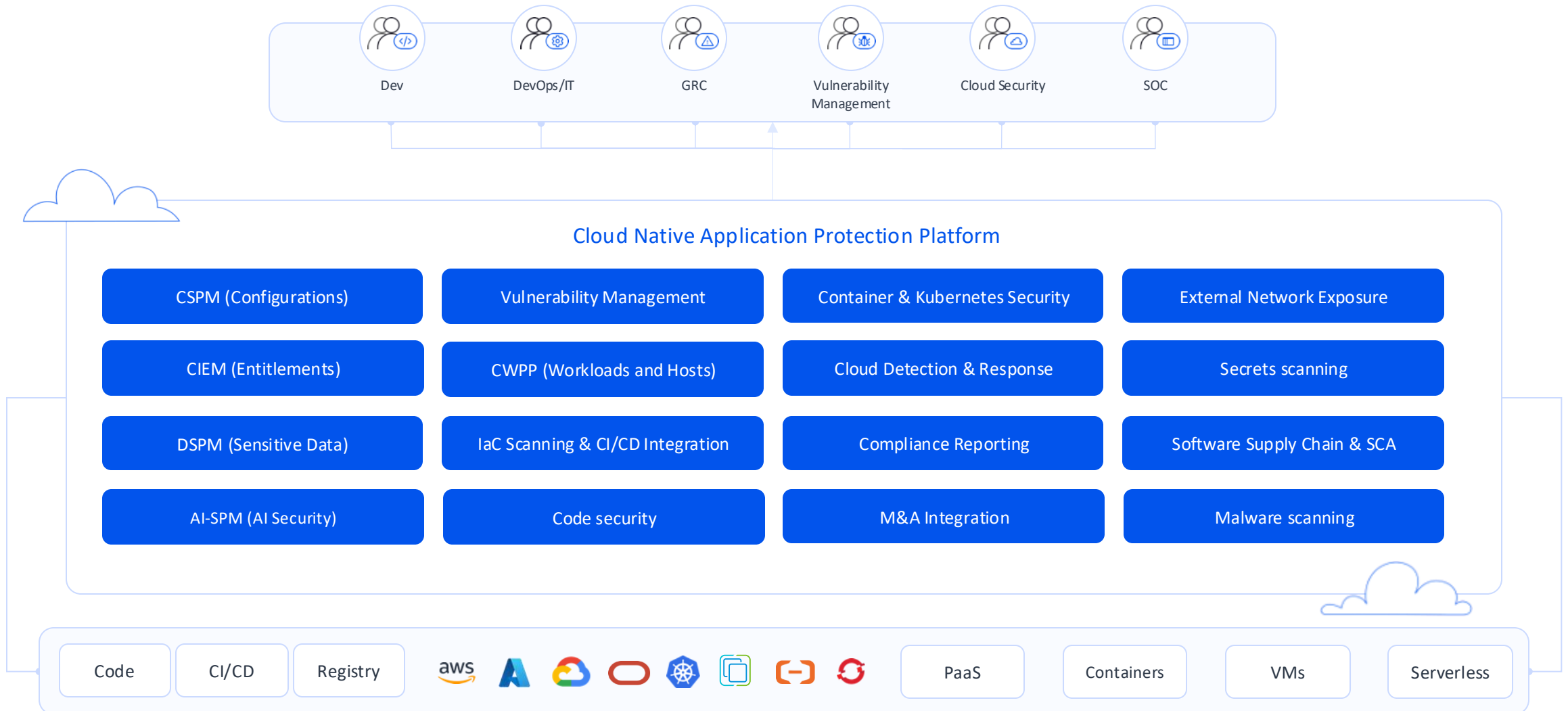
We can't keep doing the same thing

Cloud
security
needs
a new
operating
model

Cloud security is a team sport



One unified platform for your cloud security journey



AI changes everything
again!

AI has taken over the cloud:

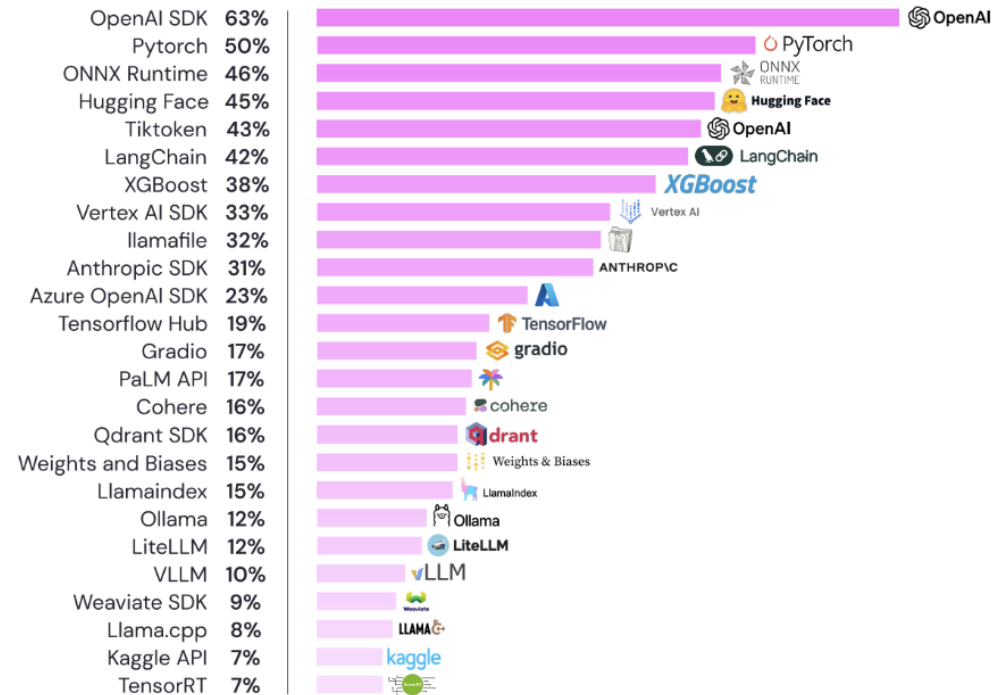
Cloud-based managed and self-hosted AI can be found in over 85% of environments



