




powered by 

How Cyera Enhances Data Security for *Amazon Bedrock*

WRITTEN BY

Yuri Duchovny,
Director of Solution Architecture, *Cyera*

Michael Koyfman,
Head of Global Solution Architecture, *Cyera*

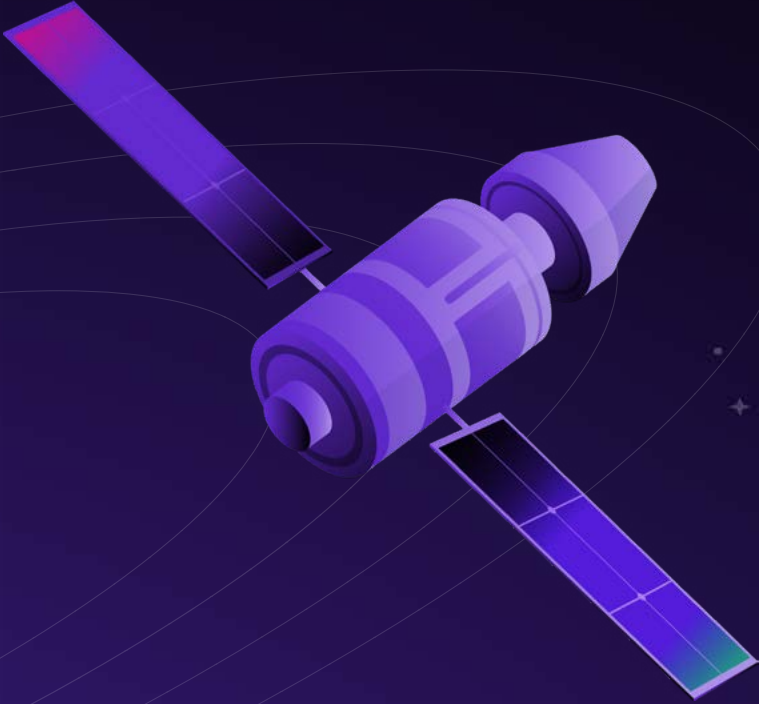
White Paper

Secure AI Adoption Starts with Data Visibility

Generative AI is revolutionizing enterprise applications, enabling organizations to solve complex problems and unlock innovative use cases powered by large language models (LLMs). However, these advancements come with unique security risks associated with managing data rapidly growing in dynamic generative AI environments.

For organizations using Amazon Bedrock to build and deploy LLM-powered applications, gaining visibility into sensitive data and its associated risks is critical. In this paper, we discuss how Cyera is purpose-built to tackle these challenges, offering advanced data security capabilities to mitigate vulnerabilities like Data Poisoning and Sensitive Information Disclosure, as outlined in the OWASP Top 10 for LLMs.





What is Amazon Bedrock?

Amazon Bedrock is a fully managed service that provides access to a selection of high-performing foundation models (FMs) from leading AI innovators, including AI21 Labs, Anthropic, Cohere, Meta, Mistral AI, Stability AI, and Amazon itself, all accessible through a single API. It offers the essential tools and capabilities needed to build generative AI applications with a focus on security, privacy, and responsible AI. To learn more about Amazon Bedrock please follow [this link](#).

Many AWS customers leverage Amazon Bedrock to infuse their applications with generative AI technologies. The platform simplifies experimentation and model selection for specific use cases, enables secure, private customization with domain-specific data using techniques like fine-tuning and Retrieval-Augmented Generation (RAG), and supports the development of agents that integrate seamlessly with existing enterprise systems and data sources to execute tasks.



What is Cyera?

Cyera is a unified data security platform purpose-built to tackle the unique challenges of securing data in LLM applications. It empowers businesses to manage sensitive data across highly permissive and widely distributed environments with precision and efficiency.

The platform's non-invasive, fully automated data discovery provides a comprehensive inventory of sensitive data across structured and unstructured sources. This capability enables organizations to address critical challenges like data proliferation. Powered by AI-driven classification, Cyera goes beyond traditional methods by understanding context, intent, and nuance. This deep insight helps uncover data risks, reduce false positives, and mitigate threats like data breaches and ransomware — areas where conventional data loss prevention and data governance tools fall short.

