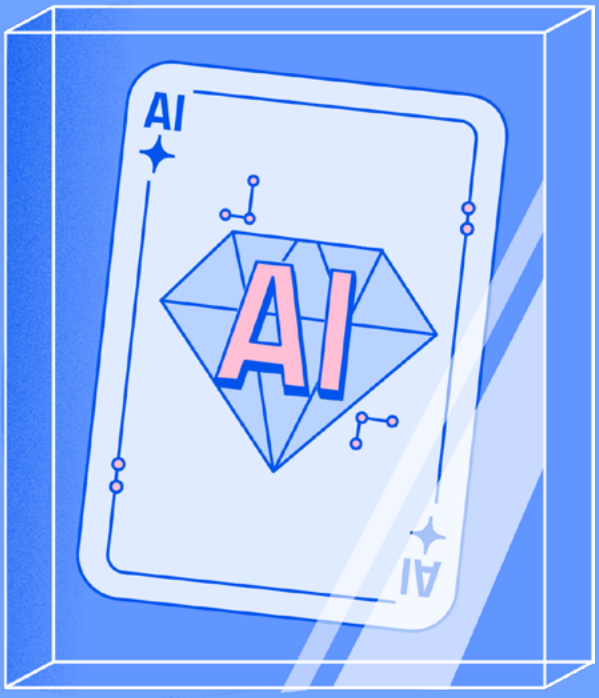
WIZ

# AI Security Assessment Sample Report



# Table of Contents

# Introduction

This is a sample assessment report for the types of AI security insights Wiz AI Security Posture Management (AI-SPM) provides you with. In this report, you will learn about AI Bill of Materials (AI-BOM) that provides you with visibility into any AI technology in your environment, across managed AI services and hosted models. You will see examples of AI-SPM configuration checks for AI services that help you follow AI security best practices. We will review the AI Security Dashboard that centralizes all the AI security insights practitioners need to know in order to focus on the most critical risk. We will introduce the concept of Wiz Issues, which are combinations of risks in your AI pipelines across sensitive data, permissions, misconfigurations, secrets, and vulnerabilities that create an attack path to your models. For each Issue, you will see a summary of the findings and the evidence on the Wiz Security Graph.

These are just a small subset of the security insights you get with Wiz to help you get familiarized with the platform.

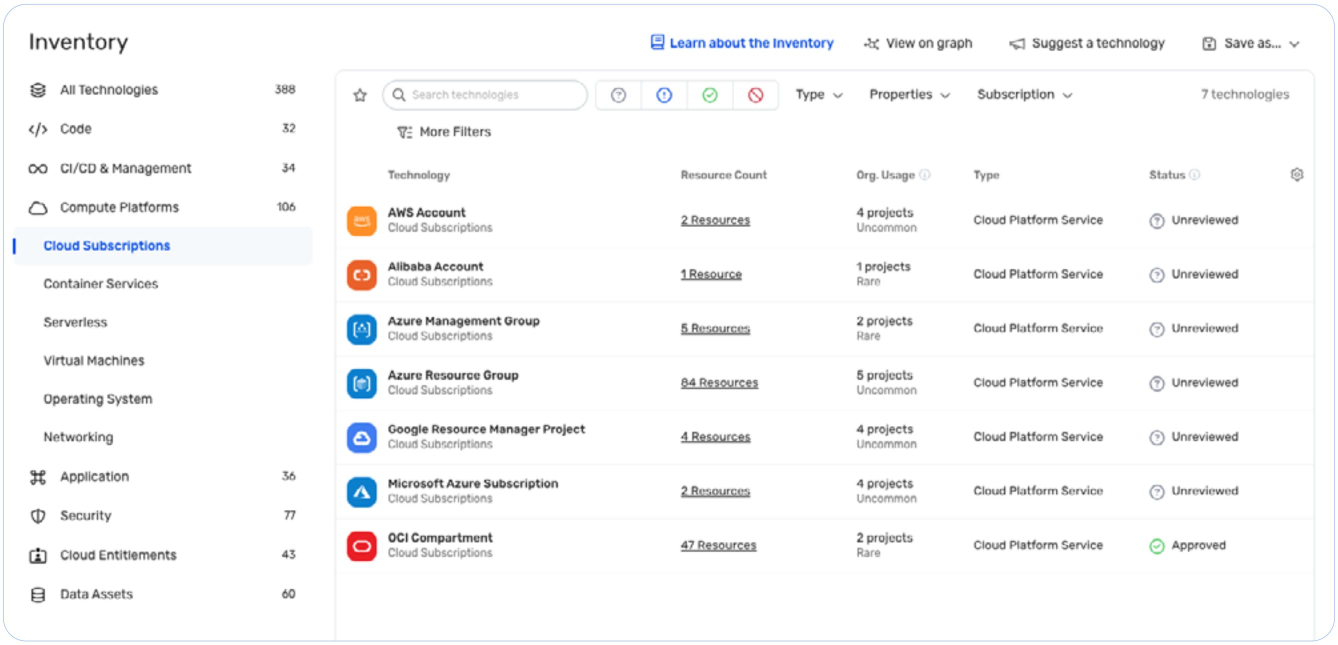
## Let’s get started!

# Deployment Scope

## Cloud subscriptions inventory

Wiz connects to your cloud environment using the cloud provider's APIs and scans your entire technology stack without any agents. Wiz is connected to these cloud environments:

- Amazon Web Services
- Alibaba Cloud
- Microsoft Azure
- Google Cloud Platform
- Oracle Cloud



