TRUSTIBLE™

⬢ databricks

Whitepaper

How Trustible Helps Operationalize the Databricks AI Governance Framework (DAGF)

TRUSTIBLE

Table of Contents

TRUSTIBLE

How Trustible Helps Operationalize the Databricks AI Governance Framework (DAGF)

As AI continues to evolve and transform industries, governance will become increasingly complex. Already, enterprises are overwhelmed by the pace of innovation, by the increasing compute and data needs for AI use cases, and by a fast-changing regulatory environment. Infrastructure platforms are beginning to offer organizations native tools for data engineering, AI development, and infrastructure, and enterprise-grade security. However, these deep technical capabilities need to be complemented by proper governance of the models and systems being developed, the teams deploying them, and the people using them. To help with these governance concerns, Databricks has developed its own best-in-class Databricks AI Governance Framework (DAGF) that outlines the key things organizations need to consider for AI governance, and how Databricks' platform helps with many of those concerns, especially for the technical resources involved.

However, AI governance involves more than just the technical personas. Business owners, legal, and risk/compliance professionals are increasingly getting involved in AI, especially as businesses adopt it for more critical projects, and additional regulations for AI come into force.

This is where Trustible comes in. Trustible enables organizations to adopt trustworthy and responsible AI practices, assess AI risk, and comply with AI regulations. Our platform blends best in class human expertise in machine learning and policy with the simplicity and ease of use of technology to accelerate the time to value of an organization's AI stack, measure and quantify multiple dimensions of AI risk, and align AI users and AI stakeholders together within organizations in one, unified platform.

Here's a quick breakdown of Trustible's core features and capabilities

AI Inventory

Understanding how and where AI is being used is a necessary first step towards effective governance. Trustible's AI Inventory helps organizations create a single source of truth for their AI use cases, models, and vendors. Trustible's platform helps organizations build and maintain their inventory by scanning ModelOps tools, cloud platforms, vendor management systems, and organization-wide registration forms.

Automated Risk Management

Trustible offers in-depth taxonomies of AI risks, mitigations, benefits, and evaluations to help organizations understand which systems pose more risks and constantly implement best practices for mitigating those risks. This helps organizations navigate the rapidly changing environment of AI research, regulatory guidance, and evolving value chain.

Organizational Workflows

Trustible's workflows offer out-of-the-box sequences of tasks for common AI governance activities, such as triaging proposed AI use cases or conducting impact assessments. These workflows help organizations solve the problem of 'who' needs to do 'what' across the AI development lifecycle.

AI Policy Center

Trustible helps organizations early in their AI governance journey with out-of-the-box policy templates written by Trustible's AI Legal Insights team that are aligned with the NIST AI RMF, ISO 42001, and EU AI Act. In addition, Trustible analyzes any supplementary policies to ensure they conform and align with regulatory expectations.

Regulatory Mappings

Trustible reads every AI-relevant law, and maps them to a standardized AI control set so that organizations only need to do things once to comply with multiple standards and frameworks. In addition, each control in Trustible has automated integrations to help ensure organizations stay compliant across their entire AI stack.

AI Governance Insights

Trustible's AI Insights team reads every new relevant AI research paper, evaluates leading models, analyzes the latest regulatory guidance, and creates actionable guidance for AI governance teams. Trustible's AI Insights helps organizations proactively monitor for new potential AI risks, mitigation best practices, or technological breakthroughs.

DAGF +Trustible

AI Governance is as much about enabling and promoting AI use as it is about ensuring regulatory compliance and risk management. One of the biggest challenges for many AI teams is building trust between their AI tools and their users. The Databricks AI Governance Framework outlines a clear path for how organizations can get started on this trust-building exercise, and integrate it directly into their ML environment. Databricks Mosaic AI has leading tools for all the technical aspects of AI, ranging from ensuring models are trained on high-quality data to acting as an AI gateway to monitor relevant model inputs and outputs. Meanwhile, Trustible helps organizations operationalize all of the most tedious aspects of AI governance, including all the paperwork, documentation, and legal processes needed to satisfy a broad set of internal stakeholders. Together, the DAGF and Trustible help organizations achieve their strategic AI and business goals and enable the wide adoption of trustworthy and responsible AI.