Seamlessly integrated with AWS, Cyera enables scalable, automated data security without disrupting existing workflows or operations. By combining advanced technology with ease of use, Cyera empowers organizations to confidently secure their data, maintain compliance, and unlock the full potential of LLM applications to drive innovation.



Challenges in Securing LLM Applications

Enriching LLM models with domain-specific data and enabling LLM agents to access that data unlocks the full potential of generative AI applications. However, it also raises legitimate data security concerns. Without proper governance, sensitive information becomes vulnerable to unauthorized access, accidental exposure, and other data-related risks.

The OWASP Top 10 for Large Language Models highlights potential security risks in developing and deploying LLM-based applications. Among these, Data Poisoning and Sensitive Information Disclosure are especially relevant, as they directly involve the data used or accessed by LLM applications.



Data Flows in LLM Applications Built on Amazon Bedrock

To understand how applications on the Amazon Bedrock platform use your data, let's refer to a typical LLM application diagram from the OWASP Top 10 publication.

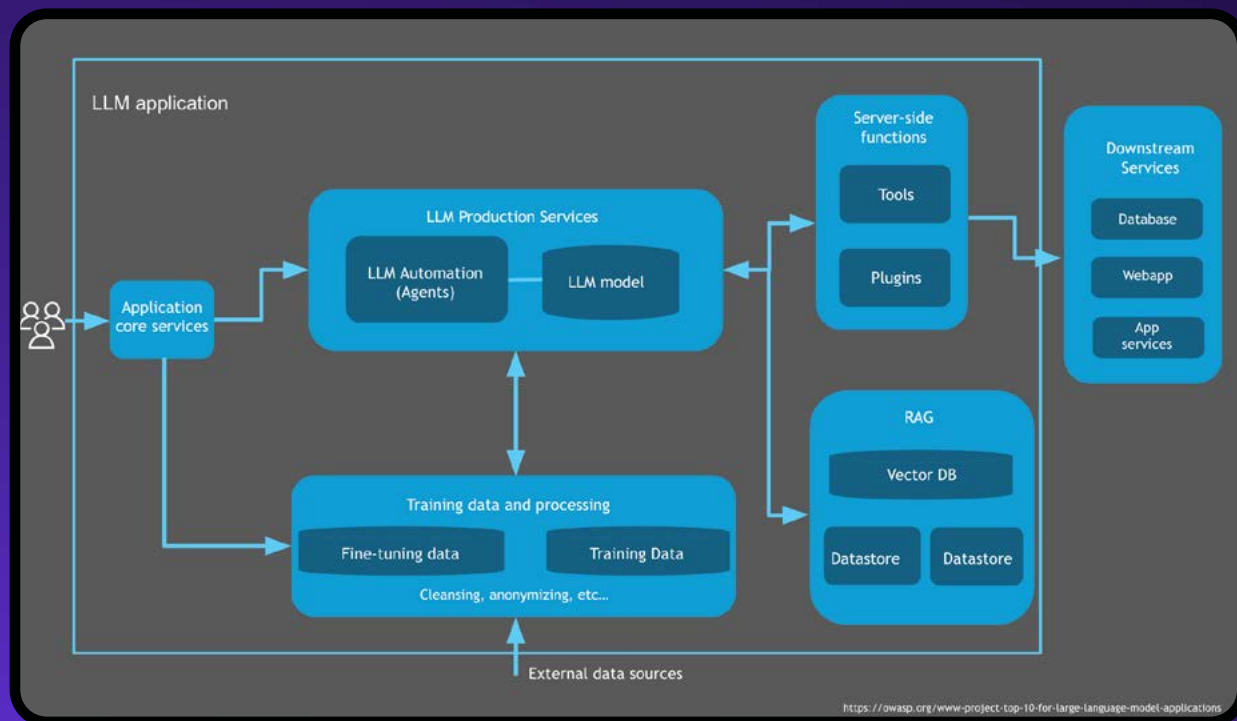


Figure 1: High-Level Architecture of a Hypothetical LLM Application

(Source: OWASP Top 10 for LLM Applications)



In this diagram, data and data processing are fundamental components of any LLM application. Based on your business needs, your use case will fall into one of the five security scopes defined in the [AWS Security Scoping Matrix](#). You may be training your LLM model externally and importing it to Amazon Bedrock, using Bedrock Model Customization to fine-tune or retrain a foundation model, or, as many Bedrock customers do, creating a knowledge base for retrieval-augmented generation (RAG). To maximize the potential of generative AI, you'll likely create LLM Agents on Amazon Bedrock, which connect to downstream data stores and execute some actions on your behalf via Amazon Bedrock Agent Action Groups, essentially functioning as tools in OWASP terminology.

