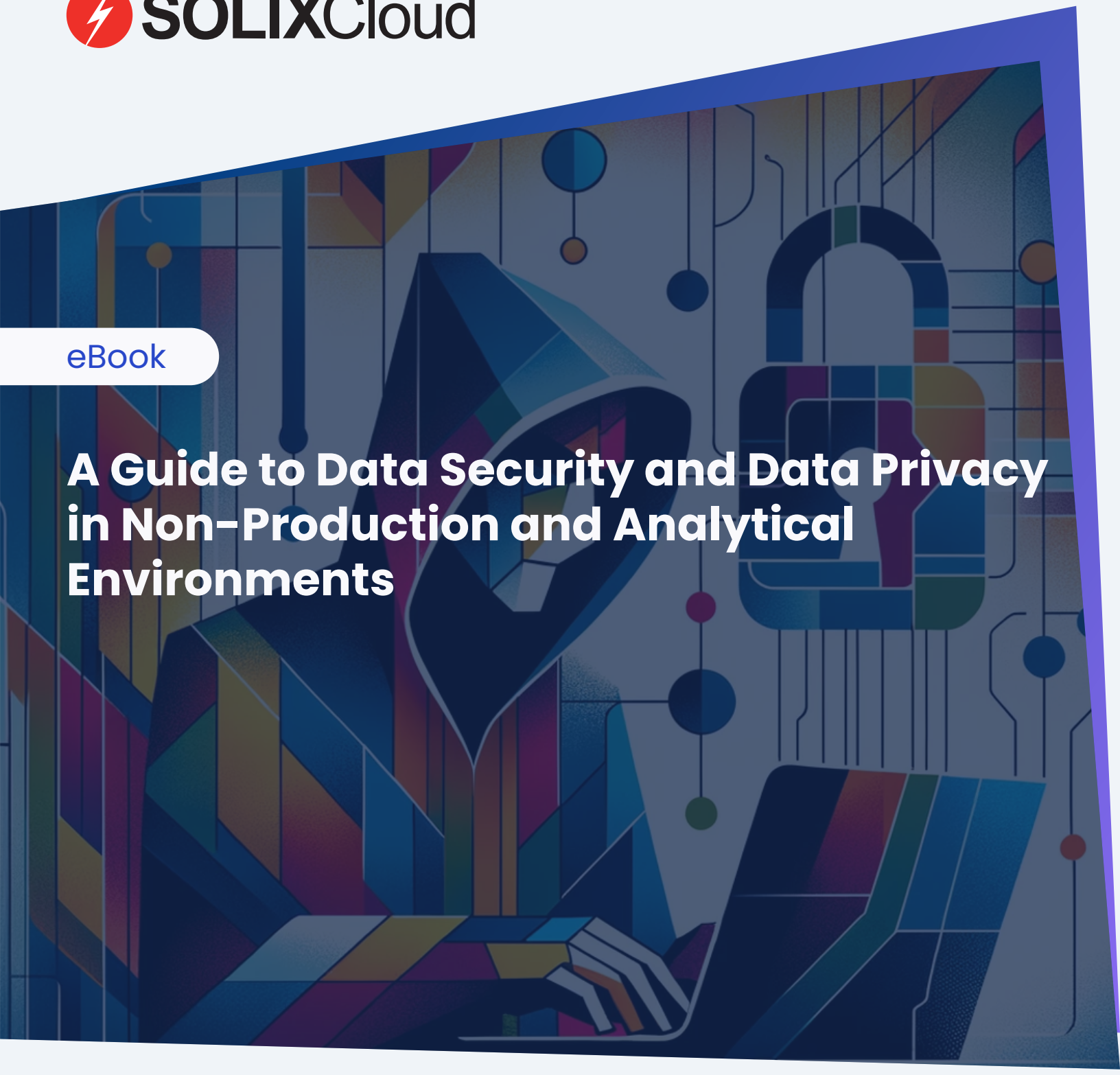
SOLIXCloud



eBook

A Guide to Data Security and Data Privacy in Non-Production and Analytical Environments

Call to find out more
+1.888-GO-SOLIX

Visit our website for more
www.solix.com

Overview

In the era of data-driven enterprises, where data has emerged as a cornerstone of business strategy and decision-making, the significance of safeguarding sensitive data has reached unprecedented heights. A growing number of organizations are integrating AI into their business strategies, and we are witnessing the emergence of new AI use cases, including Large Language Model (LLM) applications. Concurrently, these organizations are embarking on cloud migration, embracing multi-cloud and hybrid IT architectures. As a result, data is increasingly dispersed across multiple non-production, analytical, and production environments, leading to organizational data sprawl. This sprawl makes data more vulnerable.

