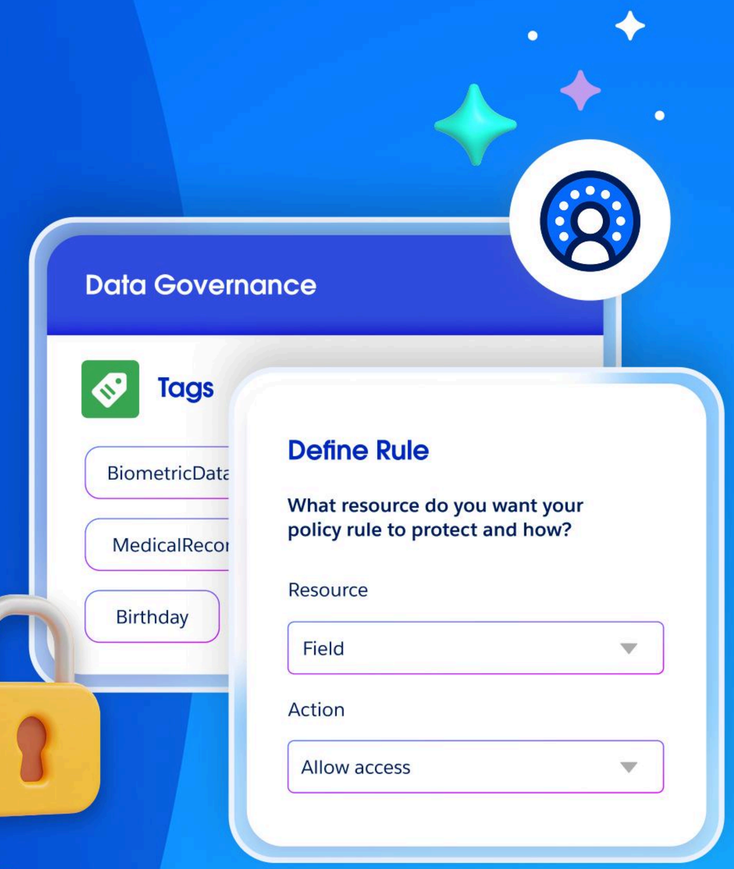
Top 3 Considerations for Governing Data and Metadata at Scale

Guide



Introduction

Governance at Scale: The Key to AI Success

Imagine you're an IT leader at a fast-growing tech company, on the brink of an AI transformation that will streamline how you service customers. Your team is buzzing with excitement about the potential of AI agents to personalize customer experiences and drive innovation. But as you dive into the project, a significant challenge emerges: **your data**. Your service tickets, transaction histories, interaction logs, emails, and chat transcripts are scattered across multiple systems – preventing AI agents from providing meaningful support to customers because they lack a unified source of truth to pull from.

As you start to integrate your customer data and dismantle silos, you realize that governance is an even bigger challenge. Traditional governance policies are complex, requiring detailed rules for each data object and user. Those policies need to be unified, but scaling unified policies is difficult as the volume and variety of your data grows. Plus, managing access and permissions for a diverse user base, from customer service representatives to data scientists, becomes cumbersome. Each type of data – structured, unstructured, and semi-structured – requires different management techniques, further complicating the governance process, increasing the risk of data breaches, compliance issues, and data hallucinations or toxicity.

2/3

Two-thirds of organizations expect AI agents to power more than a quarter of their core processes by 2025.

Introduction

But what if you could consistently unify governance across all your data from any internal source – like financial or sales data, or an external source such as IoT or social media data stored in a data lake – and ensure that every data source is governed at scale to power trusted AI agents?

In this guide, we'll explore three key considerations to help you do just that – so you can build a powerful trusted data foundation for your teams and your agentic AI solutions. You'll learn:

- **The challenges of scaling governance across all the places this data is used – whether it's for agents, analytics, segmentation, or beyond**
- **The top three considerations for governing data and metadata at scale via data platforms**
- **How Data Cloud helps you achieve top-notch governance at scale for AI success**

Effective governance is the cornerstone of success for enterprises in the era of agentic AI. By addressing these considerations, you can ensure that your teams and AI agents can safely activate all your data, driving the success of your business. Let's dive in.