Across these four primary data flows, your LLM-powered application leverages your business data—your most valuable asset. It's critical to ensure that only authorized users can access this data, safeguarding it accordingly.

Due to prompt engineering and other non-deterministic techniques not being effective security measures with today's technology, the rule is straightforward: the LLM model should never determine what data to provide to the end-user. If a user isn't authorized to access specific data, that data should not be sent to the model—simple as that.

To illustrate this, let's explore some examples.



Mitigating Data Risks in LLM Applications

Scenario 1: Preventing Sensitive Information Disclosure in Chatbots

Imagine you're building a publicly accessible chatbot for consumers of your business application. A key priority is ensuring that no private information is inadvertently exposed to the public (LLM06: Sensitive Information Disclosure). How could this happen? Several potential factors could cause this, with the most common stemming from the data flows outlined above.

You'll likely build your chatbot using Agents for Amazon Bedrock. To make the chatbot useful, you'll probably create a Knowledge Base for retrieval-augmented generation (RAG) containing business-specific data. This data is retrieved by the LLM Agent (your chatbot), passed to the LLM model, and the response is sent to the end-user. The primary mitigation technique here is to ensure your Knowledge Base contains no sensitive data that could be leaked to public users. To mitigate this risk, check the data sources for your Knowledge Base—in this case, the Amazon S3 bucket objects used for Knowledge Base injection—and confirm they don't contain private information.



If you're using an Amazon S3 bucket as interim storage and exporting data from another data store, such as Amazon Redshift or Amazon RDS, prior to the Knowledge Base injection job, you can further save time and resources by confirming the data's origin does not contain sensitive information in the first place.