In this whitepaper, we'll see how Trustible operationalizes the Databricks AI Governance Framework, and how they complement each other for organizations building with Databricks.

Frameworks / Databricks AI Governance Framework

Databricks AI Governance Fra

AI Governance Framework made by Databricks

Readiness   **Articles**

Databricks AI Governance Framew

3 / 20 articles have policy alignment requirements

17 articles need review

TRUSTIBLE.

AI Organizations

The first pillar of the DAGF is all about the organization's people and processes needed for AI governance. It focuses on the actual structure for AI oversight (centralized vs. decentralized), what processes or workflows need to be in place, and how to capture and mitigate the risks of AI systems. The first pillar outlines 12 key areas for organizations to consider, all of which Trustible helps organizations implement.

1   Business Alignment

Business Alignment is about ensuring that the organization has clear goals for its AI use and that those goals are clearly aligned with the organization's overall goals. This is relevant at the strategic level, as well as the tactical per-use case level.

How Trustible Helps:

Policy Alignment
Trustible offers out-of-the-box AI policy templates and analysis to ensure that an organization's AI policies align with both internal business goals and regulatory requirements.

Intake Workflow & Benefits Taxonomy
Trustible's guided use case intake workflow helps teams quickly identify key business requirements for an AI system, assess its risks, and document its intended benefits. Each benefit can be tied to KPIs and aligned with business goals.

Dashboard
Trustible's AI Governance dashboard helps inform executives about key information regarding AI initiatives and the organization's governance status. This includes a breakdown of AI projects by deployment status, a heatmap of AI use cases by overall benefit and risk level, and statistics on major AI risks the organization is facing.

2   Governance Model

Every organization has a different structure, AI goals, and level of risk tolerance. Therefore, the AI governance structure inside each organization may be slightly different, ranging from having a centralized department or committee that assesses every AI system, to having a distributed system with each team responsible for their own use or development of AI.

How Trustible Helps

Centralized Inventory
Trustible helps organizations build and maintain a centralized inventory of their AI use cases, models, and vendors to support internal visibility regardless of the governance structure. Each item in the inventory can be assigned to any number of owners, departments, or teams to help track who is accountable for each system.

Workflows
Trustible's AI governance workflows can accommodate a wide range of governance structures and risk escalation policies. Trustible can help ensure every AI system is reviewed, or can be configured to escalate only high-risk AI use cases to an AI governance committee.

3   Governance Oversight

Regardless of the operational structure for AI governance, there needs to be the 'governance of the governance' as well. In other words, who is responsible for overseeing the governance program, and who is accountable to them, including whether there is any board-level oversight.

How Trustible Helps

AI Policies
Trustible's AI policy templates include recommendations for the appropriate C-suite or board-level oversight of AI governance required by many AI governance standards. Trustible's AI policy analyzer can also help organizations align their policies with new regulations or standards that emerge, and can generate reports about the potential gaps.

Dashboard & Reporting
Trustible's dashboarding feature allows organizations to track key AI governance metrics, ranging from the number of AI use cases to an overview of risk management, to how quickly the organization conducts AI governance workflows. This information can be exported into a ready-to-share report for regular reporting to organizational leaders.

4   Guiding Values

Given the pace of AI development, many organizations are very likely to encounter new ethical dilemmas, AI use cases, and technologies. It's impossible to have a pre-defined policy for every situation; therefore, organizations need to have some clear guiding principles and values that help them deliberate on all the novel issues, especially in situations without perfect information.

How Trustible Helps

Benefits & Risk Taxonomy
It is difficult to make ethical decisions without a full understanding of a system's benefits compared to its risks. Trustible's out-of-the-box benefits and risk taxonomies can help organizations identify, measure, and track whether a system's benefits outweigh its risks.