Despite 60% of organizations stating that AI is a key influence on data programs, only 12% report that their data is of sufficient quality and accessibility for effective AI implementation.

What's in this guide?

Chapter 1

The Challenges of Governing Data and Metadata at Scale

Chapter 1

The Challenges of Governing Data and Metadata at Scale

AI is the driving force behind business transformation, but its potential is only as strong as the unified data that fuels it and the governance guardrails that protect that unified data. If governance policies aren't effectively enforced across datasets, user types, and use cases at all times, the risks can be catastrophic:

Unauthorized access and data breaches.
If data isn't encrypted and properly tagged and classified and if policies aren't enforced accordingly, unauthorized personnel might access sensitive information. Additionally, without secure private connections, network traffic can be intercepted, increasing the risk of data breaches.

Noncompliance penalties.
Failing to follow regulations like HIPAA, GDPR, and PII can lead to severe legal penalties and a loss of customer trust. The average cost of noncompliance is a staggering $14.82 million.

But achieving this level of policy enforcement is easier said than done due to:

A mix of structured and unstructured data.
Data from different sources often has varying formats and storage methods, making it hard to apply the same level of enforcement.

Different contexts and trust levels.
Data from sources like CRM, Slack, and public data sources have different contexts and levels of trust, complicating the governance process.

The traditional approach to applying policies in data governance involves creating permission sets for specific objects, such as a database table or a specific field, and assigning those permission sets to specific users. For example, a permission set might allow a user to view certain fields in a customer record but not others.

Chapter 1

Simply put, this approach just doesn't work for AI solutions, especially agentic AI, due to several key limitations:

Lack of scalability: As the volume of data and the number of users grow, manually managing permission sets for each object and user becomes unsustainable. This leads to increased administrative overhead and a higher risk of errors or inconsistencies.

Inflexibility: AI solutions, particularly those involving agents, handle data from multiple sources with varying contexts and trust levels. The traditional approach is rigid and doesn't easily adapt to the dynamic nature of AI data environments.

Complexity: The granular nature of object-level, field-level, and role-level security policies can lead to a convoluted governance structure that hinders the efficiency of AI systems.

Inconsistent enforcement: Manual configuration and updates can result in inconsistent policy enforcement, which can compromise data security and compliance. AI solutions require consistent and reliable governance to operate autonomously and maintain trust.

Dynamic data handling: AI, especially agentic AI, often requires real-time or near-real-time data access and processing. The manual and static nature of the traditional approach cannot keep up with the dynamic and fast-paced requirements of AI.

Chapter 2

Top 3 Considerations to Elevate Data Governance at Scale with Data Cloud

Chapter 2

Top 3 Considerations to Elevate Data Governance at Scale with Data Cloud

Efficient, scalable governance of unified data is the cornerstone of a successful digital workforce. Below are three key considerations to ensure your unified data is consistently protected, accessible, and managed according to its classification and use case. Each consideration – ranging from consistent data tagging and scalable governance to granular access control and security – plays a pivotal role in building a trustworthy environment for all your data. Let's take a closer look.

3 Key Considerations for Scalable Data Governance

- How can I make governance highly scalable for different data types and users?
- How can I ensure granular access control for different data types?
- How can I ensure data security, privacy, ethics, and compliance in a unified data environment?

Chapter 2

How can I make governance highly scalable for different data types and users?

Why it matters

To effectively manage data from diverse sources such as data lakes, warehouses, internal databases, and even different data types like structured and unstructured, scalable governance is key, especially when using zero copy to integrate everything into a single environment. This type of governance allows you to consistently apply access and masking policies using metadata and data tags, no matter who's using the data, be it AI, human agents, customers, or employees.