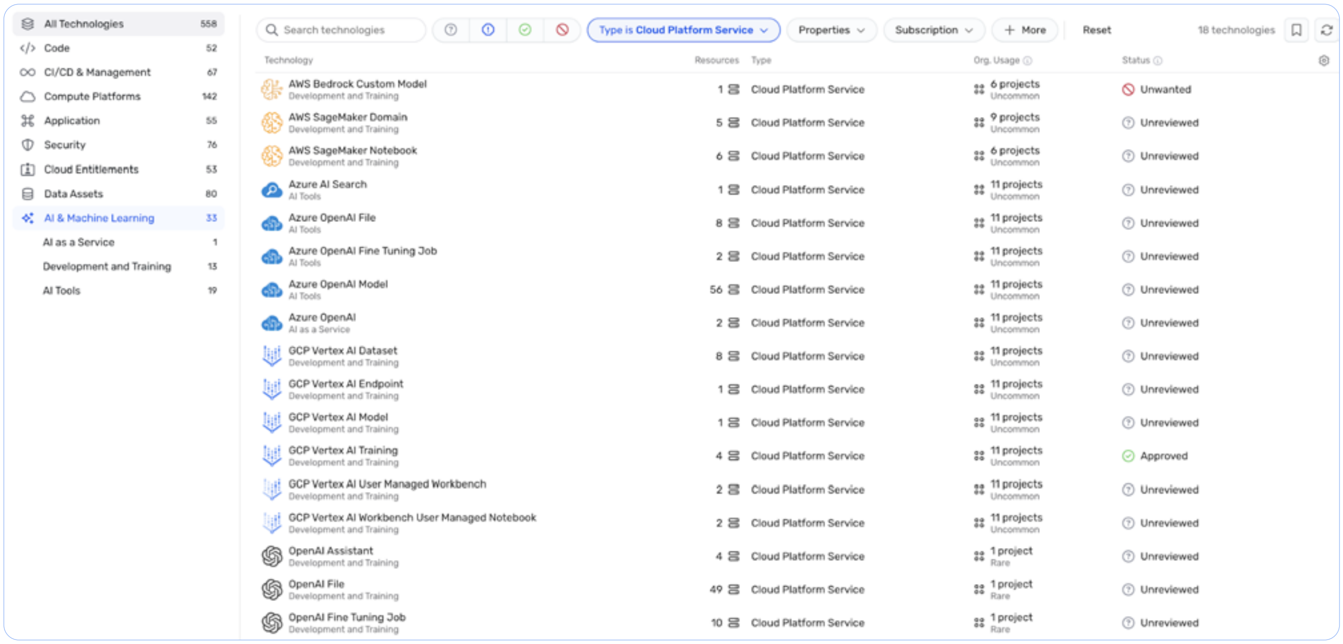
| | |
|---|---|
| All Technologies | 388 |
| Code | 32 |
| CI/CD & Management | 34 |
| Compute Platforms | 106 |
| Cloud Subscriptions | |
| Container Services | |
| Serverless | |
| Virtual Machines | |
| Operating System | |
| Networking | |
| Application | 36 |
| Security | 77 |
| Cloud Entitlements | 43 |
| Data Assets | 60 |

| Technology | Resource Count | Org. Usage | Type | Status |
|---|---|---|---|---|
| AWS Account<br>Cloud Subscriptions | 2 Resources | 4 projects<br>Uncommon | Cloud Platform Service | Unreviewed |
| Alibaba Account<br>Cloud Subscriptions | 1 Resource | 1 projects<br>Rare | Cloud Platform Service | Unreviewed |
| Azure Management Group<br>Cloud Subscriptions | 5 Resources | 2 projects<br>Rare | Cloud Platform Service | Unreviewed |
| Azure Resource Group<br>Cloud Subscriptions | 84 Resources | 5 projects<br>Uncommon | Cloud Platform Service | Unreviewed |
| Google Resource Manager Project<br>Cloud Subscriptions | 4 Resources | 4 projects<br>Uncommon | Cloud Platform Service | Unreviewed |
| Microsoft Azure Subscription<br>Cloud Subscriptions | 2 Resources | 4 projects<br>Uncommon | Cloud Platform Service | Unreviewed |
| OCI Compartment<br>Cloud Subscriptions | 47 Resources | 2 projects<br>Rare | Cloud Platform Service | Approved |

# Overview of the AI technologies in your environment

## Managed AI services

There are 18 managed AI technologies in use in your cloud environment

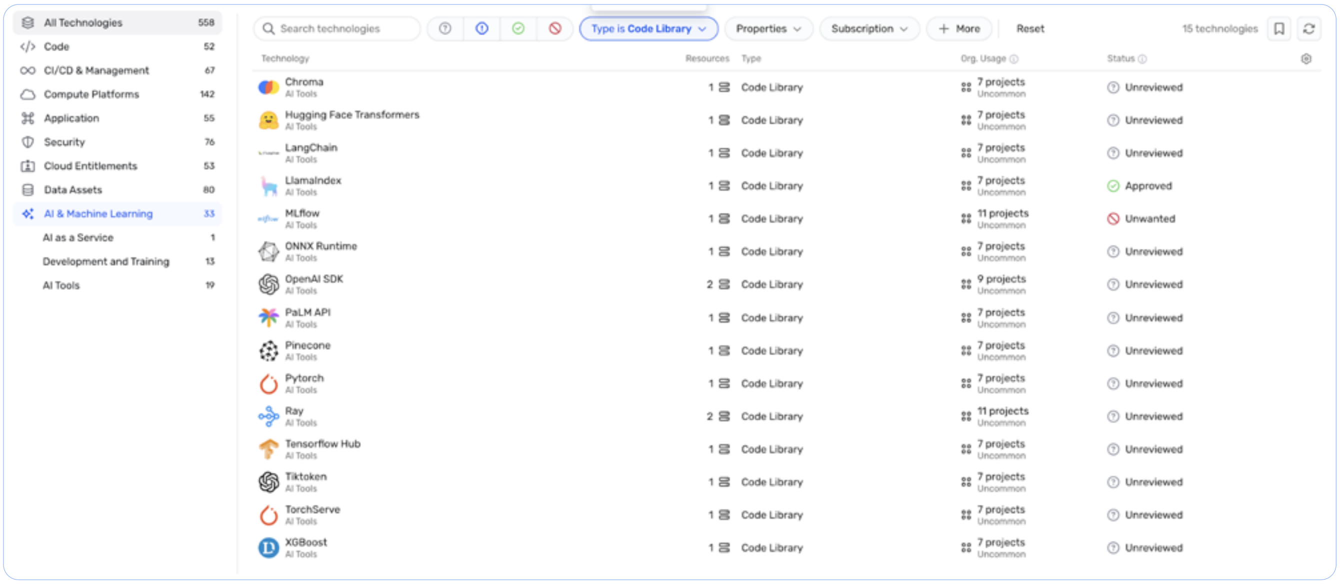


| | |
|---|---|
| All Technologies | 558 |
| Code | 52 |
| CI/CD & Management | 67 |
| Compute Platforms | 142 |
| Application | 55 |
| Security | 76 |
| Cloud Entitlements | 53 |
| Data Assets | 80 |
| AI & Machine Learning | 33 |
| AI as a Service | 1 |
| Development and Training | 13 |
| AI Tools | 19 |

| Technology | Resources | Type | Org. Usage | Status |
|---|---|---|---|---|
| AWS Bedrock Custom Model<br>Development and Training | 1 | Cloud Platform Service | 6 projects<br>Uncommon | Unwanted |
| AWS SageMaker Domain<br>Development and Training | 5 | Cloud Platform Service | 9 projects<br>Uncommon | Unreviewed |
| AWS SageMaker Notebook<br>Development and Training | 6 | Cloud Platform Service | 6 projects<br>Uncommon | Unreviewed |
| Azure AI Search<br>AI Tools | 1 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| Azure OpenAI File<br>AI Tools | 8 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| Azure OpenAI Fine Tuning Job<br>AI Tools | 2 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| Azure OpenAI Model<br>AI Tools | 56 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| Azure OpenAI<br>AI as a Service | 2 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| GCP Vertex AI Dataset<br>Development and Training | 8 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| GCP Vertex AI Endpoint<br>Development and Training | 1 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| GCP Vertex AI Model<br>Development and Training | 1 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| GCP Vertex AI Training<br>Development and Training | 4 | Cloud Platform Service | 11 projects<br>Uncommon | Approved |
| GCP Vertex AI User Managed Workbench<br>Development and Training | 2 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| GCP Vertex AI Workbench User Managed Notebook<br>Development and Training | 2 | Cloud Platform Service | 11 projects<br>Uncommon | Unreviewed |
| OpenAI Assistant<br>Development and Training | 4 | Cloud Platform Service | 1 project<br>Rare | Unreviewed |
| OpenAI File<br>Development and Training | 49 | Cloud Platform Service | 1 project<br>Rare | Unreviewed |
| OpenAI Fine Tuning Job<br>Development and Training | 10 | Cloud Platform Service | 1 project<br>Rare | Unreviewed |