Stakeholder Taxonomy & Impact Assessments
AI systems can affect different groups in positive and negative ways, which is why many AI standards require AI impact assessments for high-risk uses of AI. Conducting these assessments is a key way of understanding whether an AI system will violate any of the proposed ethical guidelines or values defined by an organization.

5   Strategy

AI without adequate justification or purpose is not responsible. AI is a powerful tool that can yield widespread benefits for an organization, but it can also rack up a lot of costs, damage an organization's brand, and impact company morale. Organizations should have a clear strategy for why they will adopt AI and how they will measure whether those initiatives are being successful or not.

How Trustible Helps

Business Case Builder & ROI Calculator
Trustible's default intake workflow helps organizations build a business case for each AI use case, including a basic assessment of the technological feasibility, user demand, and intended benefits of the system. Organizations that want to go further can also fill out and track information on Trustible's AI ROI Calculator to assess whether the intended benefits will outweigh the potential costs.

6 Roles and Responsibilities

Many stakeholders need to be involved with AI governance, ranging from business leaders and subject matter experts to the technical development team to the legal, risk, and compliance team members who may advise on governance requirements or impacts of an AI system. It's essential that each of the roles involved with governance is clearly defined and understood.

How Trustible Helps

AI Policy Analyzer
Trustible's AI policy analyzer can scan any organization's AI policy and ensure that the sum of its policies aligns with AI standards and regulations. This includes ensuring that clear roles are appointed for AI oversight, monitoring, and approvals.

Workflows
Trustible's workflows are intended to help organizations figure out 'who' needs to do 'what' at each stage of an AI project. Essentially, our workflows help organizations operationalize each distinct role involved with AI governance, and can direct each role on what they have to do, whether it's filling out system documentation, testing a model, or reviewing and approving AI proposals.

AI Literacy
If your organization wants to have 'humans in the loop' for an AI system, then it's essential that those humans have a sufficient understanding of AI to recognize a system's limitations and when it may not be functioning as expected. Trustible has AI literacy modules to help educate non-technical staff members about key concepts of AI and key AI risks and how they contribute to mitigating them.

7 Policies

An organization's AI policies often define how it will implement AI governance. These policies may include defining how AI systems should be developed, how to handle privacy and cyber risks, and how incidents or system feedback is reported.

How Trustible Helps

AI Policies & Analyzer
Trustible has out-of-the-box AI policy templates aligned with major standards and regulations, like the NIST AI RM and EU AI Act, to help organizations start with a good foundation of governance. For organizations with pre-existing policies, Trustible's AI policy analyzer can automatically scan policies and map them to relevant AI governance controls.

8 Standards

Organizations developing or deploying AI tools should have clear standards about acceptable performance for each AI system. This should be validated through automated or manual testing.

How Trustible Helps

Model Evaluations Taxonomy — Trustible helps organizations identify which set of tests, benchmarks, or evaluations is most appropriate for different AI use cases. For each evaluation, Trustible offers actionable guidance and tool recommendations to help organizations implement them. and integrates into leading ML platforms such as Databricks to help them run on their models.

9 Processes and Procedures

AI governance potentially involves many discrete tasks, ranging from writing highly technical documentation to conducting widespread impact assessments to simply getting any number of sign-offs and approvals for an AI system. Organizations should define clear processes for conducting each of these tasks and ensure that anyone involved with AI development understands them.

How Trustible Helps

Workflows
Trustible is the platform for operationalizing AI governance. It offers pre-configured workflows for common AI governance processes, including triaging proposed AI use cases, conducting periodic reviews of AI systems, and conducting an internal audit of the organization's AI risk management system.

10 Risk Management

Many AI standards and regulations require organizations to implement formal risk management systems for AI and consider the potential risks of AI to individuals and society, not just the operational risks it potentially poses. Key steps in AI risk management include risk identification, measurement, mitigation, monitoring, audit, and incident management.

