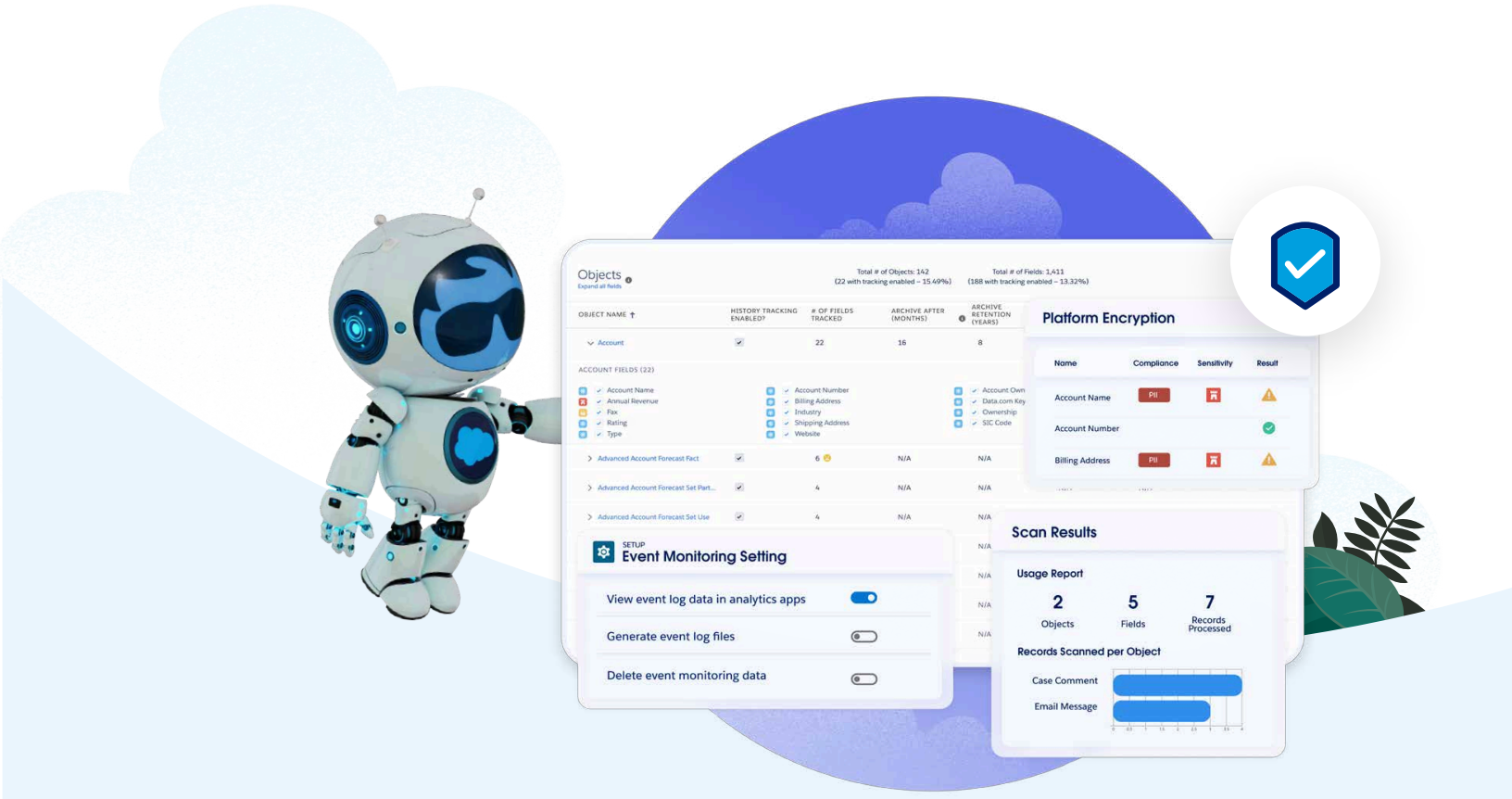
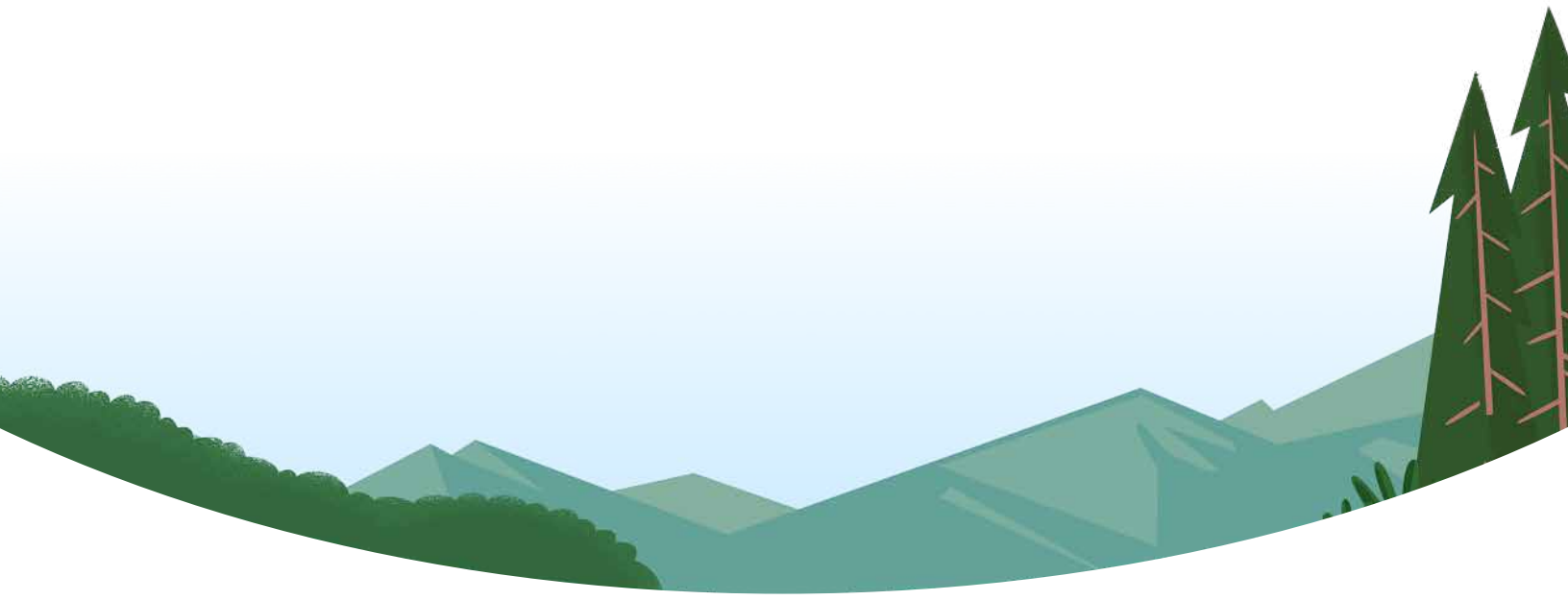
WHITEPAPER