## Hosted AI technologies

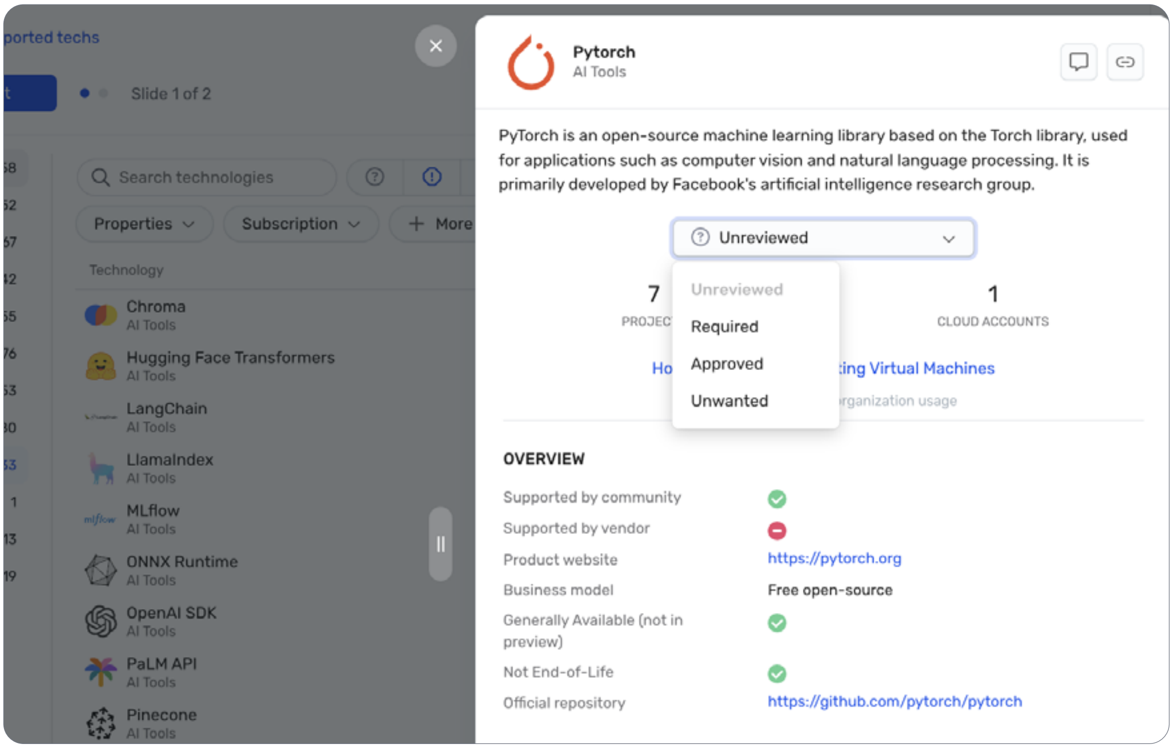
You have 15 different AI hosted technologies in your environment



| | |
|---|---|
| All Technologies | 558 |
| Code | 52 |
| CI/CD & Management | 67 |
| Compute Platforms | 142 |
| Application | 55 |
| Security | 76 |
| Cloud Entitlements | 53 |
| Data Assets | 80 |
| AI & Machine Learning | 33 |
| AI as a Service | 1 |
| Development and Training | 13 |
| AI Tools | 19 |

| Technology | Resources | Type | Org. Usage | Status |
|---|---|---|---|---|
| Chroma<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| Hugging Face Transformers<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| LangChain<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| LlamaIndex<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Approved |
| MLflow<br>AI Tools | 1 | Code Library | 11 projects<br>Uncommon | Unwanted |
| ONNX Runtime<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| OpenAI SDK<br>AI Tools | 2 | Code Library | 9 projects<br>Uncommon | Unreviewed |
| PaLM API<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| Pinecone<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| Pytorch<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| Ray<br>AI Tools | 2 | Code Library | 11 projects<br>Uncommon | Unreviewed |
| Tensorflow Hub<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| Tiktoken<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| TorchServe<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |
| XGBoost<br>AI Tools | 1 | Code Library | 7 projects<br>Uncommon | Unreviewed |

As data scientists introduce new technologies to the environment, you can review them to and mark them as Approved or Unwanted



**Pytorch**
AI Tools

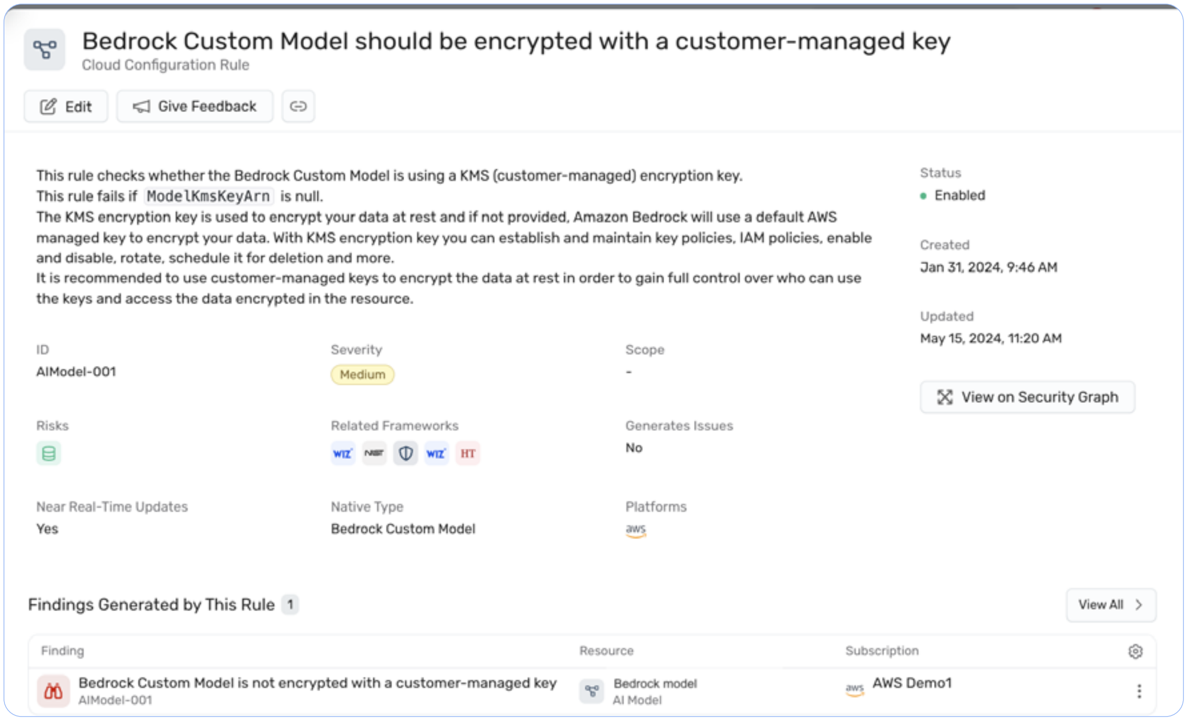
PyTorch is an open-source machine learning library based on the Torch library, used for applications such as computer vision and natural language processing. It is primarily developed by Facebook's artificial intelligence research group.