How Trustible Helps

Risk Management Module - AI risk management is one of the core capabilities of Trustible. Trustible can help organizations automatically identify relevant risks per AI use case, provide guidance and tools for risk measurement, and provide specific mitigation recommendations covering both technical and non-technical mitigation methods. As the landscape of AI evolves, and new risk/mitigation advice emerges, Trustible can alert users about new possible risks or new recommended mitigation practices.

11 Key Performance Indicators and Monitoring

AI governance is constantly evolving, and understanding how to keep an AI governance function up-to-date requires data on its performance. This includes collecting information about how effective it is at capturing potential issues and addressing them, tracking all the people involved in the process, and ensuring there are no bottlenecks or limitations.

How Trustible Helps

Dashboard - Trustible tracks how effective and efficient organizations are with AI governance. This includes looking at how many mitigations are put in place for AI risks, how effective organizations are in conducting AI governance-related workflows, and tracking the number of user-reported incidents. In addition, users can see where they stand against a list of compliance controls, and whether key systems or use cases have fallen out of regulatory compliance.

12 Reporting

AI can be technical and complex, but many non-technical stakeholders need to understand both information about the systems and information about the governance structure around them. It's essential to generate and share regular reports about both specific AI systems and the governance structure as a whole.

How Trustible Helps

Reporting — Trustible can generate both specific reports on AI systems and widespread reports on how effective the organization is at specific AI governance tasks or what the major AI risks posed to the organization are.

TRUSTIBLE.

Pillar 2
Legal and Regulatory Compliance

This pillar is focused on ensuring that AI is developed and deployed in a compliant manner. The scope of this includes not only AI-specific laws like the EU AI Act but also existing laws regarding data privacy, torts, consumer fraud, and others.

1   Assess: Legal and Regulatory Considerations

Each AI use case and system is different, and each legal jurisdiction may have its own set of legal traditions and regulations that may impact use cases differently. Simply understanding what these implications are is a challenging task on its own. This is especially true as complex AI systems or agents may span several systems or tasks, and therefore could have a very complex legal footprint across industry-specific law, privacy regulations, and emerging AI-specific regulations.

How Trustible Helps

AI Inventory
Trustible helps organizations track their AI use cases, models, and vendors, and can track which regulations may apply to each. Trustible offers out-of-the-box use case templates and pre-populated model cards that come with pre-configured regulations requirements or recommendations to help organizations implement sensitive AI use cases compliantly.

Automated Workflows
Trustible offers out-of-the-box workflows to help organizations determine what their regulatory requirements are, and how to execute on any relevant requirement, from conducting an AI risk assessment to full pre-audit reviews for laws such as the EU AI Act. Trustible's default AI Use Case Intake assessment includes recommendations on what regulations may apply, and what an organization's requirements may be under that regulation.

2   Prioritize: Liability and Risk Management

The current legal landscape for AI liability is unclear, and therefore, many organizations need some way to prioritize AI use cases and systems for risk management activities. This is especially true considering many AI systems are now built with complex supply chains of data providers, model developers, and hosting platforms.

How Trustible Helps

Use Case Risk Management
Trustible has built-in AI risk management capabilities, including risk assessment questionnaires, risk measurement guidance, and recommended mitigation measures. Trustible can help teams quickly identify relevant AI risks and create a risk treatment plan spanning technical, product, and legal mitigation options.

Workflows
Trustible offers out-of-the-box workflows for triaging use cases and ensuring that higher risk uses are their respective legal compliance requirements are quickly identified and documented.

3   Plan: Comprehensive Legal Planning

The fast-changing legal landscape around AI means teams need a strategy for managing the legal implications of changing technology, regulations, and use cases. This includes ensuring that any internally created AI systems have appropriate documentation in place and that externally acquired AI has been contractually reviewed for potential legal risks.