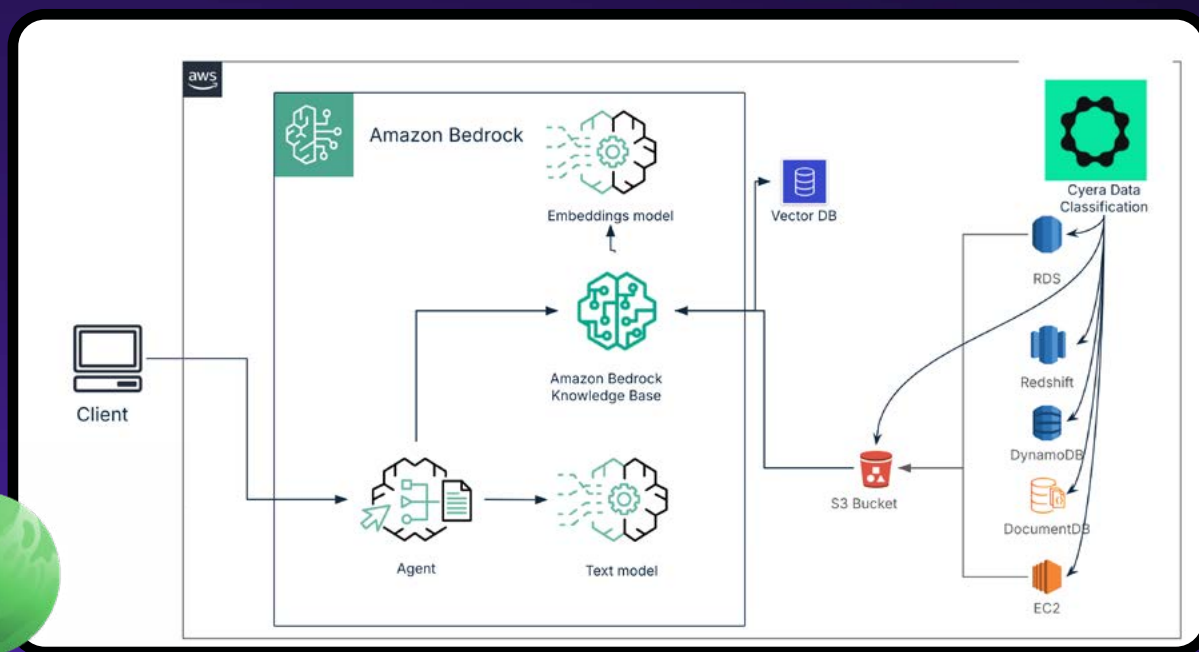


Figure 2: Mitigating Sensitive Information Disclosure in LLM Applications

— Utilizing RAG with Cyera's Data Classification

Additionally, if you add functionality to your chatbot by enabling it to search other data sources via Agent Action Groups represented by AWS Lambda functions, it's essential to ensure that these data stores do not contain sensitive information. As a rule, no information that the user should not see should be sent to the LLM model.

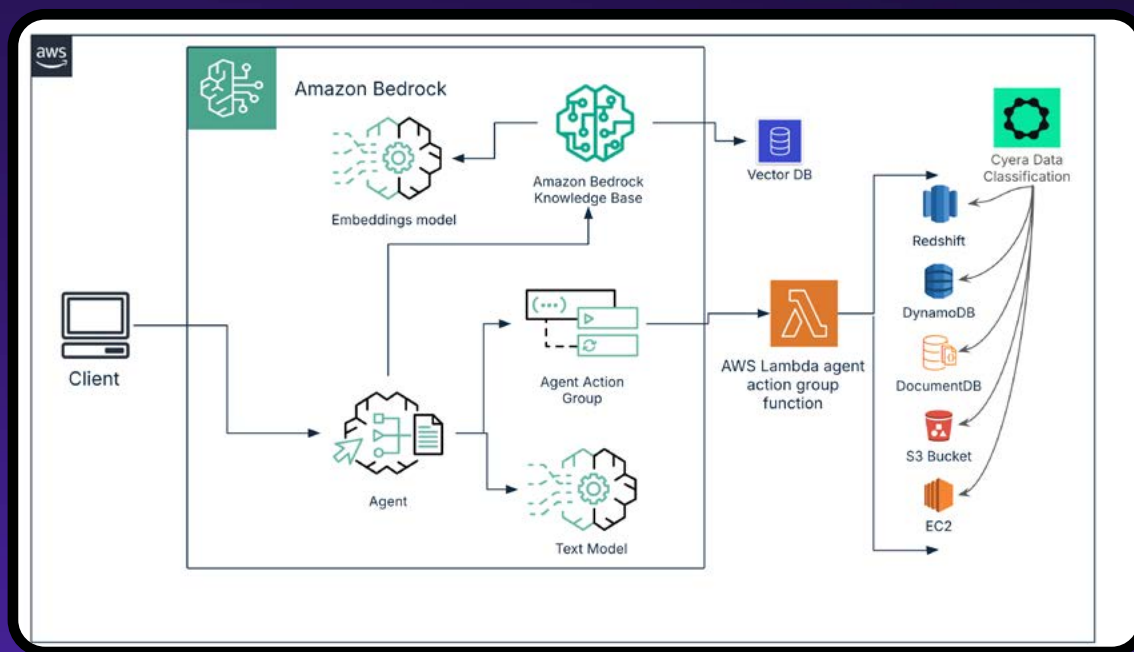


Figure 3: Mitigating Sensitive Information Disclosure in LLM Applications
Querying Heterogeneous Data Sources Using Cyera's Data Classification

Scenario 2: Securing Training Data Against Poisoning

In another example, you may decide to use the Bedrock Model Customization framework to fine-tune or continually retrain an existing foundation model or even to train your own model and import it to Amazon Bedrock. In all these scenarios, it's critical to mitigate the risk of sensitive information appearing in the training dataset (LLM03: Training Data Poisoning), which could potentially lead to information leakage to the end user (LLM06: Sensitive Information Disclosure). This means you need to scan your data, either at the training source Amazon S3 bucket or, as above, at the heterogeneous origin(s) of the data, to ensure it does not contain sensitive or unwanted information.

