ACCUTIVE SECURITY

The Auth + Crypto products and services company

White Paper

ENCRYPTION OR DATA MASKING: EVALUATING STRATEGIES FOR GDPR COMPLIANCE

Optimized Data Protections Ensure Both Privacy and Usefulness of Data Enterprise-Wide

Executive Summary

This white paper evaluates strategies for data privacy compliance, especially under the EU's GDPR, contrasting encryption and data masking techniques. Encryption, while secure, renders data unusable without decryption, necessitates secure key management, and results in significant overhead. Data masking, however, retains data properties while making data unidentifiable, offering usability for development and testing purposes without the need for decryption keys. Enterprise data masking is particularly beneficial in non-production environments where securing data can be costly. Masking ensures the generation of 'almost-production-like' data that keeps enterprise systems functional in testing phases while significantly reducing data compromise risks. It also aids in adhering to regulatory requirements, offering cost-effective, performance-efficient data protection while preserving data value and consistency across different sources. Automated discovery, scalability, high-volume handling, and user access segregation underline an optimal data masking solution. Finally, data masking lays a flexible, enterprise-wide data security foundation, creating fictitious but functional data, aligning with risk management requirements, and maintaining operational smoothness.

Optimized Data Protections Ensure Both Privacy & Usefulness of Data Enterprise-Wide

The EU's GDPR (General Data Protection Regulation) is the latest challenge facing the data-driven enterprise – and perhaps the most demanding to date, extending significant data protections to EU citizens no matter where they work, live, or do business. This is a major requirement for data controllers, and begs the question 'how does GDPR fit into the spectrum of enterprise data management techniques and technologies?' Compliance cannot be achieved with a single, one-stop GDPR solution, as is also true with regulations such as HIPAA, PCI-DSS, GLBA, OSFI/PIPEDA, and FERPA. Ideally, compliance is handled more holistically, creating an environment of preparedness and data privacy with a smart, long-term approach to data protection. This paper will evaluate enterprise data masking as a technique to address this goal, contrasting it to encryption and demonstrating its value in maintaining data security and privacy, while ensuring data continues to be both useful and safely accessible across the enterprise.

Considering Encryption or Data Masking

Encryption is a process that "scrambles" data, so that only people with a secure key can read it. Data is turned into cipher text, and its corresponding key works like a map to unscramble the information back into plain text. The process is safe and secure and prevents data from access by unauthorized individuals; however, encrypted data is also unrecognizable and unusable, meaning it must be decrypted (and therefore exposed) for use in reporting or analytics. For example, hashing algorithms that perform only a one-way operation are not a viable option. To remain usable across the enterprise, data must be able to be decrypted and returned to its original value.

It is critical that the encryption keys used in this process are kept safe from compromise. This creates the need for additional security protocols and processes, increasing overhead associated with management and storage of keys and related digital certificates. Costs of data acquisition and maintenance are generally high with encryption.

Solutions based on format-preserving encryption, or tokenization, yield some improvement over traditional encryption by replacing data with similarly structured but random replacements. However, as with encryption, this requires management of key protection and lifecycle processes. Additionally, it adds the overhead of yet another protected storage location that retains the link between the encrypted data and its replacement. This data store must be protected by protocol and processes that are at least as effective as what is deployed to its most sensitive production data source. Depending on regulatory or enterprise security requirements, the downstream data may also require some level of protection, as it contains information that can be linked back to the original data source.

Cont. p4

Data Masking vs. Data Encryption

DATA ENCRYPTION	DATA MASKING
Encryption by default makes the data completely unreadable and unusable for development and testing purposes	Data maintains original formatting—it is readable and usable for development and testing purposes
Requires encryption keys, which may be compromised	Uses random data generation; data cannot be reverse engineered
Impacts performance—is time consuming to encrypt large datasets	Does not impact performance
High overhead, particularly with regard to the management of keys or certificates	No keys or certificates to manage
High acquisition and maintenance costs	Lower acquisition and maintenance cost

In contrast to encryption, enterprise data masking provides the ability to locate and mask critical, sensitive data, while ensuring the data properties and fields remain intact across any number of sources.

Enterprise data masking is inherently different from encryption, retaining statistical properties, formatting integrity, and realism of actual data, without retaining any link between the original data and its replacement. Once processed through the masking operation, data appears real, but has in fact become fictitious. While the output data remains fully functional, the information is no longer sensitive. For example, a 16-digit credit card number in the source data remains a 16-digit credit card number after being masked. This is a critical advantage in meeting data controls of regulations like GDPR: data remains both valuable and private throughout reporting and analytics, even as it is disconnected from any PII across the enterprise customer base. There are no encryption keys to manage or store, and no performance limitations compared to the processing required for encryption and decryption of large datasets.