85% of organizations are using some form of AI (either managed or self-hosted)

AI hosted technologies by percentage of organizations

WIZ Research

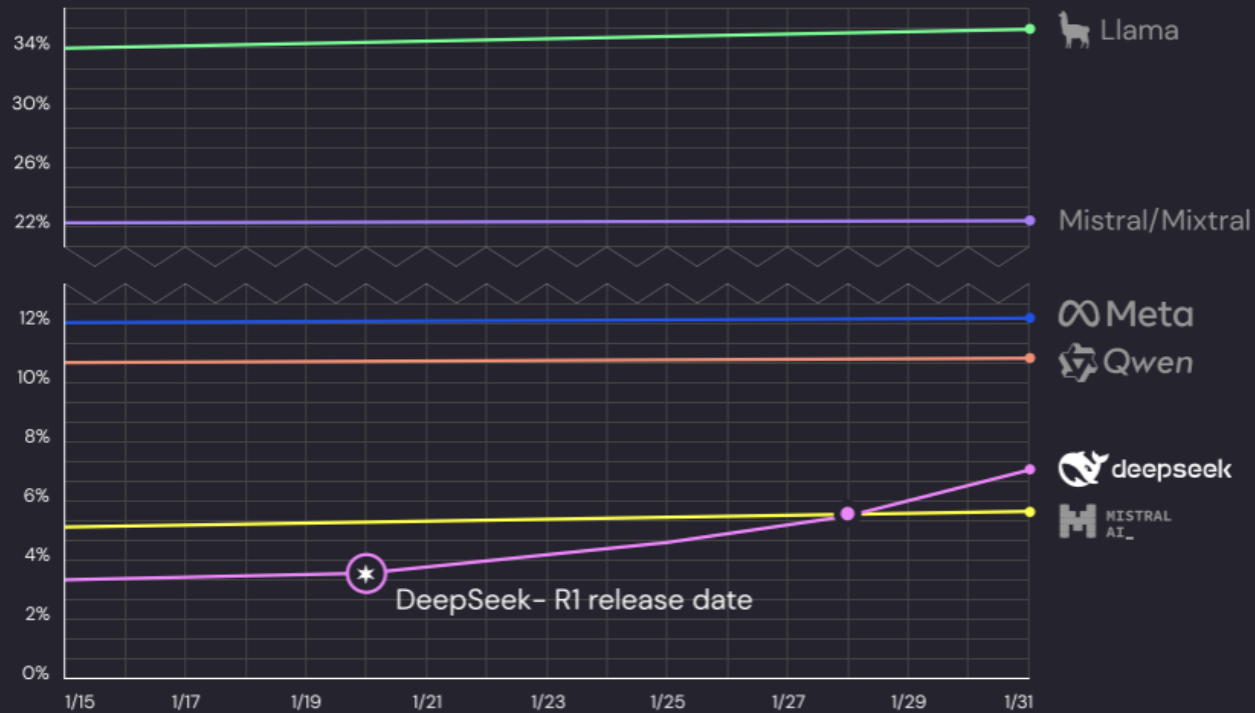


Organizations quickly adopt new innovative technologies

DeepSeek rapid growth

Percent of organizations self-hosting AI models in January 2025

WIZ Research



Wiz Research on DeepSeek

AI applications face the same security risks

← Blog

Wiz Research Uncovers Exposed DeepSeek Database Leaking Sensitive Information, Including Chat History

A publicly accessible database belonging to DeepSeek allowed full control over database operations, including the ability to access internal data. The exposure includes over a million lines of log streams with highly sensitive information.

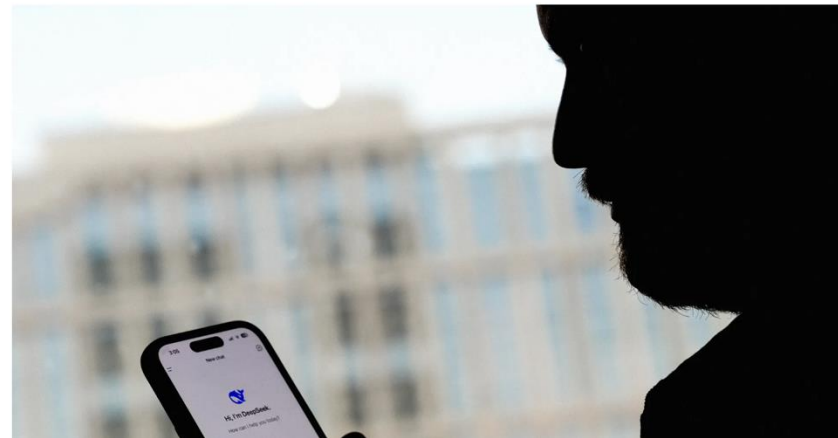
WIRED SECURITY POLITICS GEAR THE BIG STORY BUSINESS SCIENCE CULTURE IDEAS MERCH

Gal Nagli
January 29, 2025 3 minute read

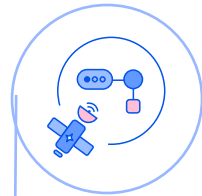


Exposed DeepSeek Database Revealed Chat Prompts and Internal Data

China-based DeepSeek has exploded in popularity, drawing greater scrutiny. Case in point: Security researchers found more than 1 million records, including user data and API keys, in an open database.