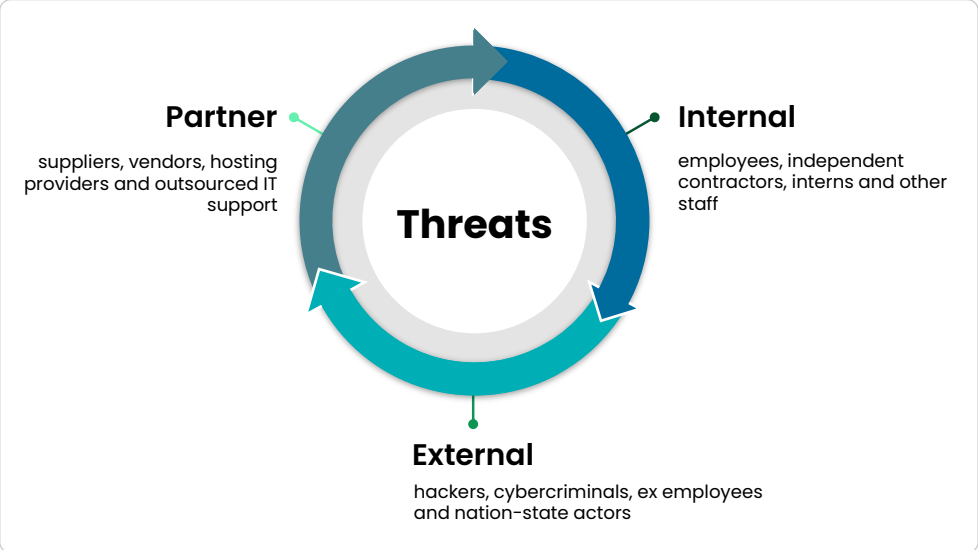
Traditionally, enterprises have focused on securing data in production environments, often overlooking the critical need for data security in non-production and analytical areas. This oversight exposes organizations to malicious external attacks and inadvertent internal mishandling of sensitive data. Furthermore, the landscape is complicated by the stringent web of industry and privacy regulations that demand meticulous data protection, making data security a non-negotiable imperative for enterprises worldwide. These risks not only have financial implications but also impact the organization's reputation and customer trust. Responding to these challenges reactively instead of proactively can adversely affect business continuity.

This e-book discusses the key drivers for securing sensitive data and the vital role of data masking in safeguarding sensitive data in often neglected non-production and analytical environments. We will also discuss what you need to look for while evaluating a data masking solution for your organization. Finally, we will review the Solix Data Masking solution in detail for your consideration.

The Challenges and Risks

Internal and External Threats

Today's enterprises are navigating a challenging and expanding terrain of data security threats, particularly in non-production and analytical environments. These internal and external threats present distinct challenges and risks, underscoring the need to protect sensitive data in these environments.



External threats often involve sophisticated cyberattacks aimed at exploiting system vulnerabilities. These attacks can be particularly damaging in analytical environments, where large volumes of data are processed and stored. Attackers may use advanced methods like social engineering, malware, or network intrusion to access sensitive data, including intellectual property and personal customer information.

