DATA SECURITY TECH BRIEF

Data Security Technology Comparisons

Data encryption, tokenization and masking – what they are and when to use them

Numerous data protection solutions exist in the marketplace today – each designed to protect sensitive data in different ways – making it nearly impossible for data scientists or IT professionals to know which data protection solutions are the best fit for their needs. In this paper, we'll look at three of the more common data security technologies and how they apply to various use cases.

1 Data Encryption

What it is

Data encryption is the process of using an encryption key to alter data to render it unreadable to anyone who does not possess the encryption key or who is not an authorized user. The original, sensitive data or plaintext is encrypted via sophisticated algorithms that convert it to unreadable text or ciphertext. Several algorithms exist for encryption, with the most sophisticated being AES, followed by DES, ECC, QKD, and more.
A decryption key is necessary to revert the encrypted data to a readable format.

Where tokenization uses a token to protect data, data encryption uses a key. Tokenization swaps sensitive data for an irreversible, non-sensitive token and stores the original data outside of its original environment. Encryption encodes the content of a data element where it resides with a key shared between those with the right user access controls.

When to use it

Data encryption traditionally has been used to protect data at rest or data-in-motion. However, new encryption solutions – known as data-in-use encryption, also protect data while it is being used or queried. Encryption has multiple use cases, starting with decreasing the impact of

a possible ransomware attack, allowing data analysts both internally as well as with third parties to analyze encrypted data without compromising security, as well as protecting from data theft and data breach.

A relatively new use case is the ability to share data with partners and third parties to take advantage of new business opportunities without being concerned that any data can be compromised.

Until recently, organizations frequently avoided encryption solutions as they were difficult and time-consuming to deploy, and they did not allow for encrypted data to be queried or analyzed. But data-in-use encryption removes these pain points by requiring no changes to applications; by having next to no latency; and by not requiring a team of experts to manage the solution. In addition, with data-in-use encryption, encrypted data can be queried or analyzed without the need to decrypt, and they enable organizations to encrypt and manage data in multiple data stores from a centralized platform.

To learn more about data-in-use encryption, here's a white paper that takes a deep dive into in-use encryption, what it is and how it works.

2 Tokenization

What it is

Originally, tokenization was launched to protect payment card data to help retailers reduce their obligations under PCI-DSS. Tokenization converts a data placeholder to a token placeholder, replacing sensitive elements with randomly generated data mapped one-to-one within the environment. The original information is no longer contained within the tokenized version; therefore, the token cannot be easily reversed back to the original sensitive data.

Tokenization uses a database called a token vault that stores the relationship between the sensitive element and the token. No key can reverse the tokenized data back to the original data values without accessing the original data file, as no mathematical relationship to the real data tokens exist. Unlike encryption, tokenization doesn't use a mathematical process to replace sensitive data with a token. Instead, tokens consist of meaningless placeholder text that replace sensitive elements.

When to use it

Tokens can be used in applications to replace highly sensitive data. When the original data must be retrieved, the token is submitted to the vault. From there, the index is used to fetch the real value for use in the authorization process. Tokenization is frequently used as a solution for protecting sensitive elements when there is no need to access the original data. Also, tokenization is useful when the sensitive elements that are tokenized do not need to be queried. Tokenization does not, however, protect the original data file that is still stored outside of the environment.

3 Data Masking

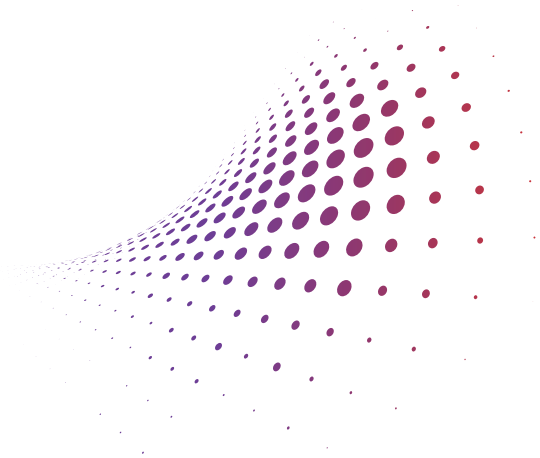
What it is

Data masking is frequently referred to as data anonymization. Data masking replaces original, sensitive data by using fictitious data or characters. Several types of data masking exist using various masking techniques such as static data masking (SDM) and dynamic data masking (DDM). Static data masking copies the original data into a test environment, from which the masked data can be shared with third parties. Dynamic data masking turns the original dataset into masked data.

Typically, algorithms mask sensitive data and replace all values with the same type of data as well as the same data length. This means the data is defined as structurally identical, as in, if a social security number is 9 digits long, the masked data will also be 9 digits long. It is important to note that there is a slim chance that a person can identify original data based on trends in masked data. This means that data masking is not the best solution for sharing data with third party vendors. The original data can be accessed out of the masked data via a reverse proxy.

When to use it

The use case for masking data is usually tied to compliance requirements, specifically for a rather quick solution for organizations that must adhere to GDPR, CCPA, or ITAR regulations. A secondary use case is for a test environment with identical data when original data is not needed to test the database environment.



→ **To learn more about your data security options, we invite you to contact Sotero to speak with a data security solution specialist. Click here to contact us and we'll get right back to you.**

ABOUT SOTERO

Sotero is the global innovator and leader in next generation data security. Sotero's KeepEncrypt™ solutions secure your data by encrypting data at rest, in motion, and in-use with virtually no latency or impact on user experience. Securing data "in-use" enables organizations to realize higher returns and faster time-to-value from their data for mission-critical business use cases, including data analysis, data sharing, and product development.

Learn more at: www.soterosoft.com

Sotero

99 S. Bedford Street, Suite 106
Burlington, MA 01803
www.soterosoft.com