Data Security Best Practices in the Age of AI



Content



Introduction: The critical inflection point of AI and data security

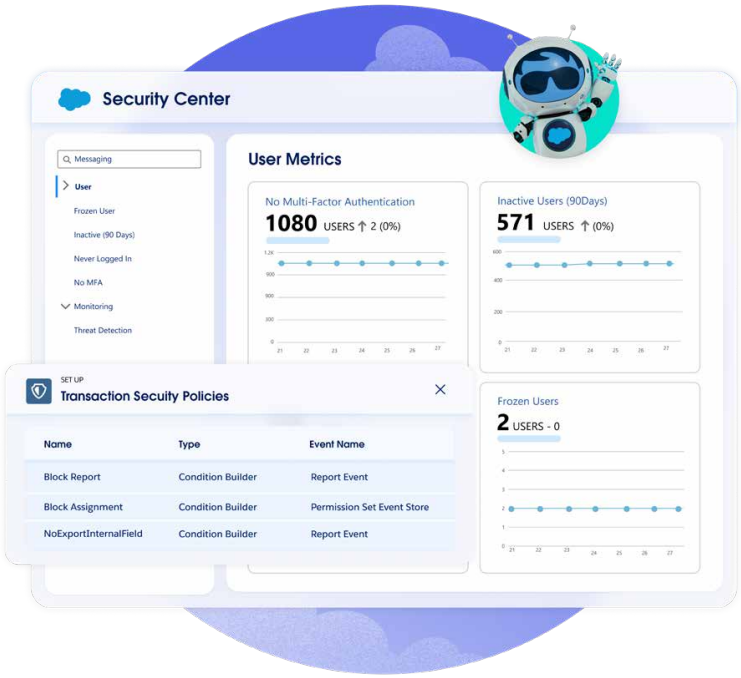
We're seeing that AI is moving fast, and it's changing everything from how we build products to how we serve customers. For IT leaders, that means navigating a new set of security challenges and opportunities. Whether it's large language models powering internal tools or intelligent agents automating tasks across teams, these systems are only as valuable as they are secure. The reality is that AI doesn't work in isolation, it relies on data – often sensitive, regulated, or proprietary – to function.

When that data isn't properly governed, the risks grow quickly. Privacy concerns. Costly data breaches. Compliance gaps. Exposure of critical IP. In a world where speed is the default, security can't be an afterthought.

That's why more organizations are putting data security at the center of their AI strategy. Because once trust is lost, whether internally or with customers, it's hard to win back. And without trust, AI adoption stalls, no matter how advanced the technology.

But here's the upside: security doesn't slow innovation; it enables it. When AI systems are built on strong data foundations, teams move faster with more confidence. Leaders can scale with less risk. And organizations create space for real transformation – not just experimentation.

In this guide, we'll explore the data security risks introduced by AI systems and outline best practices to help your teams build responsibly, reduce exposure, and stay ahead of evolving compliance standards.



Research and statistics: A portrait of the modern trust landscape

AI continues to evolve from predictive tools to copilots, and now to autonomous agents capable of executing tasks with less human oversight. These advancements unlock powerful new capabilities, but they also raise the stakes for data security. The more sophisticated the systems, the more essential it becomes to secure the data that powers them.