Where are you on your AI journey?



Experimentation

- Sporadic use of AI tools
- Shadow AI
- Small POC projects with unvalidated tools



Active development

- Dedicated teams & projects
- Partial visibility
- Limited but dedicated AI deployments

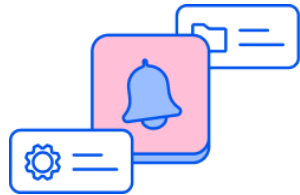


Use in production

- Dedicated AI security
- Visibility into cost and tools
- Large AI deployments



No matter where you are on your AI journey, AI adoption adds a new layer of complexity



New technologies

Where do I have AI in my environment? Do I have full visibility?



New risks

AI-specific risks, how do I detect and remove attack paths in AI pipelines?



New skills

Security teams are not AI experts, how do I upskill them?

AI is here to stay, how do I ensure security doesn't slow down innovation?

We learned from the cloud

Let's apply the learnings to AI



Visibility is the foundation

- Visibility into AI pipelines, models, technologies, toolsets, and data



Continuous risk detection

- Detect AI-specific risks and misconfigurations
- Find sensitive training data
- Identify malicious models
- Detect unintentional exposures and vulnerabilities



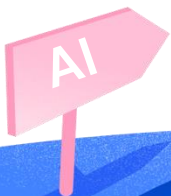
Prioritize critical risk with context

- Correlate AI risks to cloud, workload, and business context to find attack paths to models
- Prioritize the most critical risk on the Wiz Security Graph



Detect threats as the last line of defense

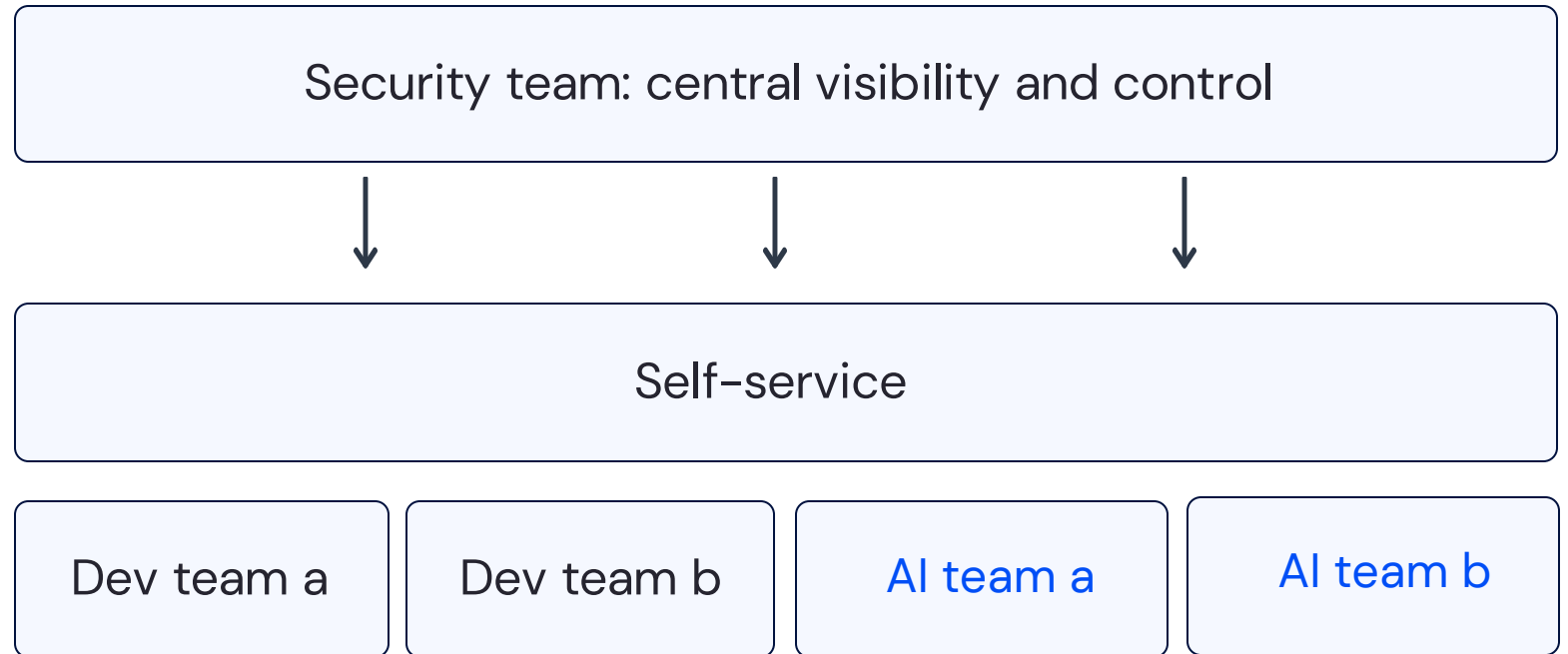
- Identify suspicious behavior and misuse of AI models
- Respond quickly to threats