**OVERVIEW**

|  |  |
|---|---|
| Supported by community | ✓ |
| Supported by vendor | ⊖ |
| Product website | https://pytorch.org |
| Business model | Free open-source |
| Generally Available (not in preview) | ✓ |
| Not End-of-Life | ✓ |
| Official repository | https://github.com/pytorch/pytorch |

# Overview of AI-SPM configuration rules

Wiz assesses your AI services for misconfigurations to help you ensure they follow security best practices. You have a configuration finding for Amazon Bedrock Custom Model



Bedrock Custom Model should be encrypted with a customer-managed key

Cloud Configuration Rule

This rule checks whether the Bedrock Custom Model is using a KMS (customer-managed) encryption key.
This rule fails if `ModelKmsKeyArn` is null.
The KMS encryption key is used to encrypt your data at rest and if not provided, Amazon Bedrock will use a default AWS managed key to encrypt your data. With KMS encryption key you can establish and maintain key policies, IAM policies, enable and disable, rotate, schedule it for deletion and more.
It is recommended to use customer-managed keys to encrypt the data at rest in order to gain full control over who can use the keys and access the data encrypted in the resource.

Status: ● Enabled

Created: Jan 31, 2024, 9:46 AM

Updated: May 15, 2024, 11:20 AM

ID: AIModel-001

Severity: Medium

Scope: -

Risks

Related Frameworks

Generates Issues: No

Near Real-Time Updates: Yes

Native Type: Bedrock Custom Model

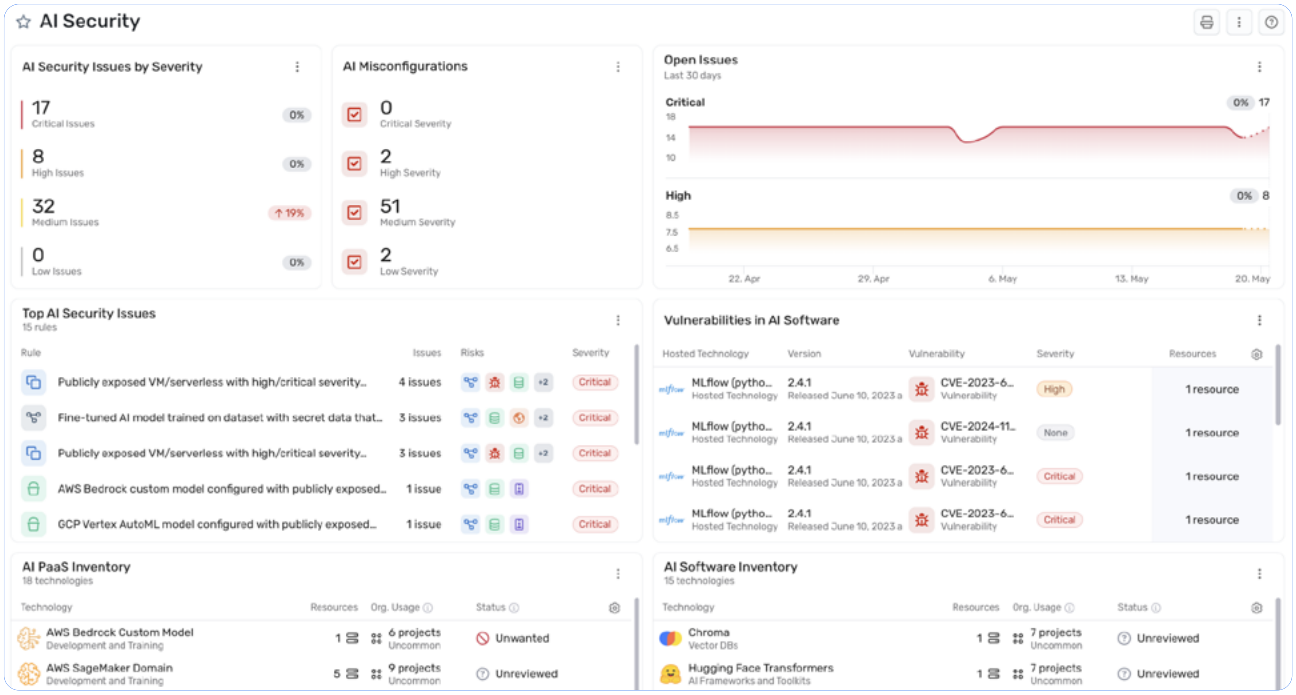
Platforms: aws

Findings Generated by This Rule 1

| Finding | Resource | Subscription |
|---|---|---|
| Bedrock Custom Model is not encrypted with a customer-managed key<br>AIModel-001 | Bedrock model<br>AI Model | aws AWS Demo1 |

# Overview of AI-SPM issues

Wiz runs deep risk assessment of AI pipelines across data, misconfigurations, vulnerabilities, identities, secrets and correlates all AI risks on the Wiz Security Graph. Wiz then identifies Issues in your environment, which are a combination of the different risks that result in an attack path in your AI pipelines. Issues are prioritized based on criticality. You can get a centralized view of your AI security posture with the built-in dashboard:



AI Security Issues by Severity

| | | |
|---|---|---|
| 17 | Critical Issues | 0% |
| 8 | High Issues | 0% |
| 32 | Medium Issues | ↑ 19% |
| 0 | Low Issues | 0% |

AI Misconfigurations

| | |
|---|---|
| 0 | Critical Severity |
| 2 | High Severity |
| 51 | Medium Severity |
| 2 | Low Severity |

Top AI Security Issues
15 rules

| Rule | Issues | Risks | Severity |
|---|---|---|---|
| Publicly exposed VM/serverless with high/critical severity... | 4 issues | +2 | Critical |
| Fine-tuned AI model trained on dataset with secret data that... | 3 issues | +2 | Critical |
| Publicly exposed VM/serverless with high/critical severity... | 3 issues | +2 | Critical |
| AWS Bedrock custom model configured with publicly exposed... | 1 issue | | Critical |
| GCP Vertex AutoML model configured with publicly exposed... | 1 issue | | Critical |