Cyberattacks are growing more advanced and harder to contain. On average, it takes nearly 300 days to identify and remediate a breach. At the same time, many organizations lack the personnel or security tools to keep pace. More than half of companies that suffered a breach reported major security staffing shortages – a clear signal that stronger tools and smarter investments are urgently needed alongside the right talent.

As security threats grow, so do compliance demands. Laws like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) continue to evolve. Failure to meet these standards can be costly, with the average non-compliance fine reaching $14.8 million. Prioritizing compliance isn't just about avoiding penalties; it's about building trust with customers, partners, and regulators.

Security threats are the top AI concern for CIOs.

57% of CIOs say that security and privacy threats are their biggest worry when it comes to AI.

Why AI raises new security concerns

AI brings a new set of security considerations. In fact, 71% of IT leaders believe generative AI introduces unique risks. Common concerns include:

- **Data leakage through improperly secured models**
- **Malicious prompt injection or model tampering**
- **Biases or harmful content generated without oversight**
- **Scalable attacks, such as AI-generated phishing or deepfakes**

Without strong safeguards, even well-intentioned AI systems can expose sensitive data, generate unintended outputs, or be weaponized by bad actors. These aren't hypothetical risks – they're already being tested in the wild.

As a result, security is becoming a first-order priority for AI deployment. 75% of IT leaders expect to increase their security investments in the next year, with spending on application and data security projected to rise by 15%. It's not just defensive: companies that heavily leverage AI in their cybersecurity programs see real benefits, saving an average of $2.2 million in breach-related costs.

The takeaway? AI can absolutely be part of the solution – but only if organizations proactively address the security challenges it introduces. What's required now is a forward-thinking, layered approach to protecting data in an AI-powered world.

Security professionals have concerns about AI — but many also see a silver lining.

As regulations rapidly evolve, **68% of security leaders** say staying compliant is becoming more difficult, and **43%** admit they're not fully prepared for potential AI-related rules.

Still, there's optimism: **80%** believe AI agents will create new opportunities to strengthen security and improve compliance.*

say staying compliant is becoming more difficult

admit they're not fully prepared for potential AI-related rules

believe AI agents will create new opportunities

*Salesforce State of IT, 4th Edition: Security Report

Best practices: Strike the balance between AI innovation and data security

By following the eight best practices outlined below, your organization can build a stronger foundation for securing the data that powers your AI systems, including agents. These steps are designed to help you implement a future-ready security framework that protects sensitive information, reduces risk, and builds lasting trust in your AI initiatives.

1 Understand the scope of your data

Start with a data audit and bring the right stakeholders to the table.

Before you can secure your data, you need to know exactly what you're working with. That means conducting a thorough audit of the data flowing through your AI systems, including what's being ingested, processed, or acted on by models and agents across your environment.

Work cross-functionally to map where your data lives, how it's being used, and who or what has access. Does it include sensitive categories like Personally Identifiable Information (PII), Personal Health Information (PHI), or proprietary business data? Are external systems or autonomous agents pulling from it to generate outputs or trigger actions? Understanding these dependencies is key to identifying exposure risks.

It's also important to assess the freshness and lifecycle of your data. Stale or unused datasets can increase your attack surface and complicate compliance. If data is no longer operationally necessary, it should be archived or deleted based on retention policies.

Once you have a clear picture, you can classify your data and apply appropriate controls, from encryption and obfuscation to access restrictions and monitoring. With a full understanding of your data ecosystem, you're better equipped to protect it and to operate your AI systems and agents on secure, well-governed foundations.

Key actions

- Conduct a full audit of your data
- Identify sensitive information
- Classify data based on sensitivity

2 Secure your data

Make sure your AI agents and systems are trained on secure, compliant data.

Your models, workflows, and agents are only as trustworthy as the data they rely on. After classifying your data based on sensitivity, apply strong protections to every stage of the pipeline. That means encrypting sensitive data both at rest and in transit – and just as importantly, managing your encryption keys securely, with regular rotation to prevent misuse.

If agents are accessing or acting on live data, make sure those interactions are subject to the same security protocols as human users or applications. A well-secured pipeline prevents exposure, enforces compliance, and allows your AI systems to operate safely at scale.

Key actions

- Encrypt data at rest and in transit
- Regularly rotate encryption keys to protect access
- Limit access to authorized users and regularly monitor usage

3 Control access with roles and specific permissions

Implement role-based access control (RBAC) for both people and AI systems.

One of the most effective ways to reduce risk is by limiting access – not just to data, but to model inputs, training environments, and outputs. Start by mapping who (or what) has access, what they're accessing, and why. That includes internal users, external vendors, and even autonomous agents operating within your systems.

Use role-based access control (RBAC) to enforce clear boundaries. Define permissions based on need, aligned with internal policies and external regulations. For example, not every developer needs access to raw datasets, and not every system needs write access to production environments.

Regularly audit access logs to identify anomalies and revoke unused privileges. Tight, intentional access control helps ensure that both humans and systems interact with data responsibly – minimizing the risk of leaks, misuse, or escalation.

Key actions

- Define roles and permissions for AI data access
- Conduct audits to detect unauthorized access
- Enforce strict access control policies

4 Stay vigilant

Establish continuous monitoring to protect data and ensure system integrity.