AI introduces new teams to the security operating model



Extend the cloud security operating model to AI



AI-SPM:
The four
questions
security org
need to ask

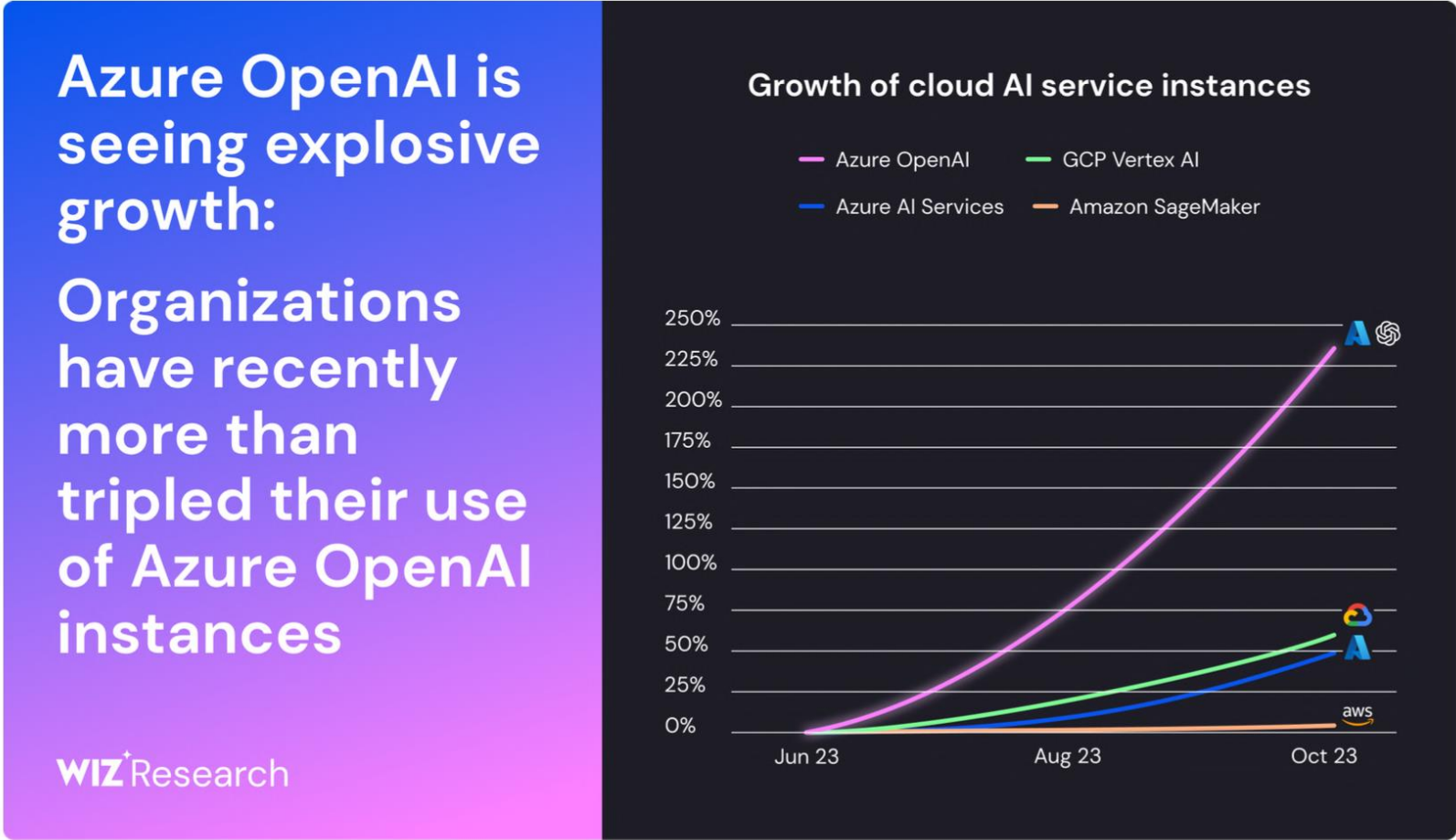
- Do I know what AI services and technologies are running in my environment?
- Do I know what risks exist in my AI pipeline?
- Can I prioritize the critical risks across the AI pipeline?
- Can I detect a misuse in my AI pipelines?



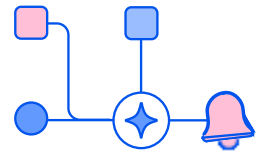
1. Do I know what AI services and technologies are running in my environment?