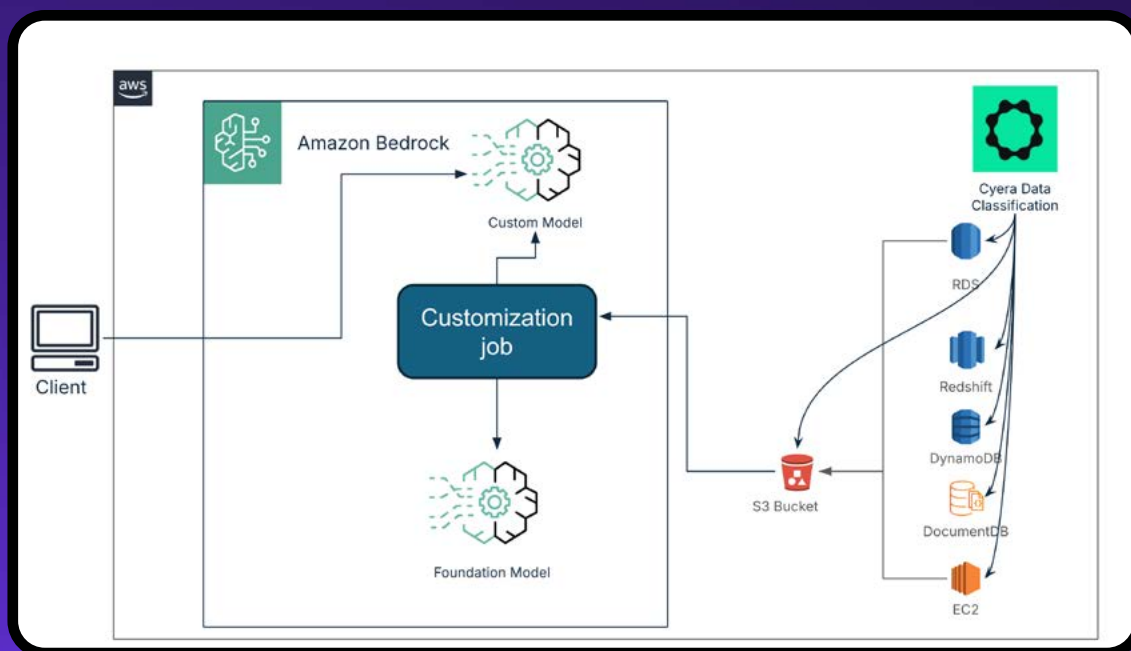


Figure 4: Mitigating Dataset Poisoning in Amazon Bedrock with Cyera's Data Classification



The Challenge of Knowing Your Data

In all these scenarios, the core mitigation technique boils down to one simple principle: **know your data**. Yet, in today's fast-paced, cloud-driven environments, achieving this level of knowledge is anything but simple. Unlike static, legacy systems, modern multi-cloud and hybrid environments are dynamic and constantly evolving, with data continuously created, stored, and moved across multiple platforms. This fluidity makes it difficult to maintain an accurate inventory of data and understand precisely where sensitive information resides. Traditional data classification solutions rely heavily on manual processes, such as tagging, regular expression tuning, and connection strings, which are not only labor-intensive but also error-prone. As data ownership is often distributed across departments, alignment between business stakeholders and technical experts is necessary to classify data accurately—a process that can be costly, complex, and fraught with errors. In this landscape, outdated discovery and classification tools often miss critical data or misclassify it, leading to both data security risks and operational inefficiencies.

Furthermore, legacy data discovery and classification methods lack the context needed for effective security in the cloud. These tools are inventory-based, meaning they often only detect data stores but fail to accurately recognize or categorize the actual data within those stores. Security policies based on these incomplete or outdated classifications result in high false-positive rates, causing alert fatigue and diverting valuable resources away from actual threats. Data classification requires more than merely identifying data—it demands a deep understanding of its value and relevance to the organization. Without the necessary contextual insights, these tools struggle to offer tailored security for different data types, especially sensitive information, which remains vulnerable to unauthorized access and misuse.