As a result, costs are reduced relative to security operations as well as data acquisition and maintenance. Data masking also ensures the value of data to the enterprise, providing that the masking solution can keep masked data consistent throughout all sources. For example, the fictitious credit card number ideally remains the same, whether it is appearing in a CRM platform or billing application. Joe Smith masked as John Doe should always be John Doe in each enterprise system using the data.

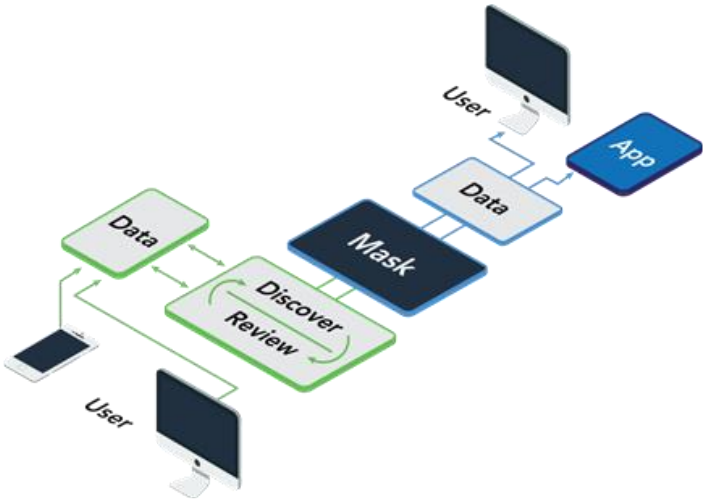
Securing Non-Production Environments

Data is used by two primary system families: production, or customer-facing, and non-production, which includes applications such as development, testing, or third-party use. Because it helps ensure optimal results, organizations often must copy or move sensitive data containing PII and payment card information from the security of their production environments to less secure, non-production environments. While encryption is a common technique used to protect data in production environments, enterprise data masking is better suited to non- production settings which are potentially difficult and costly to secure. Encrypted production data can be decrypted using the right key; therefore, any system containing the same encrypted data may be subject to the same regulatory oversight as production systems, even if the encryption key is changed between environments.

Consider the role of a testing environment and its need for representative data. Without masking, an organization needs to recreate data scenarios that cover the spectrum of sample customers and testing needs. This increases the likelihood of application failures, as a result of an incomplete range of data use cases. Because of this challenge, data masking takes a leading role in securing non-production settings—allowing realistic, useful, and 'almost-production-like data' to be applied as needed across the enterprise. Systems which communicate using a common identifier, such as customer ID, can continue to function in a test environment when the data is masked consistently. Most importantly, data masking reduces the risk of data compromise, while ensuring the enterprise meets regulatory requirements for data protections.

How Data Masking Works

Ideally, enterprise data masking performs automated discovery of sensitive data. By automatically reading and reporting data source metadata information and sampled data sets, optimized solutions can mask content between any source and destination data source, such as Oracle, DB2, MySQL and SQLServer databases, NoSQL databases, flat files, spreadsheets, and other methods of structured and unstructured data storage. All source tables and field names, as well as a subset of the actual data, are searched and mapped during discovery operations, and potentially sensitive matches are flagged for inspection. During masking, the solution maintains all field properties of the source data. After masking, data looks real but is merely representative. At the same time, the algorithms used in data masking generate random data and are not reversible, preventing the original data from being recovered by any means.



Optimized enterprise data masking is also scalable and able to meet high volume testing needs of enterprise datasets. The solution periodically refreshes and maintains masked data across different environments, ensuring a repeatable, consistent, and automated masking operation. User access is ideally managed by the masking application, segregating the duties of security administrators, database administrators, and application testers. Audit controls must be included, monitoring and tracking user activities so they can be reviewed, analyzed, and reported.

Accutive Security's Data Discovery and Data Masking solution is built upon the reliable and scalable Java platform which is optimized for high performance. It is a web-based, drag-and- drop user interface for rapid and simplified configuration and field mapping.

Interpreting Levels of GDPR Risk

Data masking offers a flexible, enterprise-wide foundation for data security, creating fictitious data that retains field properties and is fully functional across multiple systems. This is an advantage for an enterprise working to maintain smooth data operations while addressing the risk management requirements of various regulatory agencies and standards.