Real-time monitoring is essential to protecting the systems and data that power your AI strategy. By actively tracking performance, usage patterns, and access behaviors across your AI-driven workflows, you can detect irregularities early before they escalate into broader issues.

This includes monitoring the activity of systems and agents to ensure they're functioning as intended, operating within expected parameters, and adhering to internal governance policies. For example, keeping an eye on access logs and usage trends can help surface potential vulnerabilities – like excessive data requests, off-hours access, or integrations behaving abnormally – that may point to a compromised credential or policy violation.

Monitoring also helps catch unintended consequences of user misconfiguration. For instance, if a user mistakenly grants an agent broad access to sensitive data, that agent may then expose that data externally – not due to a flaw in the agent itself, but due to human error. Observability ensures that such lapses can be identified and remediated quickly, preventing privacy or security breaches.

Strong observability doesn't imply distrust in your AI systems, it reinforces confidence in them. With clear baselines, alerting protocols, and transparent audit trails, organizations can proactively maintain security, build trust, and support the safe scaling of intelligent automation.

Key actions

- Monitor AI processes for unusual behavior
- Detect anomalies early and address potential threats
- Track user activity to ensure secure access to data

5 Minimize sensitive data exposure

Mask sensitive data to reduce risk in testing and production environments.

Sensitive data doesn't always need to be visible to be useful. In both testing and production environments, use data masking to obscure personal or proprietary information. This allows teams to build and validate AI systems without exposing the underlying data to unnecessary risk.

Whether you're developing models, training agents, or running simulations, anonymizing PII and other sensitive fields is key to staying compliant and reducing your attack surface. Techniques like data substitution, scrambling, and tokenization help ensure that even if data is accessed, it has no exploitable value.

Masking isn't just a privacy measure – it's a security layer that protects customers, reduces legal exposure, and helps future-proof your AI stack against evolving regulatory expectations.

Key actions

- **Use data masking and anonymization in testing and production environments**
- **Implement different masking techniques based on AI use cases**
- **Balance security and functionality during testing**

6 Track changes to your data

Maintain a complete audit trail to ensure accountability, compliance, and resilience.

In an AI-powered environment, transparency matters. Keeping a detailed audit trail, such as with Field Audit Trail, of who accessed data, when, and what was changed is foundational for ensuring integrity.

Whether data is being accessed by internal users, external tools, or agents, logging every interaction gives you the visibility needed to spot irregularities, confirm responsible usage, and streamline regulatory reporting. Be sure to define how long audit logs should be retained based on your compliance requirements and risk tolerance.

Beyond audit logging, your broader data retention and recovery policies need to scale with your data volume and complexity. Define how long operational data should be stored, how it's secured, and how it can be restored quickly in the event of accidental loss, corruption, or unauthorized changes. A reliable recovery strategy isn't just a safeguard – it's essential so your AI systems remain accurate, auditable, and trustworthy over time.

Key actions

- **Maintain a detailed audit trail of data changes**
- **Set scalable data retention policies**
- **Implement recovery plans to revert accidental changes**

7 Test in safe environments

Use sandboxes to securely build and refine AI systems, including agents.

Safe experimentation is a cornerstone of responsible AI development. Using isolated environments like Salesforce Sandboxes allows your teams to develop, test, and iterate on models, workflows, apps, and agents without exposing sensitive production data.

These environments reduce the risk of accidental leaks, misconfigurations, or unauthorized access by keeping testing activity separate from live operations. This separation is especially important when AI systems are connected to real-time data or business-critical functions.

To further reduce risk, consider using synthetic data – realistic but anonymized datasets that allow for accurate testing without introducing privacy concerns. With the right safeguards in place, sandboxes become a powerful space to validate performance, fine-tune outputs, and build confidence before deployment.

Key actions

- **Test AI models, workflows, apps, and agents in isolated sandbox environments**
- **Use synthetic or masked data for secure testing**
- **Audit sandbox environments to prevent data exposure**

8 Take a proactive approach to AI security

Regularly test and validate your models, apps, and agents to stay ahead of threats.

The best defense is a proactive one. To protect your AI systems – including models, apps, and agents – it's critical to identify and address vulnerabilities before they can be exploited.