1. Massive adoption of AI services

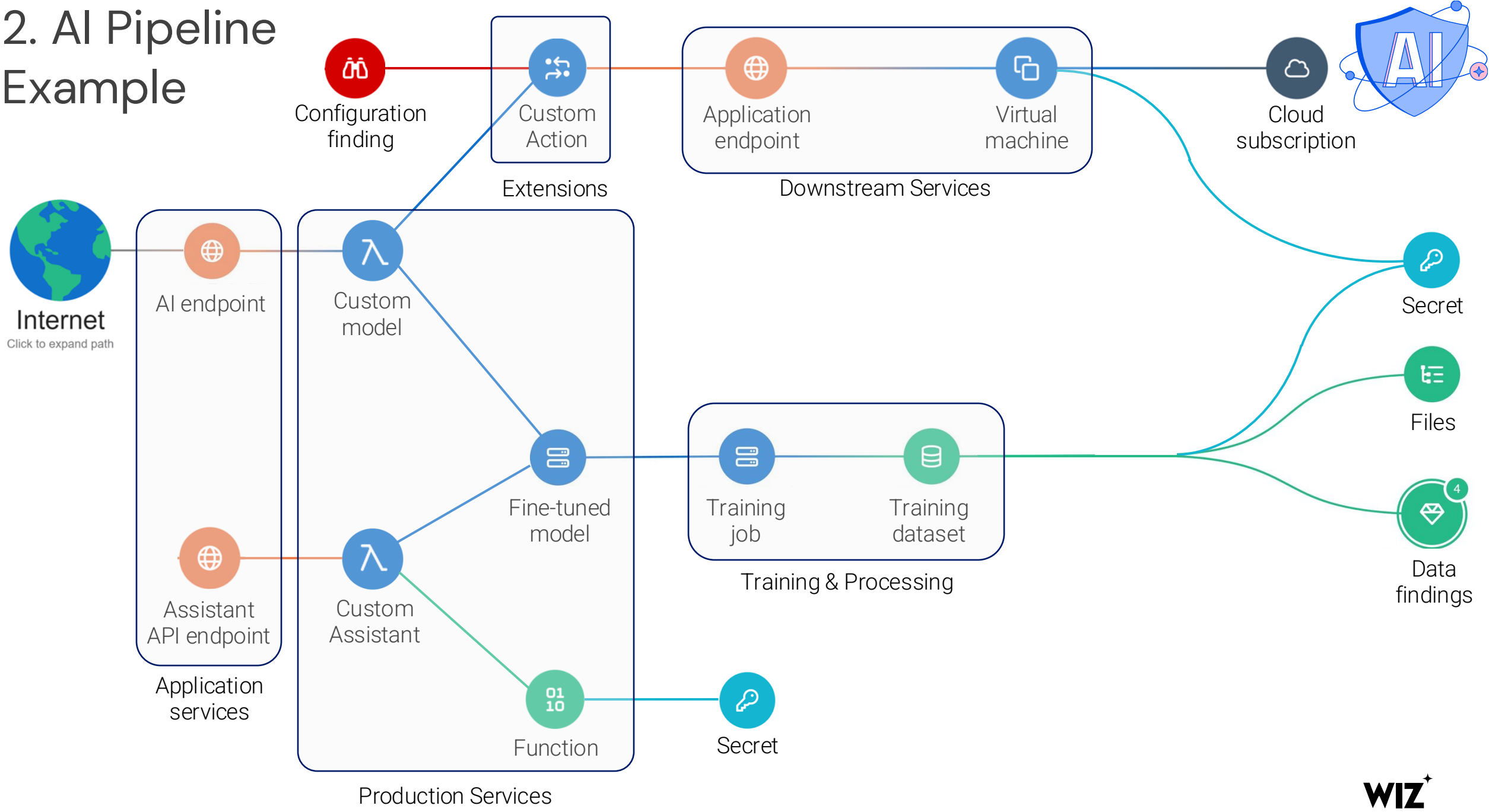


Source: [State of AI in the Cloud 2024](#)



2. Do I know what risks exist in my AI pipeline?

2. AI Pipeline Example





3. Can I prioritize the critical risks across the AI pipeline?

3. Real-life example of an AI toxic combination



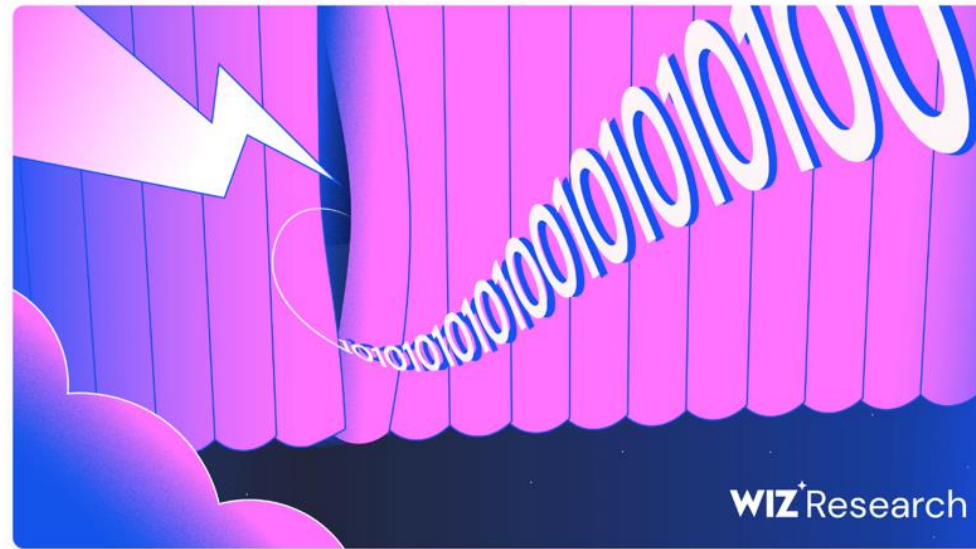
38TB of data accidentally exposed by Microsoft AI researchers

Wiz Research found a data exposure incident on Microsoft's AI GitHub repository, including over 30,000 internal Microsoft Teams messages – all caused by one misconfigured SAS token



Hillai Ben-Sasson, Ronny Greenberg
September 18, 2023

10 minutes read





4. Can I detect a misuse in my AI pipelines?

4. Can I detect a misuse in my AI Pipelines?

Threat detection to respond to AI threats in real time



Real-time Threat detection

Anomaly Detection

Detect Attack Paths

The screenshot displays a security dashboard with three overlapping alert panels. The top panel, titled "Malicious AI Model detected", shows a "Detected events" section with a timestamp of "Nov 26th 04:02:01 PM to Nov 26th 06:55:53 PM" and a "Fileless execution was detected" alert. The middle panel, titled "Anomalous activity detected in AI Model execution engine", shows a "Raw Event Details" section with "Stdin /dev/pts/0" and "Stderr pipe:[79873268]", and a "Process Tree" section. The bottom panel, titled "AI Model escape detected", shows a "Subscription" section with a "Severity" of "Critical" and a "Status" of "Open".

Malicious AI Model detected

Detected events View all 3 events >

Nov 26th 04:02:01 PM to Nov 26th 06:55:53 PM **Fileless execution was detected** 3

Process image path resolved to memfd or shared memory (shm). Memfd (memory file descriptor) and shm (shared memory) are interprocess communication mechanisms in Linux where memfd allows for the creation of anonymous memory objects that can be shared between processes using file descriptors, while shm enables the creation of shared memory segments that allow multiple processes to access and exchange data efficiently in a fast and synchronized manner. This could indicate the presence of a threat actor achieving fileless execution.