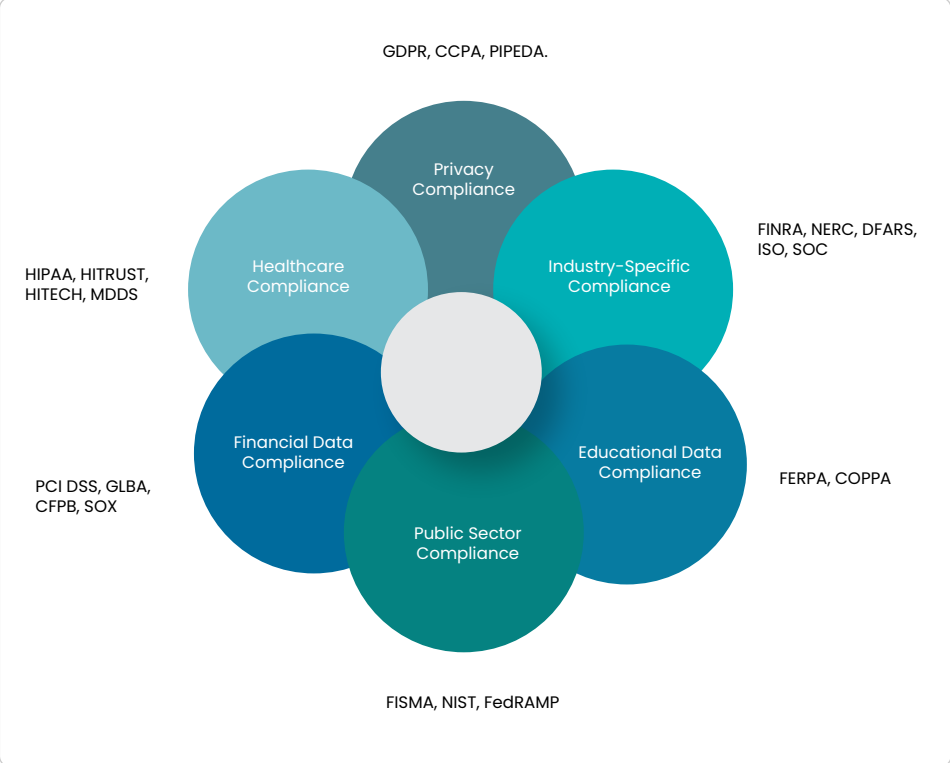
On the other hand, internal threats stem from within the organization and can be more challenging to detect and manage. In non-production environments, where data is used for development, testing, or analysis, the risk of accidental exposure or misuse is heightened. Employees with access to sensitive data in these environments might unintentionally leak information due to inadequate knowledge of security protocols or negligence. Moreover, the risk of intentional insider threats, such as data theft or sabotage by disgruntled employees, poses a constant concern. These internal threats are exacerbated by non-production environments, which often do not receive the same security scrutiny as production environments.

- The global landscape of data breaches paints a grim picture of the risks associated with poorly managed enterprise data environments. According to the 2023 Verizon Data Breach Investigations Report, there were over 16,312 security incidents and 5,199 confirmed data breaches in that year alone. Notably, 83% of these breaches were perpetrated by external actors, while 17% involved internal actors, underscoring the significant role of both external and internal threats.
- The financial sector, healthcare industry, and public sector are among the most affected by these breaches. For instance, the healthcare industry reported 436 confirmed data breaches compromising personal and medical records, with a significantly high (35%) driven by Internal Actors.
- IBM's Cost of a Data Breach Report 2023 highlights the financial implications of these breaches. The global average cost of a data breach in 2023 was USD 4.45 million, a 15% increase over 3 years.

These figures underscore the severity of data breaches and the importance of securing data in all environments, including non-production and analytical spaces.

Industry and Privacy Regulations

With the proliferation of data breaches and privacy concerns, governments and regulatory bodies worldwide have introduced a range of regulations to ensure that organizations responsibly handle and protect sensitive data.