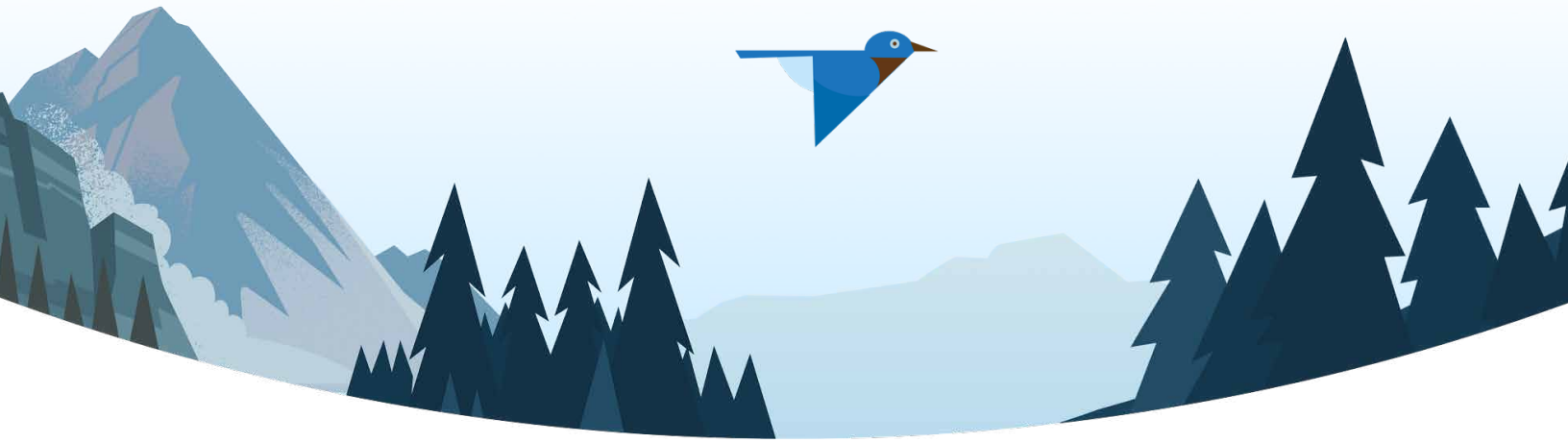
Conduct regular security assessments, including vulnerability scans and controlled penetration tests, to uncover weaknesses in how your AI systems interact with data, users, and other applications. These stress tests help surface hidden flaws that may not be obvious in routine usage but could be targeted in real-world scenarios.

Beyond technical testing, establish clear performance baselines and implement ethical and safety guardrails that validate your systems are functioning as intended. This includes validating outputs, access patterns, and decision logic to ensure your AI remains reliable, compliant, and aligned with your organization's values.

By building security into the development and deployment lifecycle, you create a resilient AI environment – one that adapts to evolving risks and enables innovation without compromise.

Key actions

- **Conduct regular vulnerability scans and simulated attacks**
- **Establish baseline performance metrics**
- **Implement safety and ethical checks to ensure model integrity**



Solutions for success: Salesforce security and privacy

Salesforce Shield: Comprehensive security solutions for critical data

Salesforce Shield delivers the tools you need to uphold strong data security while scaling AI innovation. With capabilities designed to protect sensitive information, monitor system activity, and surface risks early, Shield creates a continuous feedback loop – keeping your organization informed, resilient, and ready to act.

Shield: Event Monitoring

Prevent, mitigate, and monitor threats to sensitive data.

Event Monitoring gives you a clear view of what's happening across your Salesforce environment – who accessed data, when, and from where. With customizable security policies and real-time alerts, your teams can detect anomalies like unusual login activity or unexpected API spikes before they escalate.

Beyond threat detection, Event Monitoring helps optimize performance and drive adoption by uncovering bottlenecks and inefficiencies in how users interact with your applications.

Use cases:

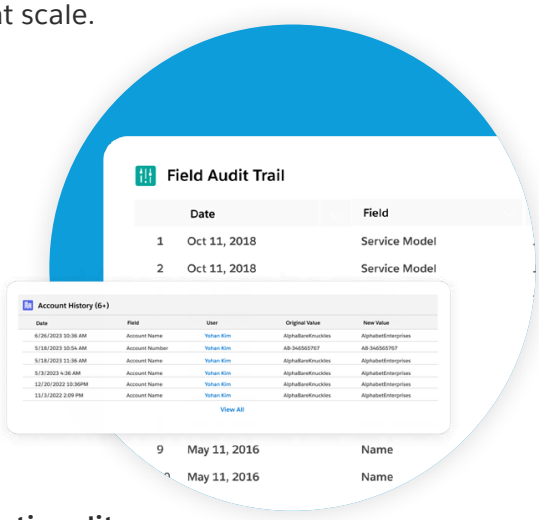
- **User and application visibility:** Monitor and audit user behavior to ensure governance and compliance.
- **Data loss prevention:** Protect sensitive information with custom security policies to mitigate internal and external risks.
- **Productivity and adoption:** Analyze workflows to enhance user productivity and encourage adoption by eliminating inefficiencies.
- **Performance monitoring:** Gain real-time insights into application performance, including page load times and API usage.
- **Low-latency threat response:** Query events directly to investigate and respond to security threats and performance issues quickly.

Shield: Field Audit Trail

Track changes to data at scale.

Field Audit Trail gives you a long-term view of how your data evolves over time. It captures a complete history of changes – including what was changed, who made the change, and when it happened – helping you maintain data integrity and meet compliance requirements.

With flexible retention policies, you can store field history data for as long as your business or regulatory needs require. This creates a reliable audit trail that not only supports governance, but also powers trend analysis and root-cause investigations at scale.