This consistency across all data sources is vital, particularly with AI, which relies on vast amounts of varied data, making governance and trust more complex.

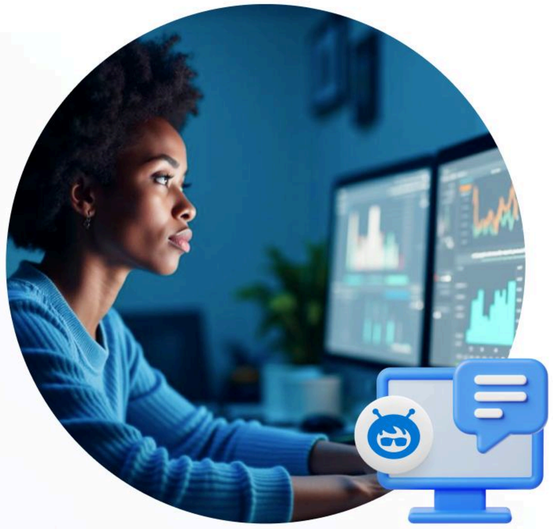
Moreover, scalable governance helps in monitoring and troubleshooting to ensure compliance and meet business needs. It also strengthens security and prevents vulnerabilities by controlling data access, ensuring only authorized users can gain access to certain information.

What's needed for success

- **Policy-based governance powered by a metadata-driven framework:** Shift from managing granular policies to a metadata-driven approach, where policies are defined based on metadata and user attributes. That means any new objects or fields that are assigned with existing tags automatically get the appropriate access governance – no manual updates needed. This ensures consistent policy application across different data types and user roles, making governance scalable and flexible.
- **Automated tagging:** Use AI to automatically suggest and apply tags to metadata and data, ensuring consistency and reducing the risk of errors.
- **Customizable tags:** Define custom tags to meet specific organizational needs, such as tagging data based on its sensitivity or origin, to support specific compliance requirements.
- **User-friendly flexible policy tools:** Define policies via a point-and-click interface that makes authoring, managing, and enforcing policies easier for different personas across your enterprise.

Use Case

Scalable Governance in Action: Employee Facing Interaction with AI Agent



Challenge

Before policy-based governance, authoring and managing different types of governance policies was a complex and error-prone process. For example, creating policies for marketing might require detailed rules for handling customer data, while finance policies needed to encompass a wide range of sensitive financial information and ensure compliance across the organization.

Solution

With AI-based tagging, data that meets defined criteria is tagged “PII”– to ensure data is managed and protected consistently. With policy-based governance enabled by a metadata-driven framework, an AI agent uses flexible policy definitions to tailor data access based on the user’s role. For instance, when a marketing employee queries the agent for customer insights, the agent ensures that any personally identifiable information (PII) is masked. Similarly, when a finance manager requests financial data, the agent applies row-level security to filter out data that is not relevant to their role.

Chapter 2

How can I ensure granular access control for different data types?

Why it matters

Granular access control, including object-level, field-level, and row-level security, ensures that AI agents have the appropriate level of access to data based on predefined policies. This allows for consistent and secure data handling across different interfaces and use cases, such as web interactions, chat, and phone calls, while maintaining data privacy and security.