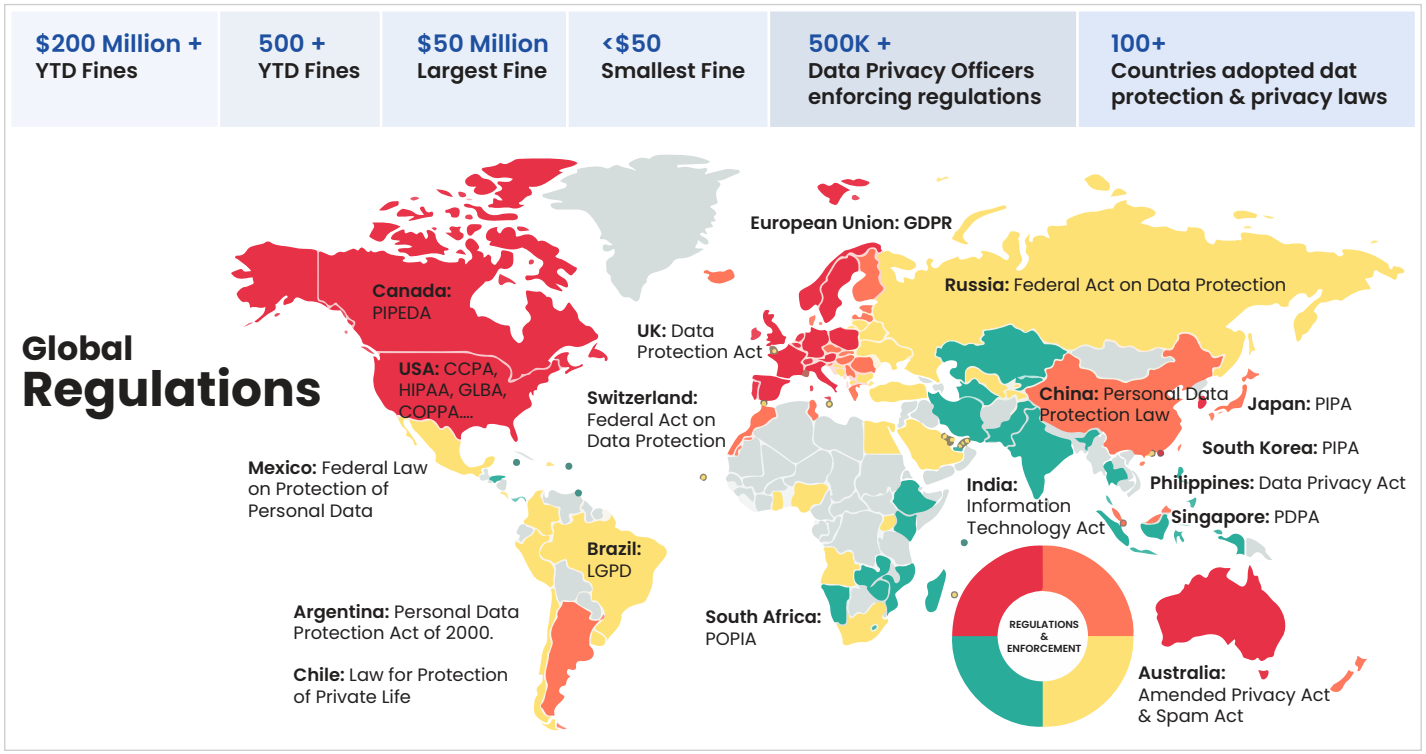
Key industry regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in healthcare, the Payment Card Industry Data Security Standard (PCI DSS) in finance, or the geo-specific privacy regulations such as the General Data Protection Regulation (GDPR) in the European Union, the California Consumer Privacy Act (CCPA) in the United States, and similar laws in other regions set strict standards for data privacy and security. These regulations mandate enterprises to implement robust data protection measures and ensure transparency in data processing.

The penalties for non-compliance with such laws are substantial, emphasizing the gravity of the responsibility organizations bear in safeguarding the privacy and integrity of the data under their purview.

Sources:
Verizon report: 2023 Data Breach Investigations Report (DBIR)
IBM - Cost of data breach report 2023

Non-compliance with such regulations carries hefty fines, reputational damage, and, in severe cases, legal actions, making adherence to data protection standards paramount for enterprises operating in today's interconnected world.

- As of October 2023, over 130 data privacy laws and regulations are in effect worldwide, with more expected to be enacted in the coming years. This complex and ever-changing landscape makes it challenging for organizations to maintain compliance.
- A 2023 study by PwC found that 73% of organizations have been fined for non-compliance with data privacy regulations. This number is up from 63% in 2022, indicating non-compliance is still a significant problem.
- Didi Global was fined $1.2 billion by China's Cyberspace Administration of China (CAC) on July 21, 2022, for violating China's data protection laws. The largest HIPAA fine in 2023 was $16.5 million, imposed on Banner Health. In 2022, Instagram was fined $403 million by the Irish Data Protection Commission (DPC) for violating the GDPR.