Vulnerabilities in AI Software

| Hosted Technology | Version | Vulnerability | Severity | Resources |
|---|---|---|---|---|
| MLflow (pytho...<br>Hosted Technology | 2.4.1<br>Released June 10, 2023 a | CVE-2023-6...<br>Vulnerability | High | 1 resource |
| MLflow (pytho...<br>Hosted Technology | 2.4.1<br>Released June 10, 2023 a | CVE-2024-11...<br>Vulnerability | None | 1 resource |
| MLflow (pytho...<br>Hosted Technology | 2.4.1<br>Released June 10, 2023 a | CVE-2023-6...<br>Vulnerability | Critical | 1 resource |
| MLflow (pytho...<br>Hosted Technology | 2.4.1<br>Released June 10, 2023 a | CVE-2023-6...<br>Vulnerability | Critical | 1 resource |

AI PaaS Inventory
18 technologies

| Technology | Resources | Org. Usage | Status |
|---|---|---|---|
| AWS Bedrock Custom Model<br>Development and Training | 1 | 6 projects<br>Uncommon | Unwanted |
| AWS SageMaker Domain<br>Development and Training | 5 | 9 projects<br>Uncommon | Unreviewed |

AI Software Inventory
15 technologies

| Technology | Resources | Org. Usage | Status |
|---|---|---|---|
| Chroma<br>Vector DBs | 1 | 7 projects<br>Uncommon | Unreviewed |
| Hugging Face Transformers<br>AI Frameworks and Toolkits | 1 | 7 projects<br>Uncommon | Unreviewed |

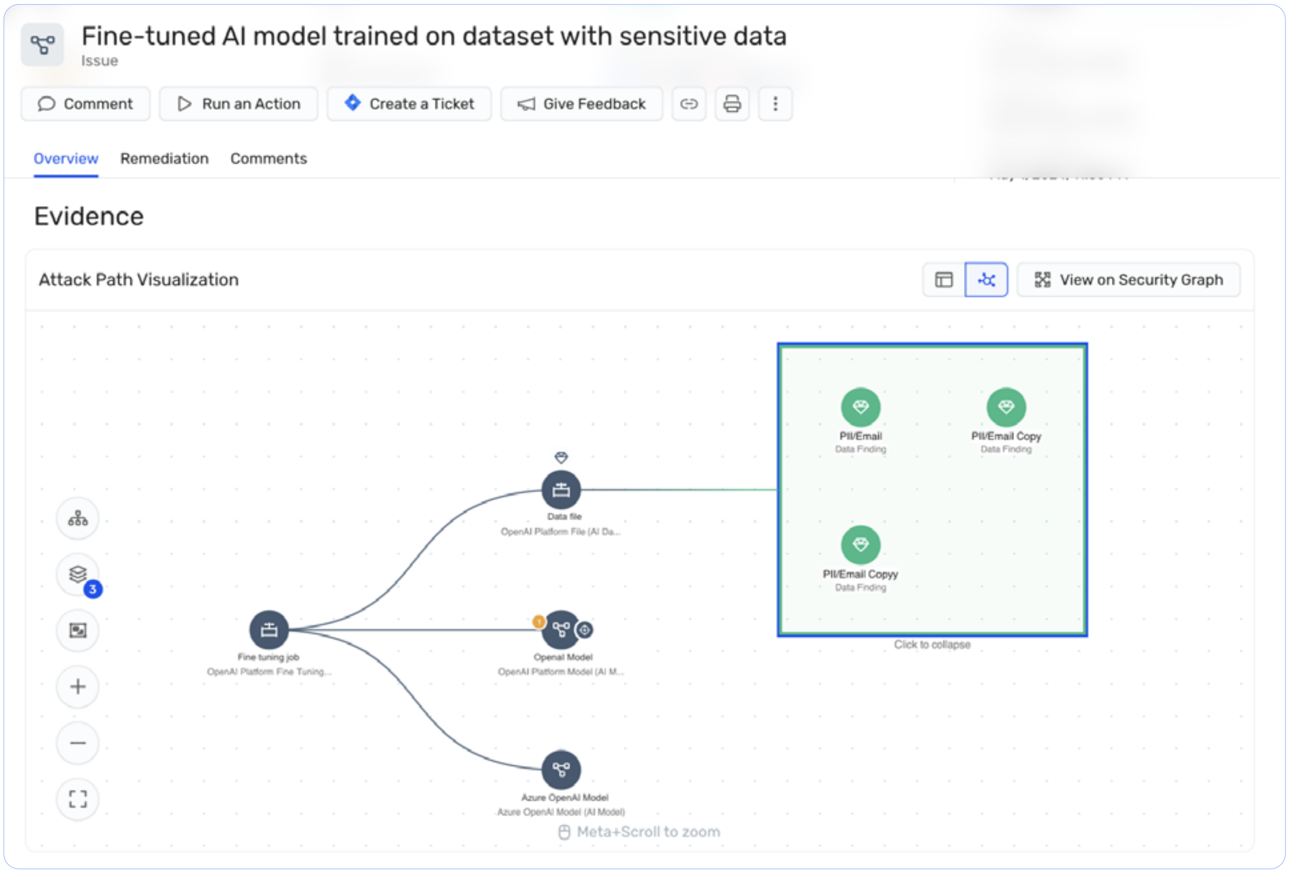
## You have 17 critical issues in your environment

# Issue Examples

1. **Sensitive training data example:** Fine-tuned AI model trained on dataset with sensitive data

## Findings

- Wiz found PII data in an OpenAI data file
- Azure OpenAI model is fine-tuned on that data file