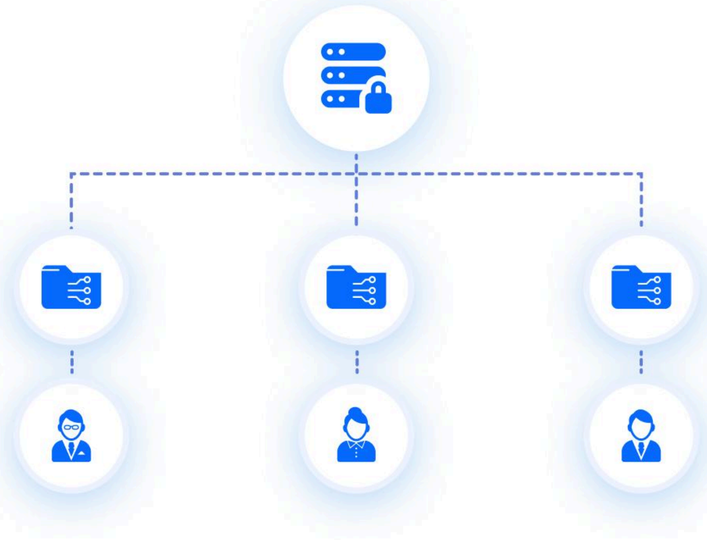
What's needed for success

- **Employ data spaces:** Segregate data, metadata, and processes by brand, business unit, and region. This allows each business unit to maintain control over its own data, regardless of the data type, while still using a unified data platform.
- **Control access at every level:**
  - **Object-level security (OLS):** Set controls to ensure users can only access the objects they are authorized to view.
  - **Field-level security (FLS):** Apply controls to specific fields within objects to ensure sensitive information is only visible to authorized users.
  - **Row-level security (RLS):** Implement controls to ensure users can only see the specific rows of data they are authorized to view.
- **Implement policy-based governance:** Author, manage, and enforce consistent policies that apply across all your structured and unstructured data. Set policies at different levels – object, field, or record – to ensure fine-grained control over data access. Ensure these policies automatically apply across all areas of your data infrastructure.

Taking these actions will ensure your data is handled securely and efficiently, no matter the context or user.

Use Case

Governance in Action: Deploying an AI Agent to Securely Share Information with External Partners



Challenge

Before implementing access governance, managing data access with external partners was a significant challenge. For example, when a partner requested data, the organization had to manually review and control who could access what information. This manual process was not only time-consuming and prone to errors but also frequently led to unauthorized access and data misuse.

Solution

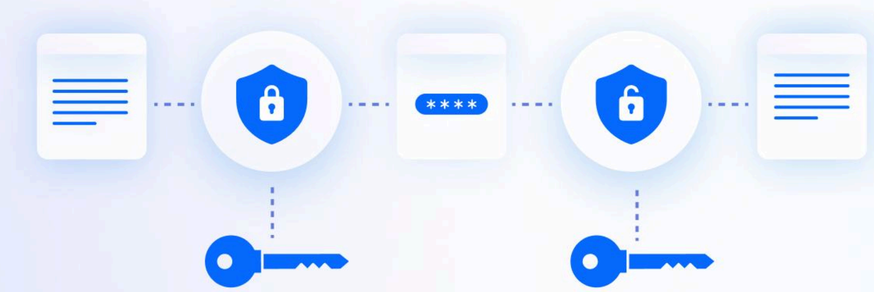
The organization deploys an AI agent that dynamically adjusts both structured and unstructured data access, including OLS, FLS, and RLS, based on the partner's location and the context of the requests. This ensures all data shared is compliant with relevant regulations and maintains the highest standards of data security.

Chapter 2

How can I ensure data security, privacy, ethics, and compliance in a unified data environment?

Why it matters

Companies are increasingly requiring vendors to allow them to manage encryption keys for data at rest, which helps maintain data security even in the event of a data access compromise. Additionally, most data security regulations mandate encryption key management, and failing to meet these criteria can result in severe penalties.



What's needed for success

- **Data encryption:** Encrypt your data both when it's stored and when it's being transferred. By managing your own encryption keys, you have full control over who can access your data. This not only helps you meet strict compliance requirements but also keeps your sensitive information safe.
- **Private connectivity:** Use private network connections to keep your data secure during transmission. By establishing direct, private connections between your data sources and your data environment, you ensure that your data never travels over the public internet – minimizing the exposure of sensitive information to potential threats.
- **Dynamic data masking policies:** Implement dynamic masking to keep sensitive information secure by automatically hiding or showing data based on who's accessing it.