How Trustible Helps

Regulatory Tracking & Mapping
Trustible actively monitors AI legislation and legal standards around the world. Key regulations with requirements on AI developers or deployers are represented as frameworks in Trustible and mapped to a common set of control requirements to assist with compliance efforts.

Compliance Documentation
Trustible's AI use case and model inventory schema includes all necessary fields for building compliant documentation and gives guidance on what information needs to be present in each field. Trustible's default reports also conform to expected regulatory formats, allowing organizations to quickly assemble, verify, and register their compliant AI systems.

4   Deploy: Legal Protections and Safeguards

Organizations can't afford to be reactive about AI compliance issues. Retroactively adding compliance can be costly and complicated, and there are serious liability issues in not taking a proactive stance. This is especially true for high-risk areas with many non-AI-specific laws that can still apply to AI activities. Non-AI-specific laws can still be broken by integrating a non-compliant AI system.

How Trustible Helps

Controls
Trustible has a standard set of controls that map to the requirements of global AI regulations and standards. These controls cover issues from ensuring policies align properly, to ensuring that AI use cases and models are appropriately documented, and what kinds of tests or evaluations need to be conducted.

Mitigations Taxonomy
Trustible can help organizations automatically identify relevant AI risks and suggest state-of-the-art mitigation strategies for each risk. Trustible's insights team constantly updates its mitigation database with the latest research, regulatory recommendations, and industry best practices and can recommend the appropriate tools or resources for implementing the mitigations.

5   Monitor: Ongoing Compliance and Audits

Even seemingly small changes to an AI system, such as adding a new training dataset, could quickly change its legal implications. In addition, the system's behavior and performance can shift over time due to model, concept, or user drift. Organizations need to have policies and processes in place to handle any new governance-related event that occurs.

How Trustible Helps

Periodic Reviews
Each use case in Trustible can be assigned a custom review period based on its risk profile. Trustible will then automatically create a new review workflow on a regular basis, pull in relevant data from connected systems, and assign review tasks to the system's owners. This helps ensure regular 'proactive' reviews of systems on a regular basis, which helps discover potential compliance or performance issues.

Automated Controls & Frameworks
Trustible's controls have integrations with leading MLOps platforms such as Databricks, to help constantly monitor and assess control statuses. This helps organizations automatically generate audit trails and identify compliance gaps in real time.

6   Prepare: Review Emerging Trends in AI Ethics and Regulation

There's no shortage of new AI-related lawsuits, novel AI incidents occurring, or capabilities being added to existing AI systems every week. This means that the legal and ethical profile of AI systems is constantly shifting as well. Organizations need to be aware and prepared for how to monitor and respond to these changes.

How Trustible Helps

Risk & Mitigation Updates
Trustible's AI insights team reviews top research publications about AI performance, risks, and mitigation best practices, and identifies the most relevant recommendations for each of them and pushes them to the Trustible platform. This allows system owners to receive insights and recommendations for how to proactively stay up-to-date with the latest risk mitigation strategies.

Regulatory Updates
Trustible's policy team actively monitors all AI-related legislation in the U.S. and major international regulations. The status of these laws, and Trustible's analysis/opinion about their significance, can be viewed and tracked inside Trustible.

TRUSTIBLE.

Pillar 3
Ethics, Transparency, and Interpretability

This pillar is focused on ensuring that ethical factors are considered during AI development and that all relevant parties have appropriate documentation and an understanding of how an AI system can impact them.

1   AI Ethics

AI Ethics include consideration of AI bias. They ensure that humans are held accountable for the development and deployment of AI systems and that systems are designed with positive intent and inclusivity.

How Trustible Helps:

AI Risk Assessment & Use Case Templates
Trustible's out-of-the-box risk assessment can help capture potential ethical issues, and Trustible provides recommended risk and mitigation profiles for common AI use cases, including an understanding of ethical dimensions to them.