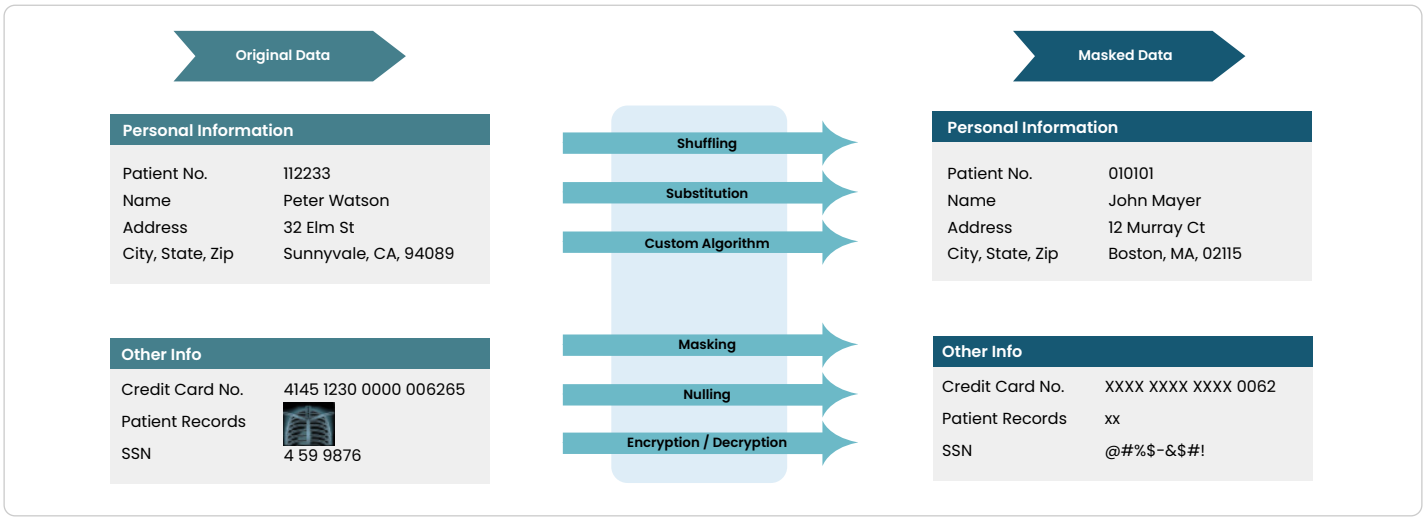


Data Masking to the Rescue

Data masking emerges as a pivotal strategy for navigating the perils of data breaches and ensuring compliance with stringent data privacy regulations in non-production and analytical environments. Data masking involves obfuscating or substituting sensitive data into a fictional but realistic version, ensuring that the data remains usable for testing or analytical purposes without exposing sensitive information.

For example, in development and testing environments, data masking is crucial for using realistic data safely. It allows developers and testers to work with true-to-life datasets without compromising sensitive details like names or identification numbers. By obscuring these elements while maintaining the data's overall structure, data masking ensures security and functional integrity.

For analytical purposes, data masking is key to harnessing data insights without breaching privacy. Advanced techniques like format-preserving encryption enable organizations to analyze data without risking exposure to sensitive data. This approach maintains the analytical utility of the data, allowing for valuable insights without legal or ethical risks associated with sensitive information exposure.