Use Case

Agentic AI Governance in Action: Ensuring Ethical and Responsible Data Use

Challenge

Before implementing governance, security, and privacy measures, ensuring that data was used ethically and responsibly was a manual and inconsistent process. For example, an AI agent might use customer data for marketing purposes without proper consent, leading to ethical concerns and potential legal issues.

Solution

With the new governance and security protocols, both human operators and AI agents can ensure data is used ethically and responsibly.

- **Data encryption** is applied both in storage and during transfer, with full control over encryption keys.

- **Data masking policies** dynamically hide or reveal sensitive data based on user roles, ensuring only necessary information is visible and maintaining appropriate usage.

- **Private connectivity establishes** direct, secure connections between data sources and the data environment, reducing the risk of data interception and breaches.

These measures ensure that customer data is only used for purposes for which consent has been given, whether the action is taken by a human or an AI agent, thereby maintaining trust and legal compliance.

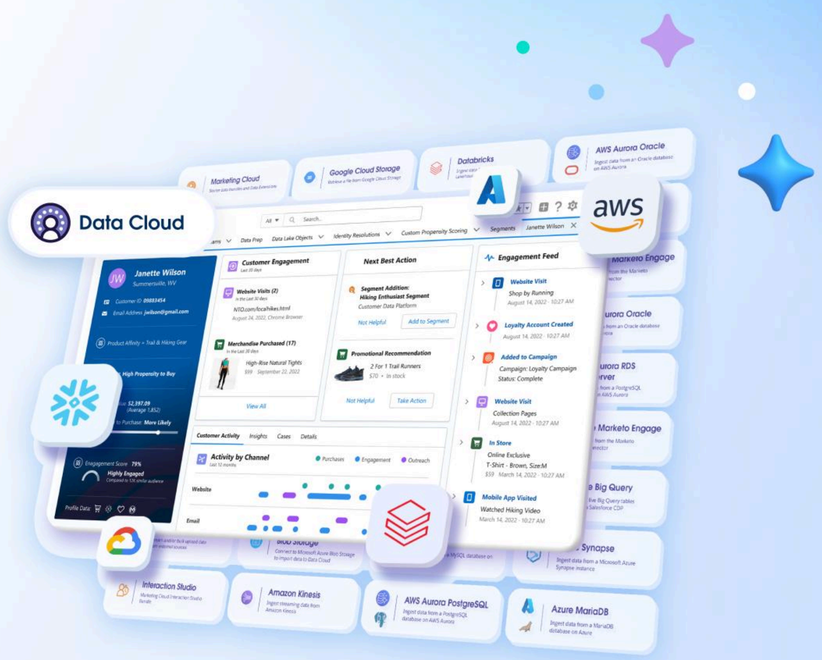
Chapter 3

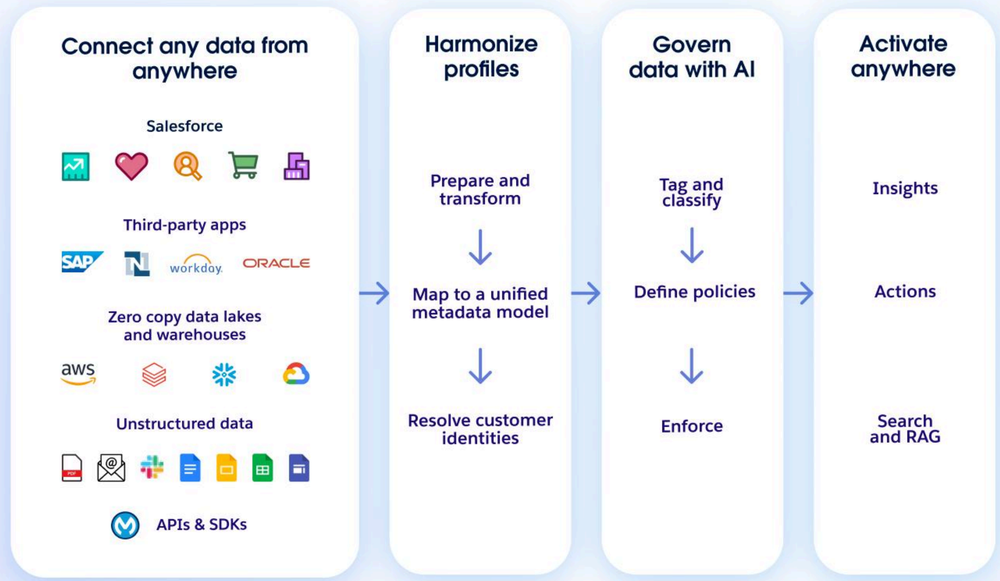
Unlock the Power of Trusted Contextual Data for Your Digital Workforce with Data Cloud