Evaluation Taxonomy
Platforms like Databricks offer best-in-class tools for conducting actual bias testing and evaluations of AI systems. However, before that point, teams need to know what tests and evaluations they should run based on the relevant use case and laws. Trustible helps with this first step by recommending specific types of evaluations that should be conducted based on the use cases and applicable regulations.

2   Transparency

Transparency is focused on ensuring that the right stakeholders have the appropriate information about an AI system so that they can ensure that AI is developed and deployed appropriately. This ranges from ensuring that model developers understand the qualities and limits of the datasets they're working with to ensuring that system users can fully understand and interpret the outputs of an AI system they are using.

How Trustible Helps:

Use Case, Model, Vendor Documentation
Trustible helps organizations build compliant documentation about AI use cases and models so that they can be shared with relevant internal and external stakeholders. In addition, Trustible can help organizations identify if a specific use case may require explainable models based on relevant risks or regulatory requirements.

TRUSTIBLE.

Pillar 4
Data, AIOps, and Infrastructure

This pillar is all about ensuring the organization has all the technical infrastructure, capabilities, and tools for building, deploying, and monitoring AI systems in production. This ranges from issues such as how organizations collect and govern the data they use to build, fine-tune, or input into AI systems, to how the organization ensures their deployed systems are robust, reliable, and scalable.

1 Data

There is no shortage of data governance concerns related to AI, both for internally built models and third party ones. Beyond just using data for model training, all the data flowing in and out of systems also needs to be tightly governed. This section is all about ensuring that there are clear data standards, ownership, and quality assessments in place.

How Trustible Helps:

Use Case & Model Documentation
Trustible helps organizations document what datasets were used for model training, fine-tuning, validation, and system inputs and generate compliant reports for customers or regulators. Trustible can also help organizations identify what data quality assessments and tests should be run to provide a degree of AI assurance.

Compliance Workflows
Trustible can help with data impact assessments, data quality reviews, and risk assessments that capture the potential privacy implications or regulatory requirements for each distinct AI use case.

2 AIOps

AIOps is all about the tools and systems used to track models, monitor them in production, and scale them for production. Databricks offers many of the leading tools for this, including MLFlow, and Unity Catalog.

How Trustible Helps:

Model Evaluations & Integrations
Databricks' workspace is well equipped for running model evaluations and tests; however, knowing what the appropriate evaluations are for each use case can be challenging. Trustible is a taxonomy of evaluations, and recommends appropriate ones based on the use case and relevant regulations, and then integrates directly with Databricks to help sync compliance documentation with the technical work being done.