Anomalous activity detected in AI Model execution engine

Comment Run an Action Create a Ticket Give Feedback

Overview Remediation Comments

Raw Event Details

Stdin /dev/pts/0

Stderr pipe:[79873268]

Process Tree

gke-sensor-ci-1 Virtual Machine

[1] /usr/lib/ Process

[2030] Process

[1513] Process

[1] Process

Subscription

Severity **Critical**

AI Model escape detected

Comment Run an Action Create a Ticket Give Feedback

Overview Remediation Comments

Process read a file resource used to manage users information by the operating system (/etc/shadow or /etc/gshadow).The "/etc/shadow" and "/etc/gshadow" files in Linux store encrypted user account information and group account information, respectively, providing an extra layer of security by keeping sensitive password-related data inaccessible to regular users. This could indicate the presence of a threat actor achieving credential theft and general user information discovery.

Subscription Projects Risks

Severity **High** Type Threat Detection Issue Related Frameworks

TA0001-T1078.001 Valid Accounts: Default Accounts

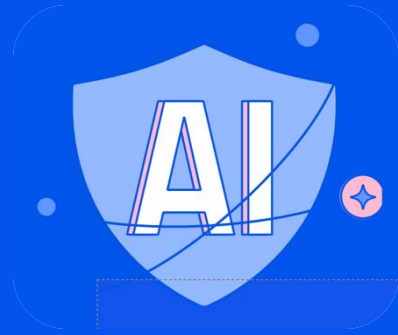
Status **Open**

Due **No due date**

Related Tickets **0 Tickets**

Created Nov 20, 2023 at 4:00 PM

Updated



Introducing **Wiz AI-SPM**

Wiz becomes the first CNAPP to provide AI-SPM capabilities, helping organizations accelerate AI adoption securely

Wiz AI-SPM

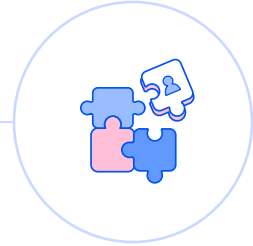
Secure AI Framework



Infrastructure

Secure your Pipelines

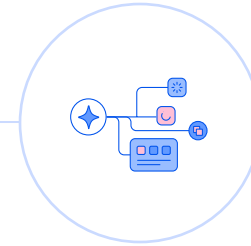
- Secure AI Research and data lifecycle
- Secure AI platforms and services (CIEM, Network, encryption etc.)
- Input & Output Guardrails
- Model inference & Serving



Model

AI Models are IP

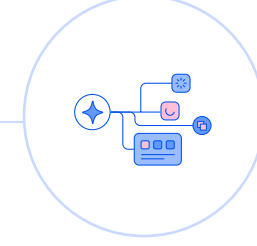
- Integrity
- Supply Chain
- Model Theft
- Malicious models
- Poisoning



Data

Secure your AI Data

- Data Sources
- Data Pipelines
- Training data
- Data integrity and bias



Application

AI is a modern App

- AI & Model SBOM
- Exposure
- Plugins and Extensions
- Vulnerabilities
- Rate limiting and gating
- DoS

Agentless visibility with AI-BOM

Detect every AI tech. with AI-BOM

AI services, SDKs, etc. all without agents. Full visibility on the technology inventory (Sagemaker, VertexAI , Bedrock)










Remove shadow-AI

As soon as your developers start using a new AI service, the security team gains visibility

End-to-end AI pipeline visibility

Detect every resource in your AI pipelines, from the virtual machines hosting the machine learning job, to the data store used for training















Technologies

 All Technologies	486
 Code	42
 CI/CD & Management	48
 Compute Platforms	130
 Application	52
 Security	73
 Cloud Entitlements	50
 Data Assets	67
 AI & Machine Learning	24
Development and Training	11
AI Tools	13

Search technologies



Type ▾

Technology	Resources	Type
 AWS Bedrock Custom Model Development and Training	1 	Cloud Platform Service
 AWS SageMaker Notebook Development and Training	5 	Cloud Platform Service
 Azure AI Search AI Tools	1 	Cloud Platform Service
 Azure AI Service Development and Training	2 	Cloud Platform Service
 Chroma AI Tools	1 	Code Library
 GCP Vertex AI Dataset Development and Training	8 	Cloud Platform Service
 GCP Vertex AI Managed Workbench Development and Training	2 	Cloud Platform Service

AI Misconfigurations

Built-in misconfigurations rules for AWS Sagemaker, Amazon Bedrock, Google Vertex AI, Azure OpenAI

Bedrock Custom Model is not encrypted with a customer-managed key
Cloud Configuration Finding

Ignore Comment Give Feedback

Overview Remediation

Project
5 Projects

Severity
Medium

Configuration Link
docs.aws.amazon.com

Expected Configuration
KMS encryption key should be configured

Has Remediation Instructions
Yes

Vertex AI Workbench user-managed notebook is not using Secure Boot
Cloud Configuration Finding

Ignore Comment Give Feedback

Overview Remediation

Project
9 Projects

Severity
Medium

Configuration Link
cloud.google.com

Expected Configuration
The field 'enableSecureBoot' should be set to 'true'

Has Remediation Instructions
Yes

Cognitive Search public network access is enabled
Cloud Configuration Finding

Ignore Comment Give Feedback

Overview Remediation

Project	Rule ID	Status
9 Projects	DataWorkload-029	Unresolved
Severity	Cloud Platform	First seen
High	Azure	Jun 12, 2023 at 7:12 AM
Configuration Link	Current Configuration	Last changed
learn.microsoft.com	Cognitive Search public network access is enabled	Oct 20, 2023 at 10:38 AM
Expected Configuration	Source	
Cognitive Search should disable public network access	Wiz Configuration Rules	
Has Remediation Instructions	Compliance Frameworks	
Yes	WIZ WIZ	