Chapter 3

Unlock the Power of Trusted Contextual Data for Your Digital Workforce with Data Cloud

As the only data platform native to the world's #1 AI CRM, Data Cloud unifies and activates all your data, no matter where it resides. This means your digital workforce can rely on one single trusted source of truth about your business and customers. And trust is at the core of everything we do.





Unify and activate all your data

- **Connect:** Link up your siloed and disparate data no matter where it lives. You can combine structured data, such as customer and order information from different Salesforce orgs or systems (SAP, NetSuite, and Workday), and other data lakes or warehouses (Snowflake, Google Cloud, AWS) with zero copy.
- **Harmonize:** Create a single, unified view of the customer that every app in your Customer 360 and Agentforce can access.
- **Secure:** Use our built-in governance tools and a deeply unified security model to manage data access and keep everything compliant and locked down.
- **Activate:** Deploy a unified profile across any Salesforce app, workflow, or AI application. With sub-second real-time performance, your teams can act faster than ever.

Chapter 3

Govern Data and Metadata at Scale

Data Cloud Governance was built from the ground up to make managing your data seamless and stress-free. It's designed to handle both structured and unstructured data while being deeply integrated into the Salesforce Platform, ensuring consistency in responses across all surface areas where your data is used.

Salesforce is already a trusted, secure platform, and Data Cloud Governance takes that trust even further. With built-in features like AI tagging and classification to keep data organized, plus Policy Based Governance to manage access and scale consistent policies across all surface areas in Data Cloud, you can stay in complete control, with the transparency and trust you need to thrive.

With Salesforce, Data Cloud Governance makes it easy for admins to manage and enforce governance across their organization, all with clicks, not code.

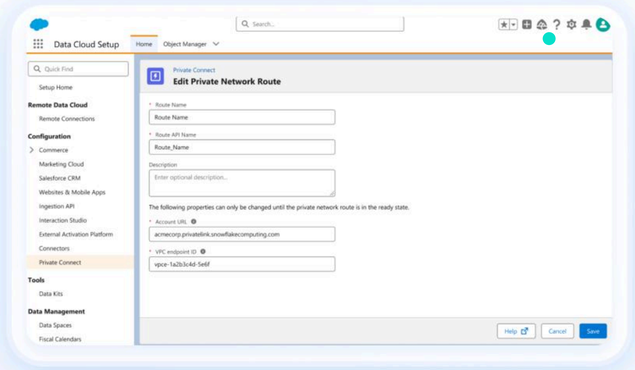
Use our built-in governance tools and a deeply unified security model to manage data access and keep everything compliant and locked down.

- **Easily govern structured and unstructured data.**
  Automate the tagging and classification of your unified data and metadata to drive policy based governance across Agentforce, analytics, segmentation, and more.
- **Enforce consistent data access across all data sources.**
  Empower users to easily author, manage, and consistently enforce policies with clicks across all ingested and zero copy data.
- **Establish a secure and private connection.**
  Prevent vulnerabilities and protect sensitive information while giving flexibility over encryption key management.