Review Workflows & Feedback Evaluation
Databricks can help organizations collect appropriate logs, metrics, and information about how a model is running in production, but this data needs to be analyzed and reviewed regularly to ensure that any model drift, user feedback, or security issues are quickly identified and responded to. Trustible offers governance workflows to ensure that these periodic reviews are conducted regularly and efficiently, ensuring that AI systems continue to meet their intended objectives.
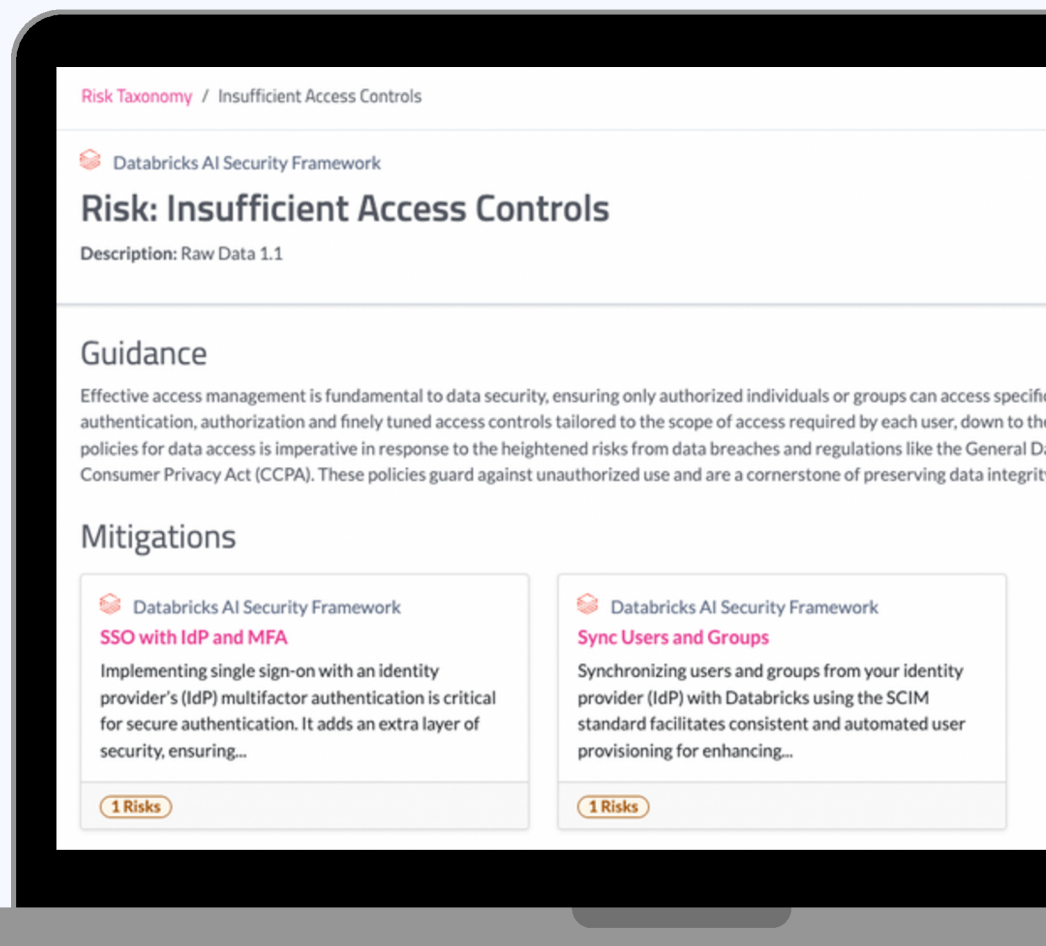
TRUSTIBLE.

AI Security

This pillar builds on the previous work done in Databricks' AI Security Framework (DASF), which outlines core security risks and relevant security controls for each one. The framework covers all parts of the ML lifecycle, from potential data poisoning of raw data to downstream user misuse of an AI system.

How Trustible Helps:

Risks & Mitigations Taxonomy
Trustible has a comprehensive set of recommended risks and mitigations for specific AI use cases. These are mapped to the DASF, and Trustible is integrated with Databricks to help organizations implement, track, and generate proof that mitigations are in place.

Risk Taxonomy / Insufficient Access Controls

Databricks AI Security Framework

Risk: Insufficient Access Controls

Description: Raw Data 1.1

Guidance

Effective access management is fundamental to data security, ensuring only authorized individuals or groups can access specifi... authentication, authorization and finely tuned access controls tailored to the scope of access required by each user, down to th... policies for data access is imperative in response to the heightened risks from data breaches and regulations like the General Da... Consumer Privacy Act (CCPA). These policies guard against unauthorized use and are a cornerstone of preserving data integrit...

Mitigations

Databricks AI Security Framework
SSO with IdP and MFA
Implementing single sign-on with an identity provider's (IdP) multifactor authentication is critical for secure authentication. It adds an extra layer of security, ensuring...

1 Risks

Databricks AI Security Framework
Sync Users and Groups
Synchronizing users and groups from your identity provider (IdP) with Databricks using the SCIM standard facilitates consistent and automated user provisioning for enhancing...

1 Risks

TRUSTIBLE.™   **databricks**

🌐 Website   www.trustible.ai

✉ E-mail   contact@trustible.ai

📍 HQ address   1201 Wilson Blvd, Floor 27
                      Arlington, VA, USA 22209