Detect attack paths to AI

Deep risk analysis in AI pipelines

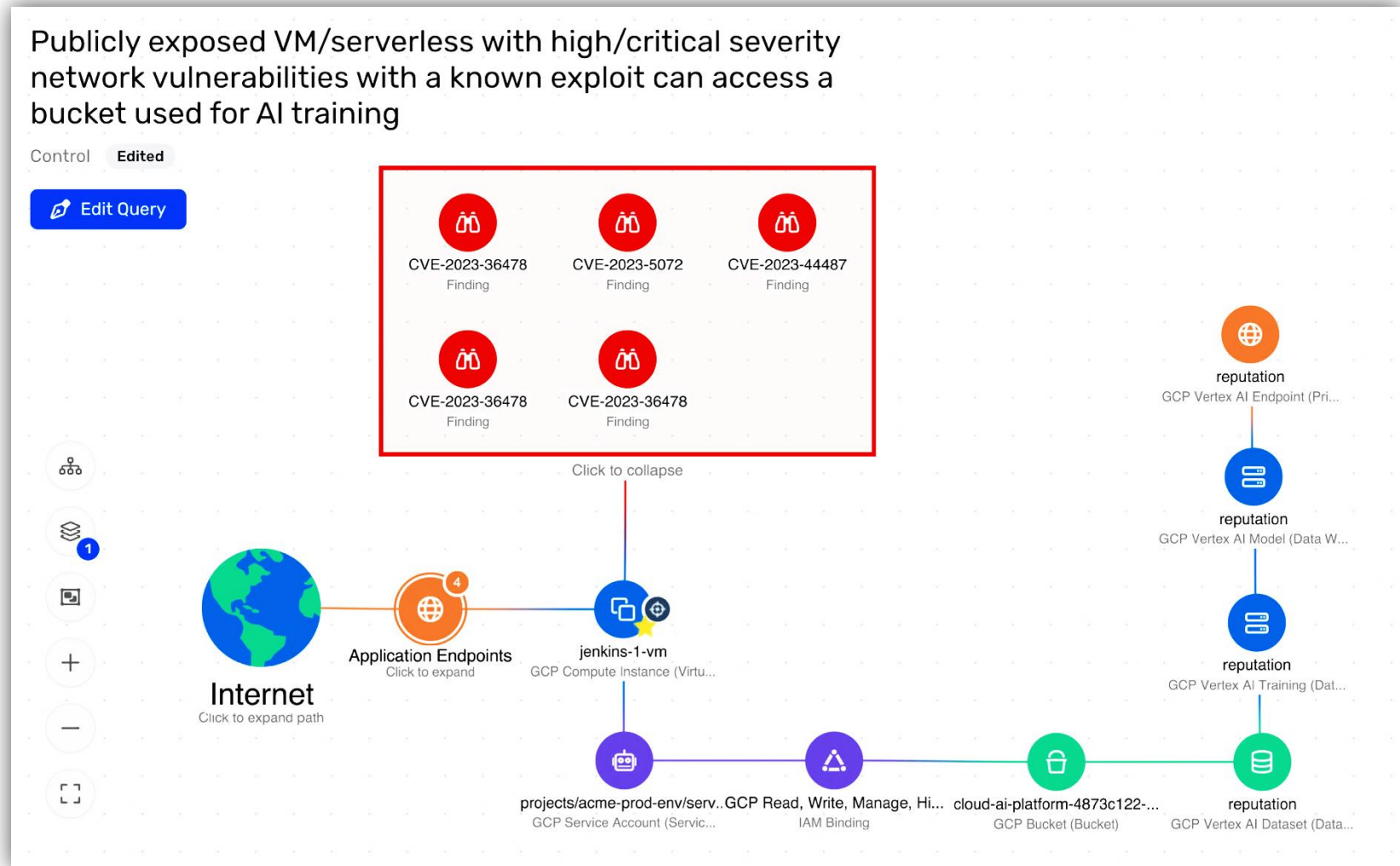
Risk analysis across AI vulnerabilities, misconfigurations, permissions, data, secrets, and network exposure

Protect sensitive training data

DSPM extended to AI to detect sensitive AI training data and protect against data risks such as data poisoning

Remove critical attack paths to AI models

Wiz identifies attack paths in your AI pipelines and allows you to proactively remove the most critical risks with context



Wiz AI-SPM

Remove AI threats in real-time with Wiz Defend



Real-time threat detection

Built-in threat detection rules to detect suspicious activity and misuse of AI models

Detect common LLM attack techniques

The Wiz Research Team adds new rules based on emerging AI attack techniques found in cloud environments

Reduce blast radius with cloud context

Threats are correlated to cloud context to understand how an attacker can move around the environment and reduce attack surface

The screenshot displays the Wiz Defend interface. At the top, a notification reads "Malicious AI Model detected". Below this, a "Detected events" section shows a "Fileless execution was detected" event from Nov 26th. The event description explains that process image paths resolved to memfd or shared memory (shm), which are interprocess communication mechanisms in Linux. Below the event, "Raw Event Details" shows stdin as /dev/pts/0 and stderr as pipe:[79873268]. A "Process Tree" section shows the process path: gke-sensor-ci-125-sensor-ci-general-n-169807b5-04r9.

Overlaid on the interface is a "Threat Detection Rule" for "Creation of IAM User Associated With A Known LLM Hijacking Campaign". The rule description states it detects the creation of an IAM user with a specific regex name. The rule is enabled, has a high severity, and is associated with the AWS CloudTrail framework. It was created on Dec 15, 2024, and last run on Dec 23, 2024.