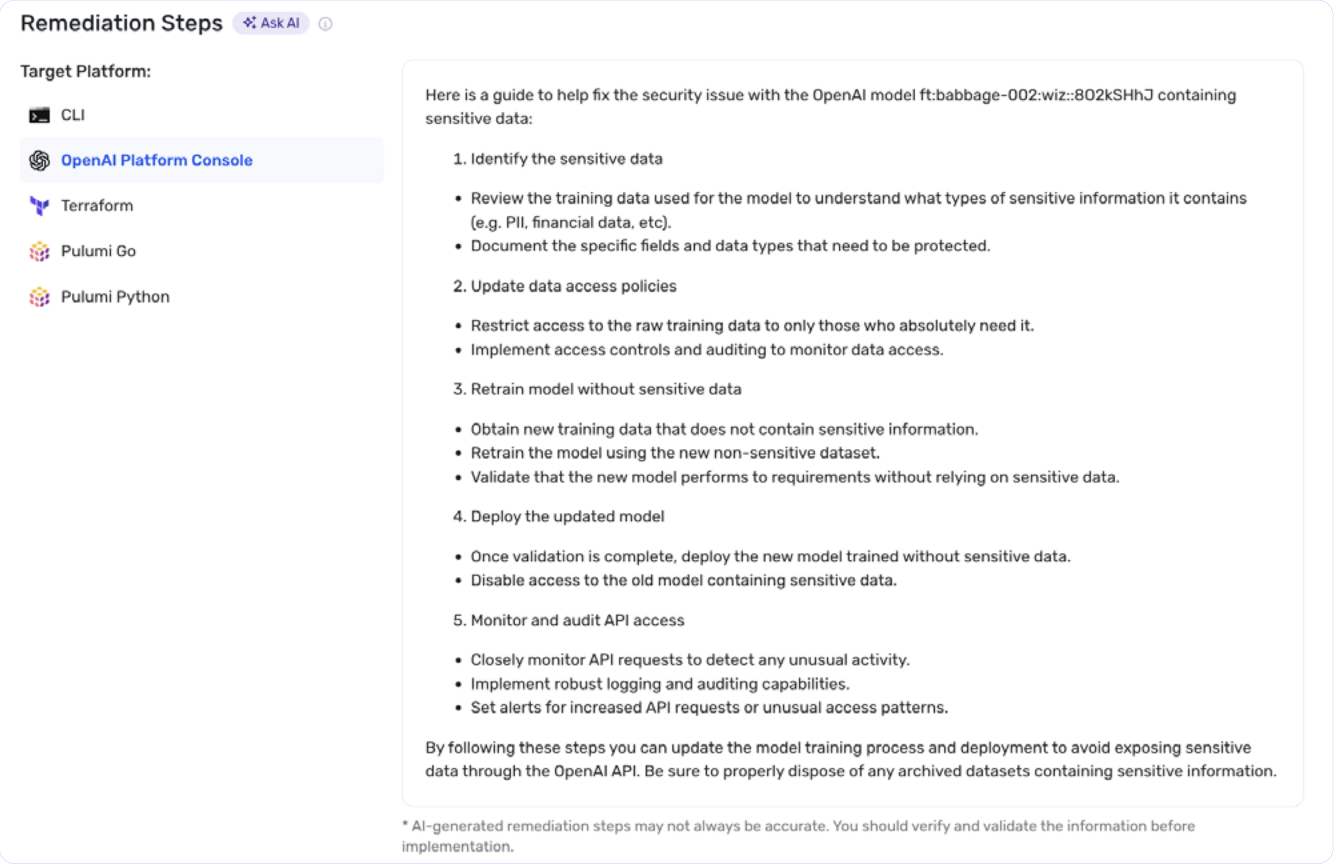
Fine-tuned AI model trained on dataset with sensitive data

Issue

Evidence

Attack Path Visualization

Wiz generates remediation steps and allows you to also generate AI-powered remediation guidance for the Issue



Here is a guide to help fix the security issue with the OpenAI model ft:babbage-002:wiz::802kSHhJ containing sensitive data:

1. Identify the sensitive data

- Review the training data used for the model to understand what types of sensitive information it contains (e.g. PII, financial data, etc).
- Document the specific fields and data types that need to be protected.

2. Update data access policies

- Restrict access to the raw training data to only those who absolutely need it.
- Implement access controls and auditing to monitor data access.

3. Retrain model without sensitive data

- Obtain new training data that does not contain sensitive information.
- Retrain the model using the new non-sensitive dataset.
- Validate that the new model performs to requirements without relying on sensitive data.

4. Deploy the updated model

- Once validation is complete, deploy the new model trained without sensitive data.
- Disable access to the old model containing sensitive data.

5. Monitor and audit API access

- Closely monitor API requests to detect any unusual activity.
- Implement robust logging and auditing capabilities.
- Set alerts for increased API requests or unusual access patterns.

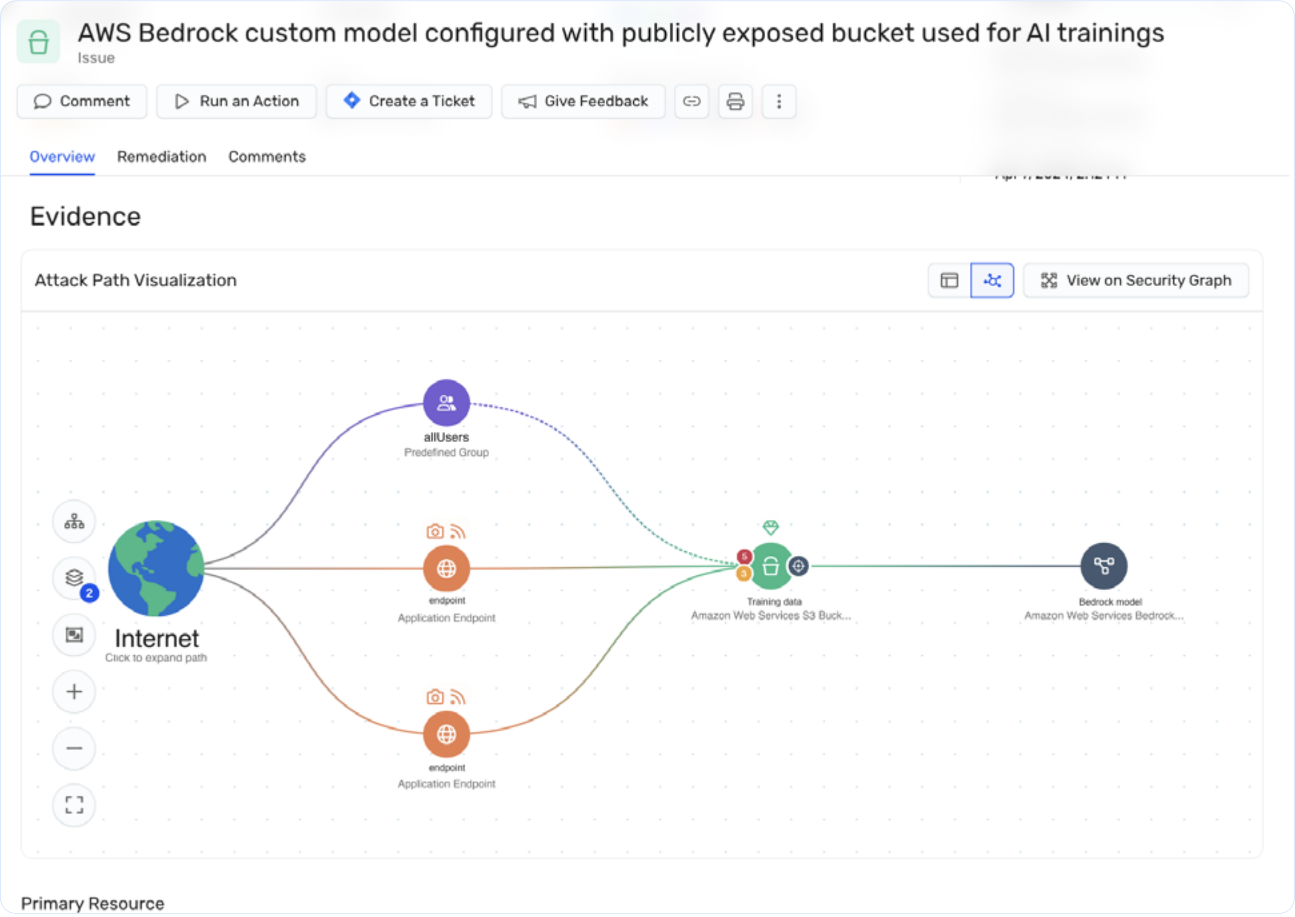
By following these steps you can update the model training process and deployment to avoid exposing sensitive data through the OpenAI API. Be sure to properly dispose of any archived datasets containing sensitive information.