Streamlining Security and Compliance with AI-powered Classification

Cyera's platform provides AWS customers with a continuous, cloud-native approach to data classification that addresses the complexities of today's multi-cloud, dynamic environments. Leveraging agentless, API-based technology, Cyera connects to AWS environments with a single read-only AWS IAM role, eliminating the need for resource-intensive manual configurations and constant network access setups. This approach enables AWS customers to automatically and continuously inventory their structured and unstructured data across various data stores, such as Amazon S3 buckets, databases, and data lakes. With fast and accurate visibility into their data landscape, Cyera helps AWS customers maintain an up-to-date understanding of where sensitive data resides, providing a robust foundation for secure data management and minimizing the risk of unauthorized data exposure.

Amazon customers that use Amazon Bedrock don't just need to protect themselves against OWASP Top 10 Vulnerabilities for LLMs. They also need to ensure that their applications are compliant with various regulations and frameworks that require enterprises to classify the data they have in their environments, as failure to do so can lead to fines, brand tarnishment, and legal action, to name a few.



Examples of Frameworks and Regulations Requiring Data Classification

Framework/Regulation	Type of Data	Classification Requirement
GDPR	Personal data of EU residents	Mandatory; classify by data subject type and sensitivity
HIPAA	Protected Health Information (PHI)	Required; identify and classify PHI
PCI DSS	Cardholder data	Required; classify by data type (cardholder data)
FISMA	Government information	Required; based on information and system criticality
CCPA	Personal data of CA residents	Recommended; based on consumer personal data categories
ISO 27001/27002	All information assets	Required; information to be classified per risk level
GLBA	Financial data	Required; to determine appropriate safeguarding practices
SOX	Financial reporting data	Implied; classify by relevance to financial reporting
NIST 800 Series	All data types	Required; particularly for sensitive and high-impact data
CMMC	Controlled Unclassified Information (CUI)	Required; per CUI categories and DoD requirements
FOIA/Privacy Act	Public records, personally identifiable information (PII)	Required; classify based on accessibility and privacy requirements
Data Protection Act (DPA)	Personal data	Required; sensitive data types must be categorized
Australia Privacy Act	Personal information	Recommended; classify sensitive data types separately
Canada's PIPEDA	Personal data	Required; categorize based on sensitivity
Cyber Essentials	All business data	Recommended; classify by risk level



In addition to continuous discovery of data stores and data, Cyera's platform employs AI processing to accurately classify data based on context, enhancing the precision of sensitivity classifications with 95%+ precision. Cyera goes beyond traditional methods that rely on manual tagging or regular expression tuning by analyzing data in its actual usage context, significantly reducing false positives and alert fatigue. For AWS customers, this translates into faster, more accurate risk assessments and streamlined compliance processes, as Cyera's contextualized classification enables security teams to focus on real security concerns rather than noise. Cyera's unique capabilities in learning new data classes in both structured and unstructured data in the customer's environment aligns well with the dynamic needs and nature of cloud-native AWS applications, supporting organizations in meeting compliance requirements while allowing them to scale securely and ensure that all compliance and regulatory requirements are met.



Confidently Secure and Scale LLM Applications with Cyera

So, what does this mean for organizations using Amazon Bedrock?

With Cyera, organizations gain a powerful tool for safeguarding their LLM-powered applications. Cyera's continuous scanning of both origin and Bedrock data sources -including Amazon RDS, Amazon Dynamo DB, Amazon RedShift and Amazon S3 buckets- enables data to be classified accurately. This ensures that datasets used for knowledge base ingestion, LLM model training or customization, or as data sources for LLM Agents are secure and free from unwanted sensitive information.

By leveraging generative AI-powered, automated data classification, Cyera helps customers mitigate risks associated with OWASP Top 10 vulnerabilities, such as Sensitive Information Disclosure and Training Data Poisoning. This advanced capability not only strengthens data security but also simplifies compliance with regulatory frameworks, reducing the risk of fines and reputational harm.

With seamless integration into AWS environments, organizations unlock the full potential of their LLM-powered applications. Customers can confidently innovate, knowing their data is secure, their operations are compliant, and their scalability is supported by a strong foundation of data governance.

Better Together

Learn more about Cyera and AWS: <https://www.cyera.io/partnership/aws>