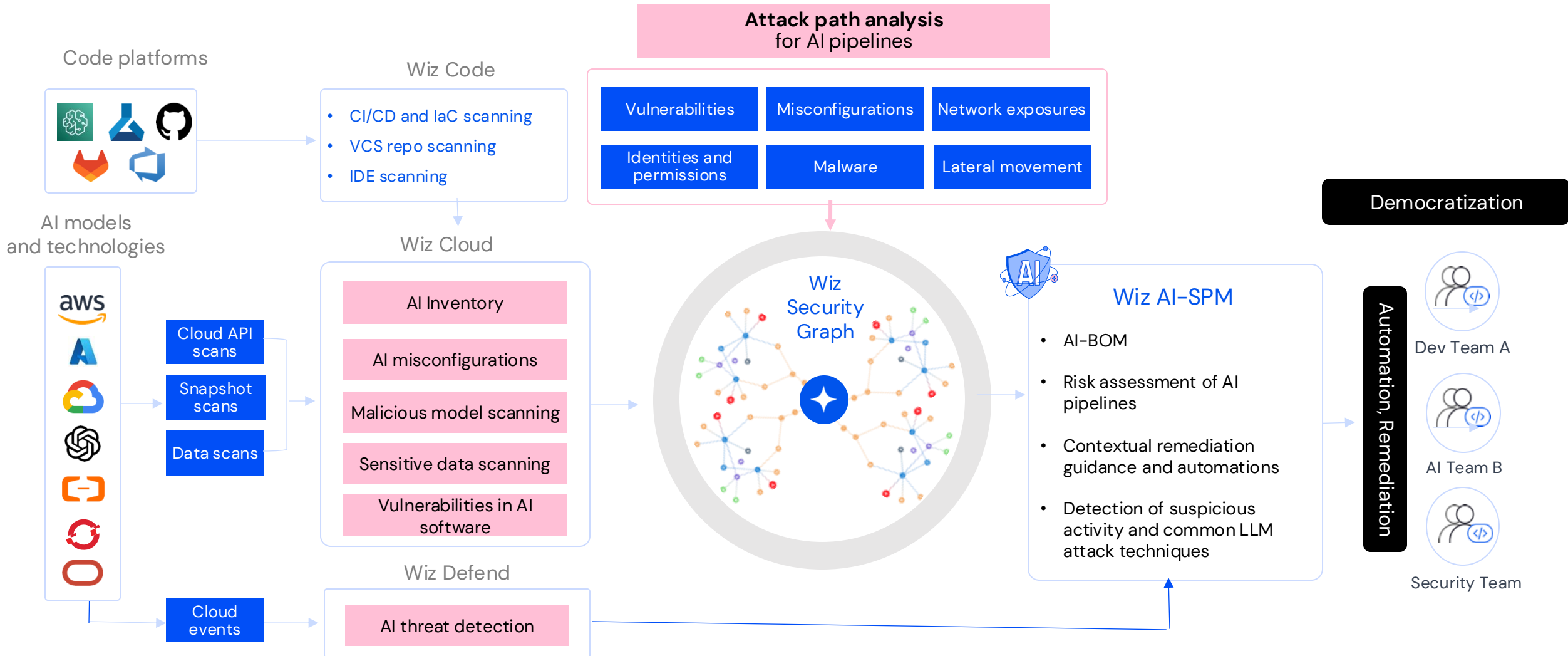
Another section shows "AI Model escape detected" with a description: "Process read a file resource used to manage users information by the operating system (/etc/shadow or /etc/gshadow). The /etc/shadow and /etc/gshadow files in Linux store encrypted user account information and group account information, respectively, providing an extra layer of security by keeping sensitive password-related data inaccessible to regular users. This could indicate the presence of a threat actor achieving credential theft and general user information discovery." This event is categorized as a "Threat Detection Issue" with a high severity.

At the bottom, there are sections for "Subscription" (No projects), "Projects" (No projects), "Risks" (No risks), "Severity" (High), "Type" (Threat Detection Issue), and "Related Frameworks" (TA0001-T1078.001 Valid Accounts: Default Accounts).



Operationalizing Wiz AI-SPM

Wiz AI-SPM empowers organizations to increase AI innovation securely by providing visibility into AI technologies and models, risk assessment for AI pipelines, and threat detection.

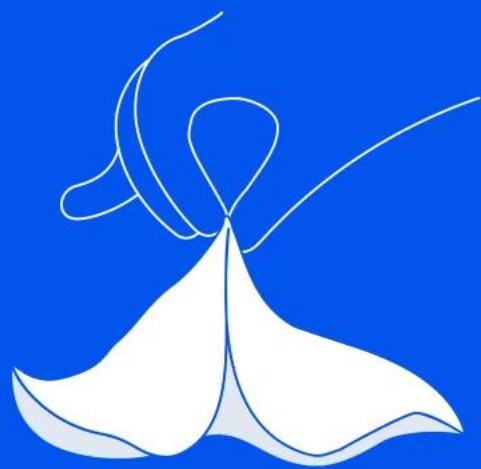




“ We use Wiz AI-SPM to accelerate the pace of AI application development and deployment, while enforcing AI security best practices. As a result, we can deploy AI applications that are secure by design and build trust with key stakeholders.

Rohit Kohli

Deputy Chief Information Officer,
Genpact





Want to see more magic about Wiz AI-SPM?

Follows us on LinkedIn

[linkedin.com/company/
wizsecurity](https://linkedin.com/company/wizsecurity)

Book or see a demo

wiz.io/demo

Learn about Wiz AI-SPM

[wiz.io/solutions/ai-security-
posture-management](https://wiz.io/solutions/ai-security-posture-management)