* AI-generated remediation steps may not always be accurate. You should verify and validate the information before implementation.

## 2. **Model poisoning example:** AWS Bedrock custom model configured with publicly exposed bucket used for AI trainings

### Findings

- Amazon Bedrock custom model is fine-tuned on a bucket with sensitive data
- The bucket is exposed to the internet and allows access to all users



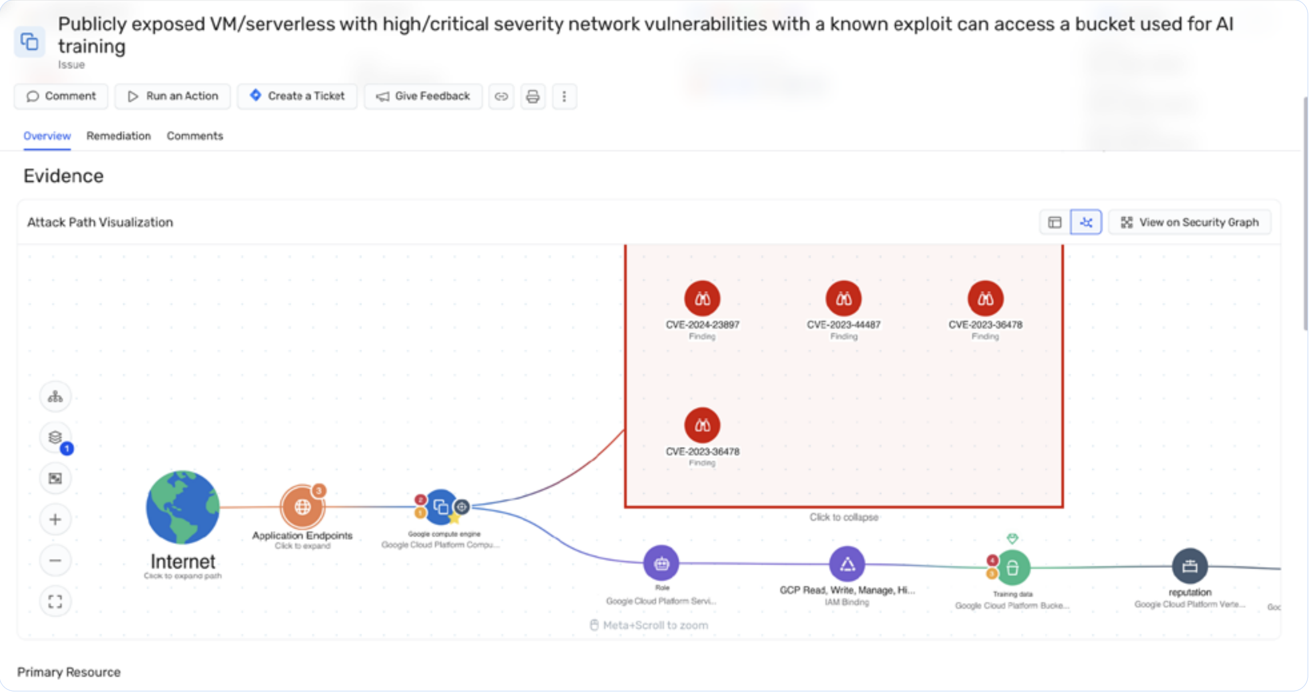
AWS Bedrock custom model configured with publicly exposed bucket used for AI trainings

Issue

## 3. **Toxic combination example:** Publicly exposed VM/ serverless with high/critical severity network vulnerabilities with a known exploit can access a bucket used for AI training

### Findings

- This VM has a public internet exposure path
- There are 7 critical vulnerabilities found on the virtual machine
- The role associated to the machine has access to storage bucket with sensitive data
- The bucket is used to store training data for a Vertex AI model



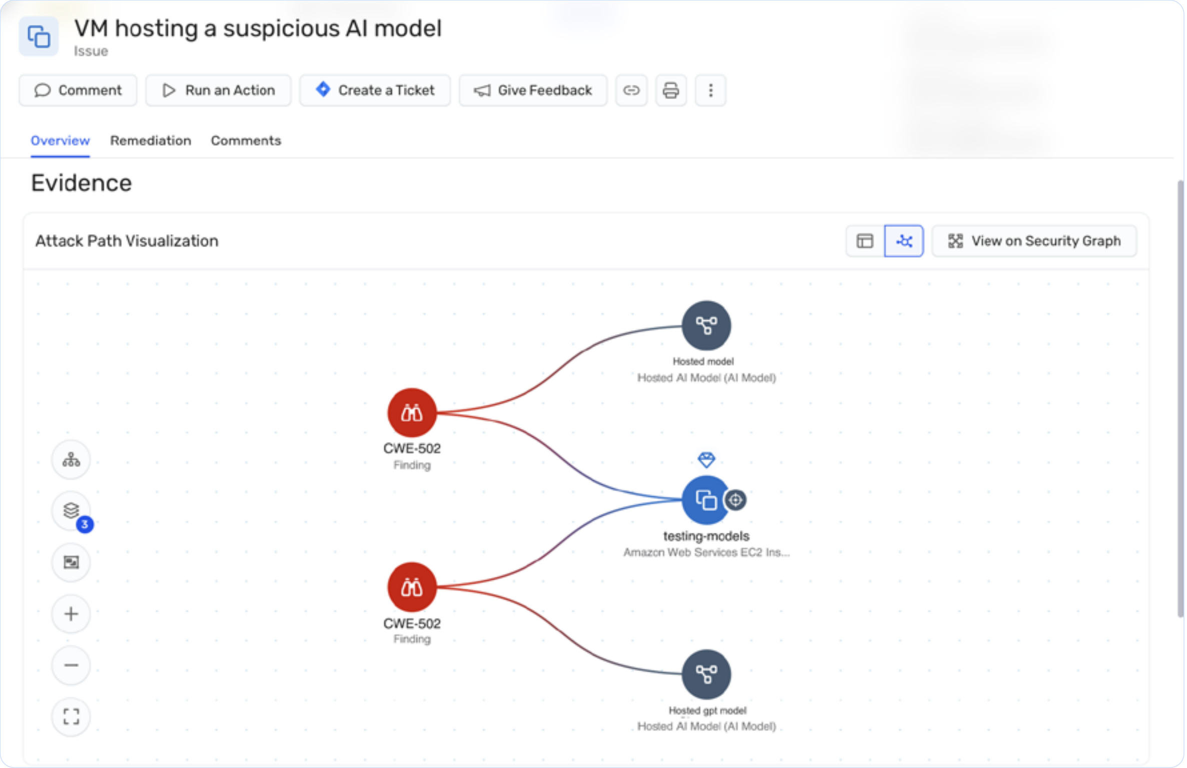
Publicly exposed VM/serverless with high/critical severity network vulnerabilities with a known exploit can access a bucket used for AI training