Use cases:

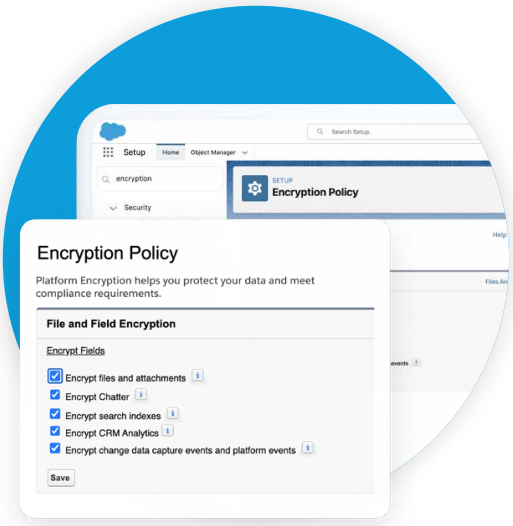
- **Audits and inspections:** Easily demonstrate compliance by maintaining records of field-level changes.
- **Data recovery:** Restore data to its original state after accidental changes.
- **Navigate compliance:** Meet industry-specific retention policies with the ability to store records indefinitely.

Shield: Platform Encryption

Encrypt sensitive data at rest while preserving business functionality.

Platform Encryption helps protect your most critical data across standard and custom fields, files, and attachments while maintaining the seamless user experience your teams rely on. Unlike traditional encryption methods that can limit usability, Platform Encryption preserves core capabilities like search, lookups, workflows, and collaboration tools.

With flexible key management, you maintain full control over how encryption is handled. Choose from bring-your-own-key (BYOK) options or use Shield Keys to encrypt data across Salesforce, including Data Cloud – ensuring privacy and compliance across every layer of your architecture.



Use cases:

- **Key management lifecycle:** Maintain control over the entire key lifecycle for enhanced visibility and data protection.
- **Contractual commitments:** Fulfill customer encryption expectations, speeding up contract negotiations.
- **Compliance with regulations:** Support compliance with data protection regulations such as HIPAA, the Financial Industry Regulatory Authority (FINRA), and the Payment Card Industry Data Security Standard (PCI DSS).

Shield: Data Detect

Find and classify sensitive data quickly.

Data Detect helps you uncover and manage sensitive information – such as credit card numbers, email addresses, and other personally identifiable data – by scanning and classifying records based on customizable detection patterns.

With flexible classification rules, you can flag high-risk data, monitor it over time, and apply security controls more effectively. Paired with other Shield features, Data Detect supports compliance with evolving data protection regulations and strengthens your overall data governance strategy.

Use cases:

- **Deeper understanding:** Uncover hidden sensitive data to improve your data governance and security posture.
- **Improved compliance:** Classify data appropriately to meet compliance requirements.
- **Increased control:** Apply more granular controls around classified data to maintain regulatory compliance and enhance data security.

Customer story: RBC Wealth Management

As part of a company-wide digital transformation, RBC Wealth Management partnered with Salesforce to consolidate multiple CRM systems and unify data from 26 legacy platforms into a single, integrated wealth management solution.

With **Salesforce Shield** at the foundation, RBC was able to ensure top-tier data security and meet strict regulatory standards from agencies like the Securities and Exchange Commission (SEC) and FINRA without slowing innovation. The result: faster operations, reduced overhead, and a dramatically improved experience for clients and advisors alike.

By embedding security into its transformation journey, RBC was able to cut IT maintenance costs by 50% and reduced client onboarding time from weeks to just 24 minutes – all while delivering the high level of trust and personalization their clients expect.

See how RBC accelerated onboarding and cut costs with Shield

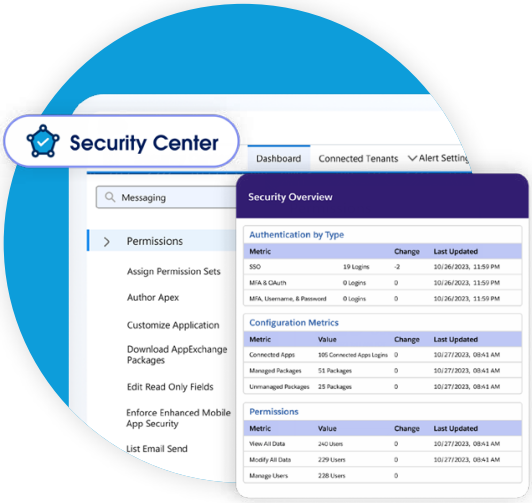
Security Center

Get comprehensive visibility into your security posture.

Security Center allows you to centrally monitor and manage your organization's security health from a single dashboard, consolidating all permissions and controls. With this unified view, you can improve your compliance and security posture and take the necessary steps to protect your data.

Use cases:

- **User permissions:** Consolidate and review user permissions to identify users with critical system access and monitor changes to key system settings.
- **User activity and authentication:** Identify users without access, such as contractors or third parties, and track how users are logging into your systems.
- **Configurations:** Detect misconfigurations that could pose vulnerabilities and ensure settings align with your security intentions.