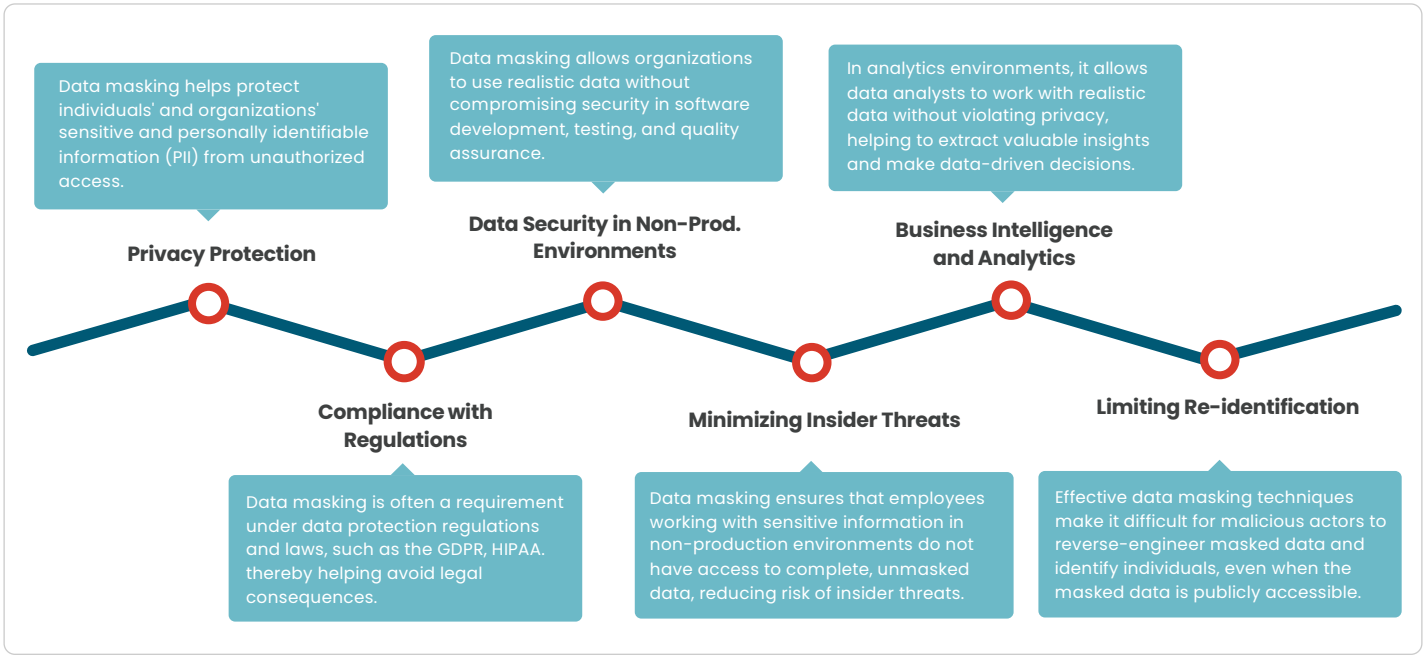


Data masking is a proven solution to address the dual challenge of safeguarding data from both internal and external threats; it reduces the risk of data leaks from within the organization and diminishes the value of the data for external attackers, thereby lessening the attractiveness of the target.

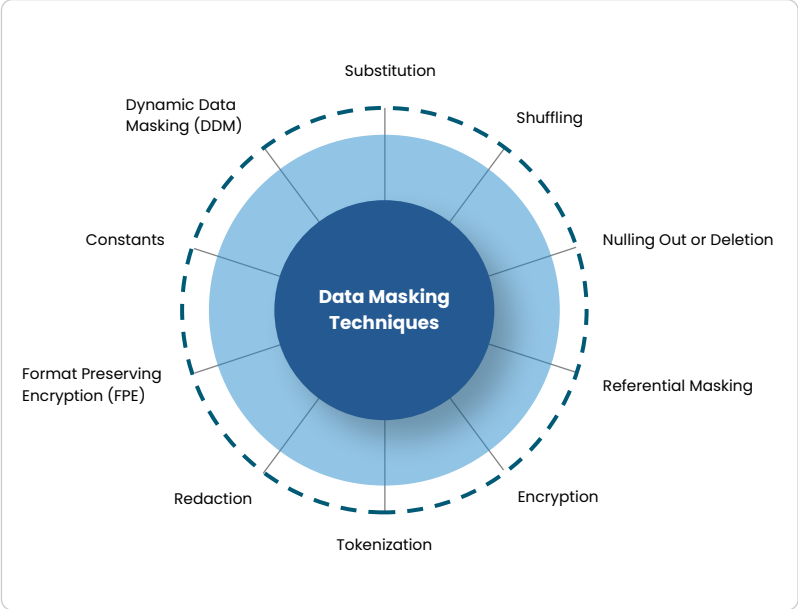
Moreover, data masking is a proactive measure in aligning with data privacy regulations like HIPAA, GDPR, and CCPA. By anonymizing personal data, enterprises can significantly reduce the risk of non-compliance and the consequent heavy fines and reputational damage.

Data Masking is increasingly recognized as an essential component in the modern enterprise's toolkit for managing the complex landscape of data security and privacy.

Benefits of Data Masking



Key Data Masking Techniques



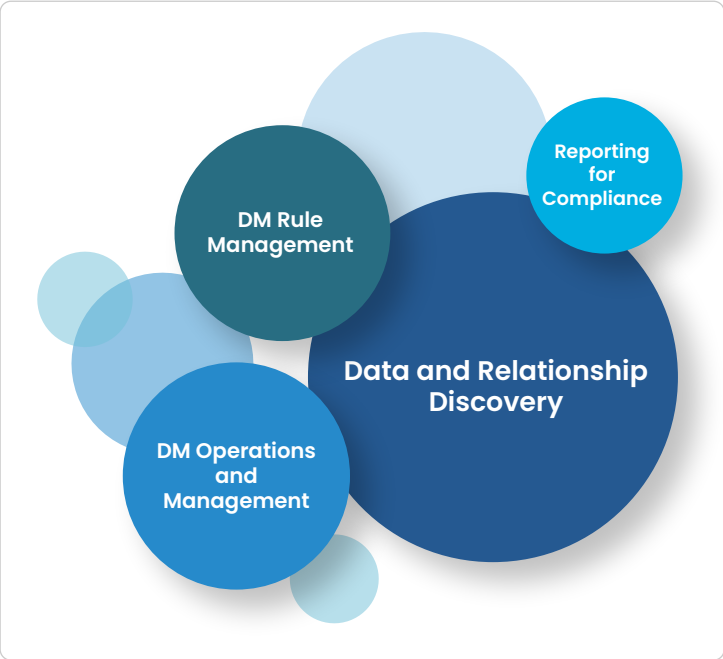
Masking offers multiple techniques to protect sensitive data. Each technique offers a unique approach to safeguard data while maintaining its usability. Let's explore some of these essential data masking techniques:

- **Substitution:** This technique involves replacing sensitive data with fictional but plausible data. For example, a real name might be replaced with a fictitious name. The substitute data should maintain the same format and type to ensure functional testing and analysis.
- **Shuffling:** Shuffling rearranges the values within a column. For instance, names in a customer database might be shuffled to no longer correspond to the original rows. This maintains data integrity while protecting individual identities.
- **Nulling Out or Deletion:** This straightforward technique involves removing or replacing sensitive data with null values or a standard placeholder. It's useful when the specific data isn't required for a particular test or analysis scenario.
- **Encryption:** Sensitive data is encrypted and can only be decrypted with the correct key. While secure, it may limit the usefulness of the data in some non-production scenarios unless users have access to decryption keys.
- **Tokenization:** This involves replacing sensitive data with unique identification symbols (tokens) that retain all the essential information about the data without compromising its security. It's particularly useful for protecting payment cards or personal identification numbers.
- **Redaction:** This technique removes or blacks out specific data within a dataset. For instance, part of a text string, like an email address or social security number, might be replaced with X's or asterisks.
- **Format Preserving Encryption (FPE):** FPE encrypts data in such a way that the output (the encrypted data) has the same format and length as the input (the original data). This is particularly useful for protecting structured sensitive data like credit card numbers, as it maintains the original data format while ensuring security.
- **Constants:** This technique involves replacing sensitive data with a constant value. For instance, every instance of a name or ID number in a dataset might be replaced with a specific, consistent value like "Name" or "1234." It is useful in scenarios where the data value is less important than the data format or structure.
- **Referential Masking:** Referential masking maintains data consistency across different tables or databases. For example, if a customer ID is masked in one table, referential masking ensures the same masked value is used wherever that customer ID appears in other tables. This maintains data integrity and relationships within the database.
- **Dynamic Data Masking (DDM):** DDM masks data on the fly as it is queried from the database without altering the actual data stored on the disk. It allows different users to see different data based on their permissions, ensuring that sensitive data is only visible to authorized users.

Each of these techniques adds a different layer or method of protection, and their effectiveness can vary depending on the nature of the data and the specific requirements of the system in which they are implemented. Choosing the right data masking technique, or a combination of techniques, is crucial for achieving the desired balance between data utility and data security.

Each of these data masking techniques has its own advantages and is suitable for different scenarios depending on the data sensitivity, the required level of protection, and the context in which the data is used.

Core Capabilities of an Enterprise Data Masking Product



An effective data masking solution is critical for maintaining data privacy and security in today's digital landscape. According to Gartner Research, a comprehensive data masking solution should encompass four key components:

- **Data and Relationship Discovery:** This component is essential for identifying sensitive data that requires masking. It involves scanning databases and applications to locate sensitive elements like personal identifiers, financial information, or confidential business data. Equally important is understanding the relationships between different data elements, ensuring that masking rules preserve data integrity and consistency across the entire dataset. Effective discovery mechanisms enable organizations to pinpoint which data needs protection and how it interconnects within their systems.

- **Data Masking Rule Management:** Once sensitive data is identified, the next step is defining and managing the rules for how this data will be masked. This involves establishing policies for how data should be anonymized or pseudonymized. Rule management should be flexible yet robust, allowing for the customization of masking techniques to suit different data types and compliance requirements. These rules must ensure that masked data remains usable for its intended purpose, such as testing or analysis, while effectively obscuring the original sensitive information.

- **Data Masking Operations and Management:** This aspect focuses on the actual implementation and execution of data masking. It includes the processes and technologies used to apply masking rules to the data, whether in a batch, real-time, or on-demand. Efficient operations and management ensure that masking does not disrupt business operations, maintains data quality, and can handle large volumes of data across diverse environments. This component also involves monitoring the masking process's effectiveness and efficiency.

- **Reporting for Compliance:** In an era of stringent data privacy regulations, reporting is crucial for demonstrating compliance with laws such as GDPR or CCPA. An effective data masking solution should provide comprehensive reporting capabilities, documenting what data has been masked, how, and when. This helps in auditing and compliance reporting and provides visibility into the effectiveness of the data masking strategies. Clear, detailed reports are essential for organizations to verify that they meet their data protection obligations and to provide transparency to regulators and stakeholders.