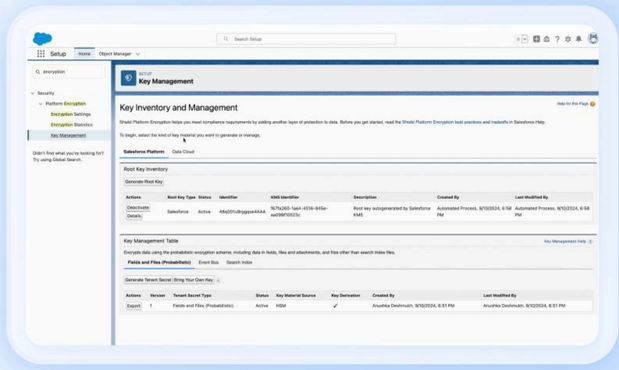
Chapter 3

Key Features and Capabilities:

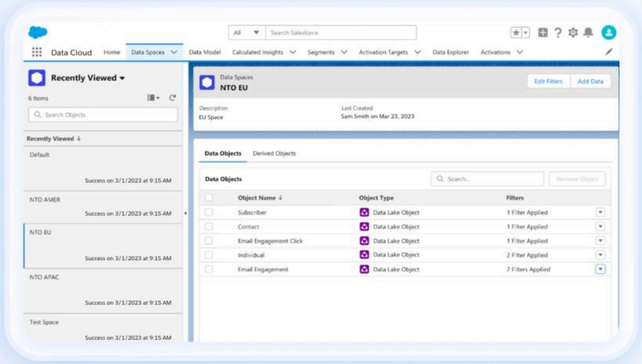
1. **Private connect for Data Cloud:** Securely manage sensitive data by establishing private, direct connections between Data Cloud and data sources such as Snowflake or Amazon Redshift. This means your data is completely isolated and never travels over the public internet.



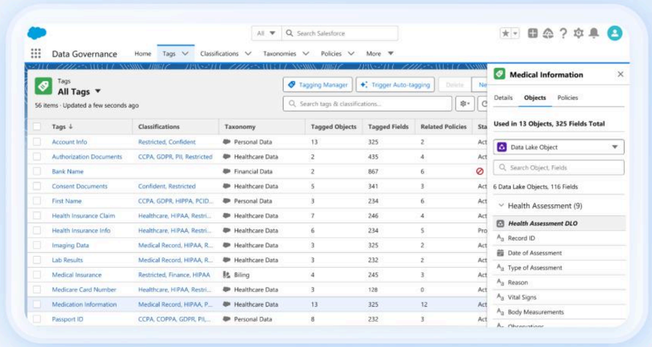
2. **Customer-managed keys:** A key feature of Platform Encryption for Data Cloud, customer-managed keys give admins more control over data security by managing their own encryption keys for data at rest. This ensures that all data at rest in Data Cloud stays protected.



3. **Data spaces:** Segregate data, metadata, and processes by brand, business unit, and region, so each business unit can maintain control over its own data while still only using one instance of Data Cloud.



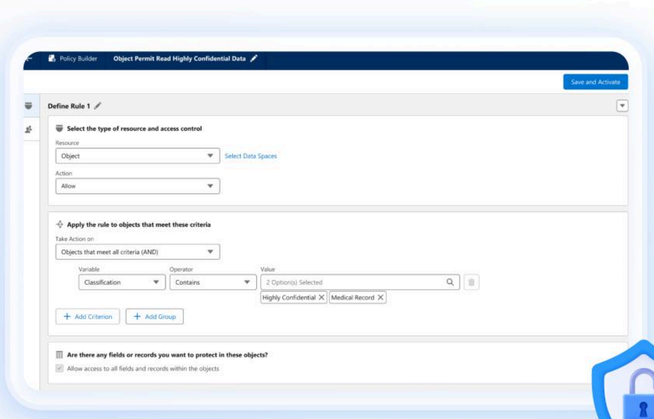
4. **AI-based tagging and classification:** With AI-recommended tags, you can automatically label and classify records – for example, marking data as “HIPAA,” “GDPR,” or “PII”– to ensure data is managed and protected consistently. These tags are designed to follow a business or compliance framework that fits your organization’s needs.