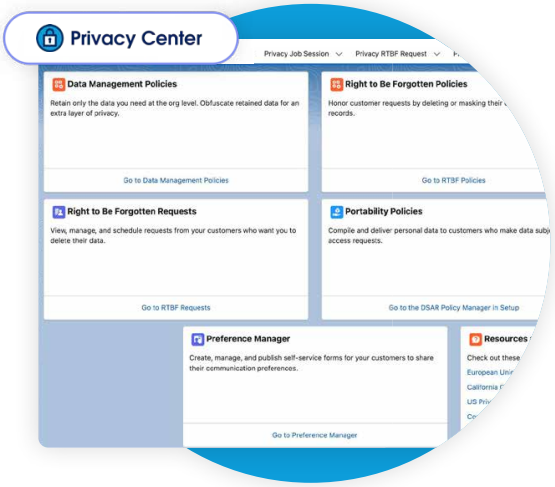
Privacy Center

Take control of data privacy.

Privacy Center empowers you to manage and safeguard your customers' personal data while navigating compliance with ever-evolving data privacy regulations. With streamlined data management capabilities, you can efficiently address customer requests and build lasting trust by protecting their sensitive information.

Use cases:

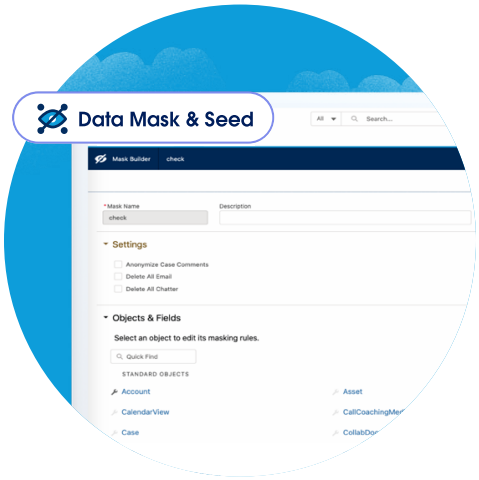
- **Data management policies:** Automate processes such as data deletion, masking, or retention to meet privacy law requirements, minimize storage risks, and reduce the chance of non-compliance.
- **Customer requests:** Seamlessly manage data subject access requests (DSARs) and right-to-be-forgotten requests, ensuring timely, accurate responses to customers while maintaining full control of personal data.
- **Consent management:** Easily capture and manage customer consent preferences with customizable forms that integrate directly into your Salesforce environment, ensuring all interactions align with the latest consent guidelines.



Data Masking

Protect your sensitive data during testing.

Data Mask & Seed allows you to provide realistic data for admins and developers while protecting sensitive customer information from leaks and unauthorized access. Rather than manually securing data for Salesforce Sandbox environments, Data Mask & Seed automatically masks sensitive customer information in whole or partial copy Sandboxes, streamlining your testing processes while ensuring data security.



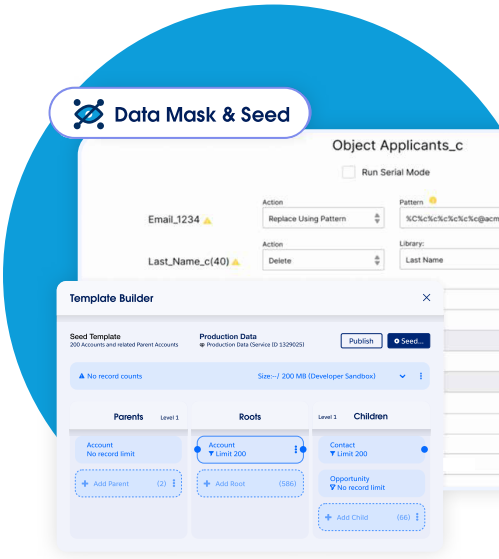
Use cases:

- **Rapidly secure your Sandbox:** De-identify your sensitive data using the fastest available masking tool for Salesforce data.
- **Develop with realistic data:** Leverage built-in field definitions to customize testing data to meet your developer's needs quickly.
- **Accelerate privacy compliance:** Manage compliance with global regulations like HIPAA, GDPR, FINRA, PCI DSS, and the Sarbanes-Oxley Act (SOX).

Data Seeding

Enhance testing and development with relevant, structured data.

Data Mask & Seed also enables admins and developers to populate Salesforce Sandboxes with high-quality, representative data by preserving data relationships. By selectively seeding data from production or other environments, teams can quickly seed relevant data for more effective testing and development. Automated seeding ensures that test environments closely mirror real-world conditions, enhancing accuracy and efficiency while seamlessly integrating with Sandboxes and Data Mask & Seed.



Use cases:

- **Ensure realistic testing:** Seed Sandboxes with structured datasets that maintain relationships, enabling more accurate and effective development.
- **Optimize data control:** Customize your Sandbox by adding new records or updating existing records, ensuring test environments contain only necessary information.
- **Streamline deployment:** Employ reusable templates and automation to efficiently provision data across multiple Sandboxes, supporting agile development and compliance needs.

Customer story: WR Immigration

With a global presence in over 100 countries, WR Immigration manages complex legal cases and sensitive personal information on behalf of its corporate clients. To support its digital transformation, the firm partnered with Salesforce to modernize its case management system and build a secure, client-facing portal: WRapid™.

By leveraging **Salesforce Shield** for encryption and **Data Mask & Seed** for secure sandbox testing, WR Immigration ensures client data remains protected – even as developers iterate and optimize workflows behind the scenes. This approach enables the team to innovate confidently without compromising compliance or confidentiality. With security embedded throughout its development process, WR Immigration continues to deliver trusted, efficient immigration services on a global scale.