Having now reviewed the key drivers to protect sensitive data, data masking and the techniques, and the core capabilities of an effective data masking solution, it is clear that choosing the right tool is paramount for enterprises aiming to safeguard sensitive data in non-production and analytical environments.

This is where Solix Data Masking comes into the picture. As a comprehensive multi-cloud solution, Solix Data Masking aligns seamlessly with the key components outlined by Gartner. It offers advanced data and relationship discovery capabilities, robust rule management, efficient operations, and thorough compliance reporting. The next section will delve deeper into Solix Data Masking capabilities and how they help meet the sophisticated needs of modern enterprises.

Introducing Solix Data Masking for Modern Data Challenges

In a time where data security threats loom large in non-production, analytical, AI/ML, and LLM environments, Solix Data Masking stands out as an all-encompassing solution. Tailored to address the intricacies of contemporary data landscapes, Solix Data Masking presents a unified strategy for protecting sensitive data across diverse platforms.

Key Features of Solix Data Masking

Support for Structured and Unstructured Data: Recognizing the diversity of data types in modern enterprises, Solix Data Masking is equipped to handle both structured and unstructured data.

- **Sensitive Data Discovery:** Solix Data Masking offers an advanced sensitive data discovery process across a wide range of structured and unstructured enterprise data repositories. With predefined and customizable patterns, the discovery feature is particularly beneficial when the organization is unsure about the presence of sensitive data and the scope of data protection needed.

- **Targeted Masking:** For use cases where organizations are aware of the location of sensitive data, Solix Data Masking offers targeted masking capabilities, allowing organizations to apply specific masking rules to identified data sets without running a sensitive data discovery process.

- **Integrated Sensitive Data Discovery-led Masking:** This integrated approach ensures that the discovery and masking processes are available as part of an integrated workflow, enhancing efficiency and reducing the risk of exclusions.

- **Support for a Wide Range of Masking Techniques:** Solix supports various data masking techniques like format-preserving encryption, tokenization, obfuscation, redaction, and many others, providing flexibility to meet different data security requirements.

- **Preconfigured customizable set of Discovery and Masking Rules:** Solix has a rich library of preconfigured patterns, simplifying the implementation of discovery and masking rules and speeding up the process.

- **Parallel and Sequential Masking:** Solix offers both parallel and sequential masking options to maintain referential integrity while ensuring high performance, making it suitable for complex and high-volume data environments.

- **Audit & Compliance Reporting:** Solix Data Masking includes comprehensive audit and compliance reporting features. These reports provide a detailed account of masking activities, including what data was masked, how, and when, which is vital for audit trails and verifying regulatory compliance.

- **Multi-Cloud Solution:** Solix Data Masking is designed to operate seamlessly in on-prem and multi-cloud environments, providing flexibility and scalability to enterprises leveraging cloud services from multiple providers. This capability ensures consistent data security across all cloud platforms, which is crucial for organizations embracing distributed and Hybrid architectures.

In summary, Solix Data Masking stands as a pioneer in the data masking market, offering a robust, flexible, and multi-cloud solution tailored to the needs of modern enterprises. Its capabilities in sensitive data discovery, support for numerous masking techniques, support for all data types, audit, and compliance make it an indispensable tool for organizations navigating the complexities of data security in a multi-cloud environment.

We invite you to explore Solix Data Masking and how it can fortify your organization's data security posture.

Analyst Recognitions

2023 MARKET GUIDE Data Masking — Gartner

2023 LANDSCAPE REPORT Data Security — Forrester

2023 LANDSCAPE REPORT Data Privacy — Forrester

2022 LANDSCAPE REPORT Data Governance — Forrester

REQUEST DEMO

VISIT WEBSITE

CONTACT US

info@solix.com

+1.888-GO-SOLIX

THE WORLD'S LEADING COMPANIES CHOOSE SOLIX

citigroup · Alberta Health Services · experian · OPTUM UNITEDHEALTH GROUP · GE APPLIANCES a Haier company · JUNIPER NETWORKS · Santander · BAE SYSTEMS INSPIRED WORK · IRON MOUNTAIN · MOLSON Coors · SONIFI · Unilever · سابك sabic · AIG · HCSC Health Care Service Corporation

4701 Patrick Henry Dr., Bldg 20,
Santa Clara, CA 95054.

Toll Free: +1.888-GO-SOLIX
Email: info@solix.com