For example, GDPR requires that organizations shift to risk-based data controls, implementing protective measures that correspond to the level of risk found within their data processing activities. While levels of risk may be up to interpretation, the onus is on the enterprise to put control of data back in the hands of consumers, ensuring capabilities such as opt-in permissions for communications and accommodating an individual's right to be forgotten. Answering these issues requires the organization to ask, 'do you know where your data flows within the organization?' This includes examining not only potential data silos, but also information on who is accessing and using data. The enterprise must carefully consider whether its data is protected in each separate application environment, each endpoint (which may include BYOD), and all communication channels.

Deeper examination may reveal that risk is increased as data flows include endpoints such as executive desktops or administrative support teams. This scenario can be as simple as staff tasked with creating reports based on specific data. Production data may commonly be extracted for use in an Excel spreadsheet, stored or viewed on a mobile device, with a file sharing service, or in a meeting. Real information is typically required for activities such as analyzing customer behavior or business trends, but it must be protected at the same time. Data masking accommodates data used in these types of individual applications—information can look the same in finance, HR, or marketing, and still be masked for compliance and privacy.

Masking Drives a Proactive Security Posture

In a landscape marred by major data breaches across a range of consumer markets, security and compliance of data have never been more important. Further, enterprises facing the strenuous new requirements of GDPR may already be dealing with data loss or exposure related to issues of staff negligence, errors, or unclear practices, rather than cyber threats.

The world is familiar with external threat actors, and too often, reports of criminal organizations generating billions of dollars in revenue from data breaches and other cybercrimes. But there are other major dangers to the enterprise, spanning intentional actions and inadvertent flaws in technology. Insider threats, such as standard and privileged users, are real and quite difficult to detect and prevent. Data is sometimes held within an organizations with few controls, for example unstructured data that may exist on multiple end-user devices in multiple formats. This can be a weak spot and therefore a risk, along with legacy systems and applications that may contain serious unrecognized vulnerabilities that leave sensitive data exposed to cyber threats.

> Compliance should not be an afterthought to ensure one is within the confines of the law, but instead be baked into core product and operational design in the first place.

Entities such as the GDPR standards body may consider intent behind data losses, but only to the extent of determining whether the infringement was intentional or merely negligent. Language of the ruling demands “data protection by design and by default,” and any lenience will only be reflected in whether fines are levied above or below the 10 million euro level. The GDPR standard sets a clear expectation that security and privacy controls must be a priority, stating “compliance should not be an afterthought to ensure one is within the confines of the law, but instead be baked into core product and operational design in the first place.”

In this context, enterprise data masking demonstrates a powerful commitment to securing data across the enterprise. It is a high-value technology investment that may play an important role in creating a defensible position of compliance and customer protection.

Thinking Beyond Compliance

GDPR is a significant departure from a landscape already rich in regulatory demands, with deep requirements and severe penalties for lack of compliance. Its success may drive similar data protections relative to other citizens of the world, and the regulatory environment will certainly continue to advance globally. Yet for the enterprise, keeping up with data protection that is driven by one compliance need at a time is reactive, costly, and dangerous. Implementing 'rule by rule' strategies force a game of catch-up with evolving data regulations, compounded by continued threats such as external forces, internal agents, legacy data and applications, and simple but dangerous human error.

Encryption continues to play a role in meeting these challenges; used appropriately, it will aid an enterprise in achieving GDPR compliance. However, its limitations become more evident as organizations better understand the need to protect data outside of production environments. Accutive Data Discovery and Data Masking offers a secure and easy-to-use alternative, keeping data hidden but useful as it is used in non-production settings enterprise-wide, such as reporting and analytics. Offering a more holistic approach, enterprise data masking keeps data secure as well as accessible. This helps meet regulatory demands and maintain the value of data to an enterprise, creating an environment of data protection as a standard of operation.

About the Author

Paul Horn, CISSP, is CTO of Accutive (www.accutive.com). His expertise in data protection strategies and public key infrastructures spans more than two decades in enterprise software development, operations, data security, and encryption. Contact him at paul.horn@accutive.com.

CONTACT INFORMATION

27068 La Paz Road, Suite 245 Aliso Viejo, CA 92656

Toll Free 888.666.8315

Contact Us

www.accutivesecurity.com

Accutive Security provides comprehensive information technology solutions which includes enterprise security; IT application development; systems integration; lending platform services; and professional services to clients in a range of industries like financial services, healthcare, technology, entertainment, and retail. Accutive Security is a specialized team of experienced professionals with years of industry, business, and IT expertise dedicated to developing IT solutions that exceed our clients' expectations.