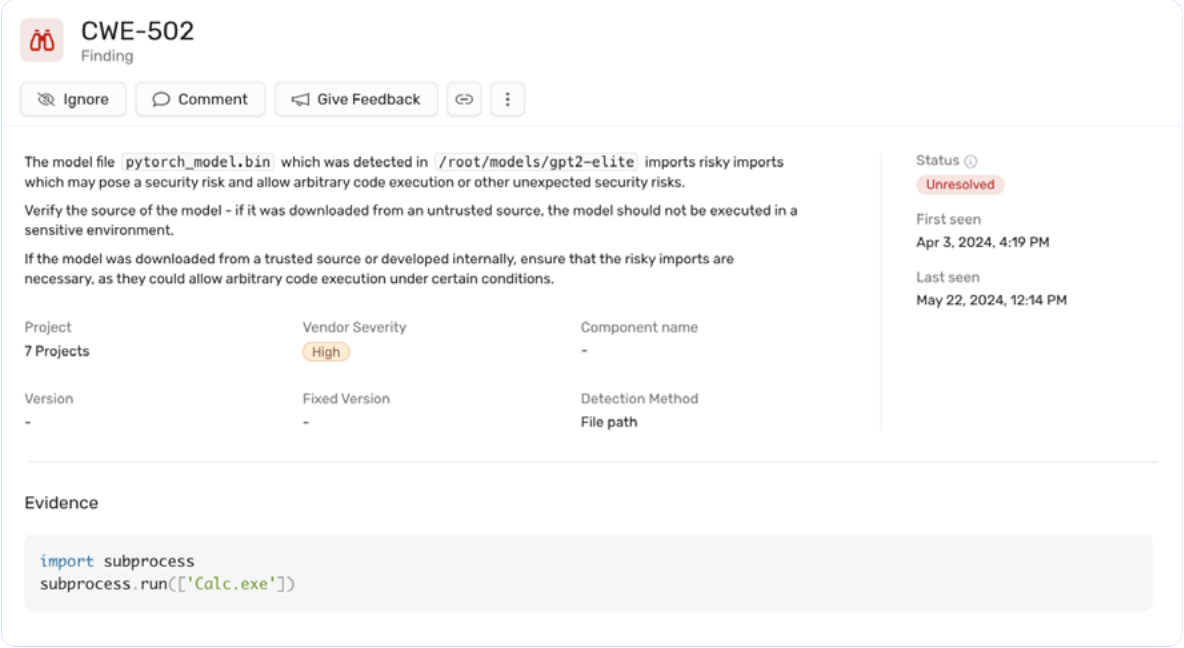
Issue

## 4. **Malicious model example:** VM hosting a suspicious model

### Findings

- There is an EC2 instance that is running two hosted AI models
- The model files have risky imports that can allow arbitrary code execution or other unexpected security risks





The model file `pytorch_model.bin` which was detected in `/root/models/gpt2-elite` imports risky imports which may pose a security risk and allow arbitrary code execution or other unexpected security risks.

Verify the source of the model - if it was downloaded from an untrusted source, the model should not be executed in a sensitive environment.

If the model was downloaded from a trusted source or developed internally, ensure that the risky imports are necessary, as they could allow arbitrary code execution under certain conditions.

Project
7 Projects

Vendor Severity
High

Component name
-

Version
-

Fixed Version
-

Detection Method
File path

Status
Unresolved

First seen
Apr 3, 2024, 4:19 PM

Last seen
May 22, 2024, 12:14 PM

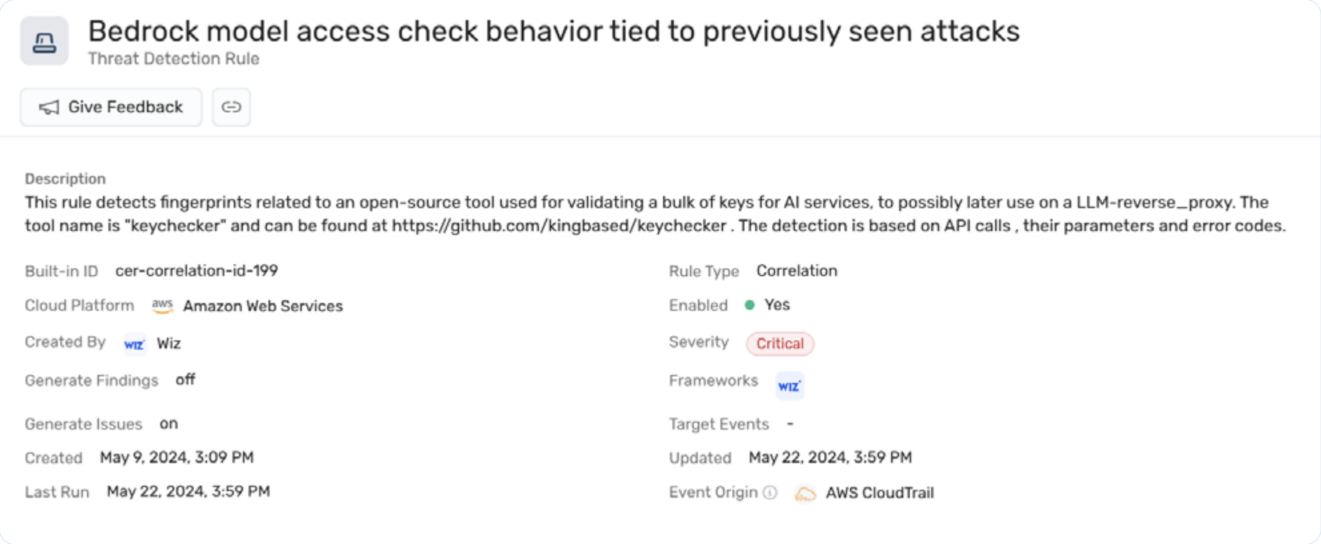
Evidence

```
import subprocess
subprocess.run(['Calc.exe'])
```

### 5. **Example threat detection:**ABedrock model access check behavior tied to previously seen attacks

Findings

- Wiz detected a suspicious behavior in Amazon Bedrock model access in near real-time



Bedrock model access check behavior tied to previously seen attacks

Threat Detection Rule

Description

This rule detects fingerprints related to an open-source tool used for validating a bulk of keys for AI services, to possibly later use on a LLM-reverse_proxy. The tool name is "keychecker" and can be found at https://github.com/kingbased/keychecker . The detection is based on API calls , their parameters and error codes.

| | | | |
|---|---|---|---|
| Built-in ID | cer-correlation-id-199 | Rule Type | Correlation |
| Cloud Platform | Amazon Web Services | Enabled | ● Yes |
| Created By | Wiz | Severity | Critical |
| Generate Findings | off | Frameworks | WIZ |
| Generate Issues | on | Target Events | - |
| Created | May 9, 2024, 3:09 PM | Updated | May 22, 2024, 3:59 PM |
| Last Run | May 22, 2024, 3:59 PM | Event Origin | AWS CloudTrail |

## Next Steps

Get started now with Wiz AI-SPM to see your own AI security insights, we would love to connect with you over a live demo.