Discover how WR Immigration safeguards sensitive data with Salesforce

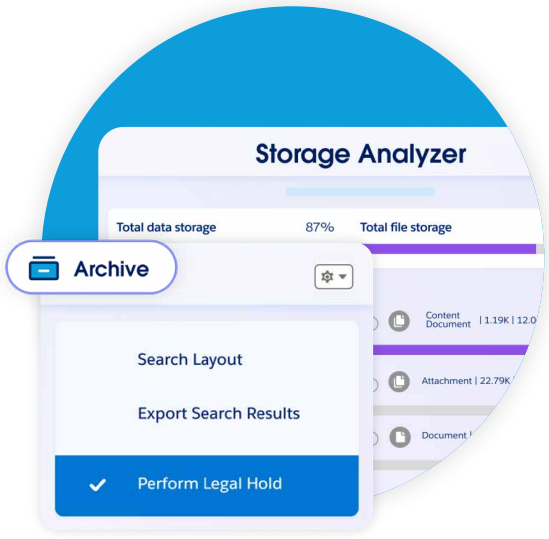
Archive

Improve governance, boost system performance, and navigate compliance by offloading inactive data.

As data volumes grow, retaining every record in your production environment can slow system performance and increase your risk of non-compliance. Archive helps you offload inactive data, while keeping it available in a centralized place for audits, reference, or compliance purposes. With automated archiving policies and the ability to seamlessly view archived records from production, you can manage your data lifecycle more efficiently – without disrupting users or workflows.

Use cases:

- **Improve governance:** Control access to your historical Salesforce data for enhanced organization and oversight.
- **Boost system performance:** Keep production orgs lean and performant by archiving historical or infrequently accessed data.
- **Navigate compliance:** Adhere to industry and government regulations for data retention by securely archiving and preserving historical information.



Backup & Recover

Protect critical data from loss and corruption to ensure business continuity.

Data loss, corruption, or accidental changes can disrupt operations and compromise business-critical information. Backup & Recover helps safeguard your data, ensuring that it remains intact and can be quickly restored when needed. By automating backups and enabling precise recovery, organizations can minimize downtime, prevent data-related disruptions, and maintain compliance with industry best practices.

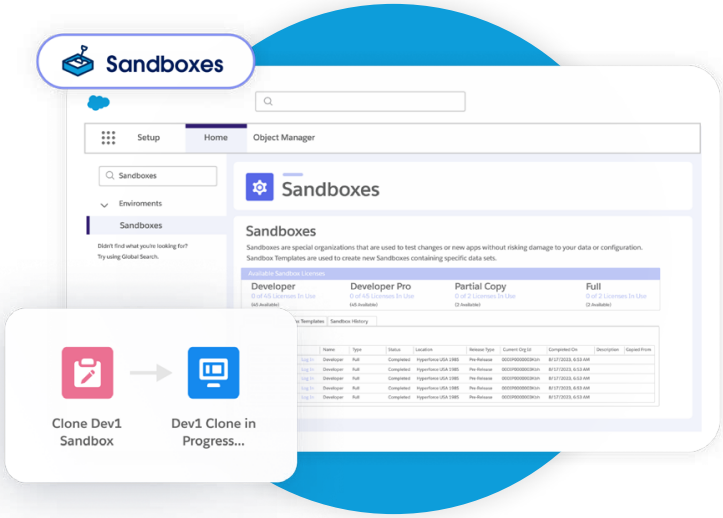
Use cases:

- **Prevent data loss:** Automate regular backups of data, metadata, and files to ensure critical information is always recoverable.
- **Restore data with precision:** Quickly and accurately recover lost or corrupted data to maintain operational continuity.
- **Detect and respond to incidents:** Set proactive alerts to monitor data activity and identify potential security risks before they escalate.

Sandboxes

Replicate the production environment to avoid impacting live data.

Salesforce Sandboxes offer isolated, trusted spaces separate from your production environment for development, testing, quality assurance, and user training. As close replicas of your live system, Sandboxes enable teams to safely test applications without impacting customers, giving them the flexibility to innovate and work independently.



Conclusion: Deploy AI quickly and securely with Salesforce

AI has the power to transform how organizations operate, but with that potential comes responsibility. Protecting sensitive data isn't just a requirement – it's a foundation for earning trust and accelerating adoption.

To strike the right balance between innovation and security, you need a strategy that spans the entire data lifecycle: understanding where your data lives, protecting it with purpose, and continuously monitoring it as your systems evolve.

The **Salesforce Platform** helps organizations move fast without cutting corners on security. With integrated solutions that support governance, compliance, and secure development, you can confidently bring your AI initiatives to life while keeping data safe and private.

Explore the full range of Salesforce security solutions available to help you protect your data while driving innovation.

1. **Explore** how AI innovation and security go hand-in-hand with insights from security experts.
   Get the State of IT: Security report
2. **Discover** how to strengthen your security posture while accelerating innovation.
   Read the guide
3. **Learn** how Salesforce Platform helps organizations stay audit-ready and resilient.
   Explore data security solutions