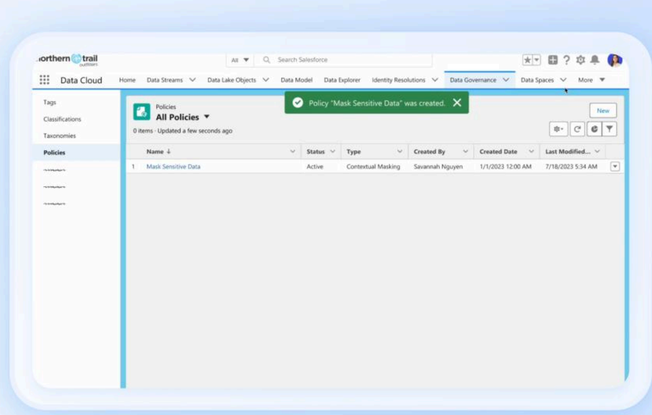
Chapter 3

5. **Policy-based governance:** Author, manage, and enforce consistent policies that surface across all your data. These policies can be set at different levels – object, field, row – giving you fine-grained control over data access. Plus, they automatically apply across all areas of Data Cloud, including Agentforce, analytics, segmentation, and more, so your data stays secure and consistent everywhere.

   - **Access policies:** Ensure everyone only gets access to the data they're allowed to see.



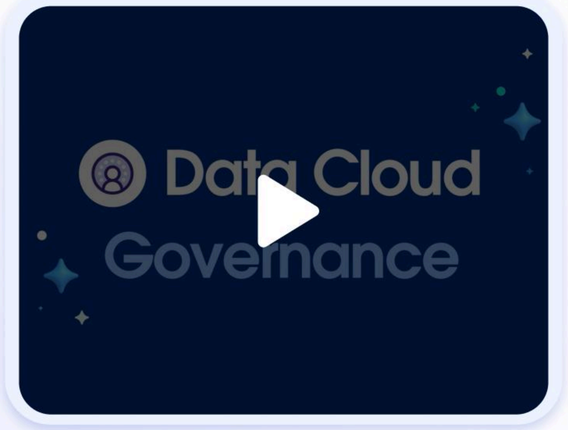
   - **Dynamic data masking policies:** Create masking policies to keep sensitive information secure by automatically hiding or showing data based on who's accessing it.



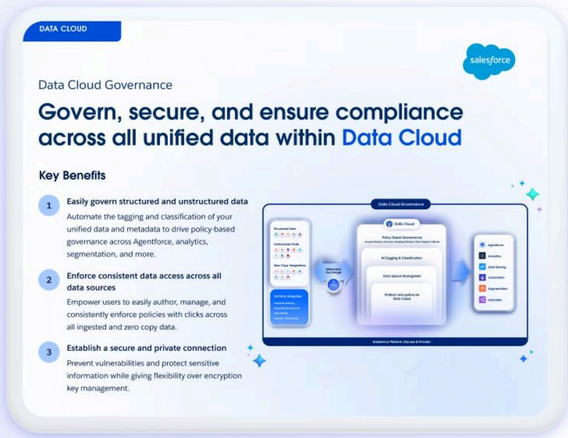
Chapter 3

See Data Cloud Governance in Action

Take control of all your data with Data Cloud Governance, where native security and governance features ensure your data remains compliant, discoverable, and trusted – no matter where it resides in the Salesforce ecosystem. From dynamic masking to private connections and customer-managed keys, see how Data Cloud empowers you to govern data and metadata at scale.



Watch demo



Read the datasheet

The world's most data-driven companies use Data Cloud.

THE ADECCO GROUP
Adecco AKKODIS LHH

See how Agentforce and Data Cloud help The Adecco Group connect 1-to-1 with each job seeker.

Read on

FedEx

See how FedEx grows shipping revenue with unified B2B data in one view.

Learn more

Saks

Discover how Saks elevates luxury shopping with unified data and AI service agents.

Read the story


salesforce