TECHNICAL OVERVIEW

Tokenization: The Data Substitute That's Best for Business

Just as a spy can be a master of disguise, data can also shift its shape to avoid detection.

When tokenized, data assumes another identity. Its sensitive elements are hidden from those who could use the information for nefarious purposes. But tokenization isn't reserved for spies. Organizations of all types and sizes can use this protection method to safeguard valuable data while still using it for business purposes.

A cloak-and-dagger approach to data security makes sense nowadays. Cybercriminals have all sorts of tricks up their sleeves–phishing attacks, web redirects, and even simple password guessing–and they can choose from dozens and dozens of inexpensive and sometimes free tools that can be used by even the most inexperienced hackers. Because traditional cybersecurity doesn't protect data as it moves between on-premises systems and cloud-based applications and servers, it must be protected end to end in its travels. Tokenization ensures sensitive data isn't left unprotected on its journey toward delivering valuable business insights.

Tokenization conceals sensitive data elements so that data sets can go to work in transactional, computational, and analytical systems. The process transforms a piece of sensitive data into a randomly generated value that has no mathematical correlation with the real value. That way, if an organization's data is ever breached, the visible tokenized data means nothing.

The Data 'Coat Check': Tokenization Retrieves Hidden Data

Considering that tokenization substitutes data, it might be easier to understand how it works by first considering a metaphor that involves substitution: a coat-check service.

A coat check follows a simple process. Before dining at a restaurant or hobnobbing at a social function in a hotel, people hand their coats, purses, and other valuables to a coat checker, who, in turn, hands them a ticket that verifies their possessions are tucked away in a safe place. With a number on the ticket corresponding with a number attached to the stored items, dining patrons and party guests retrieve what is rightfully theirs when leaving.

Tokenization is a coat checker for data as it travels between applications, devices, and servers. But instead of relying on a ticket to recapture the data that's tucked away, a digital token is the key to reclaiming the valuable data. As soon as a user with authorization needs to access the sensitive data elements, a token affixed to that data is used to reveal it, much as a coat-check ticket enables people to retrieve their valuables.

Cleartext vs. Tokenized Data

jsmith@protegrity.com → uycdoi@sioakvntis.com

In its most basic form, tokenization simply substitutes a randomly generated value, a token, for a cleartext value. A lookup table, or token vault, is kept in a secure place, mapping the cleartext value to the corresponding token.

The token data type and length usually remain the same as the cleartext value, and the token lookup table becomes a key, allowing the cleartext value to be retrieved from the token. Tokenization is reversible and a popular protection method for organizations that need to safeguard individual fields of data in transactional or analytical systems because the data type and length do not change.

Managed Though a Single Pane of Glass, Protegrity Vaultless Tokenization

Tokenization can be used to de-identify many types of sensitive data, from personally identifiable information (PII) and credit-card details, to business-related intellectual property. Yet, as tokenized datasets grow, and as IT infrastructures become increasingly complex, dynamic, vault-based token-lookup tables quickly become unmanageable. Organizations that rely on vault-based tokenization might not initially notice, but eventually the process takes longer and becomes unmanageable. It's not an ideal protection method for organizations that expect data to immediately be available for use.

A more sophisticated form of tokenization, Protegrity Vaultless Tokenization (PVT), solves the time and capacity challenge. PVT uses small, static token tables to create unique, random token values without the need for a dynamic, vaulted token-lookup table. Instead, users benefit from a highly scalable, flexible, and powerful protection method for structured and semi-structured data. Protection is applied at the data point, and the value of the token is based on a codebook that remains consistently responsive no matter how much data has been previously protected.

But it's not just PVT that empowers businesses to make the most of protected data. Protegrity's "single pane of glass" approach means the tokenized data of a global enterprise or mid-sized business is managed through a single interface, reducing the chance for error and baking consistency into policy enforcement. Authorized users see a centralized view of Protegrity protection at work. With that knowledge, organizations can set policy rules to determine who sees and doesn't see sensitive data and maintain control over who can and can't access the data–no matter if it's in use, in motion, or at rest.

The Ideal Data-protection Method for Data-driven Organizations

Businesses use data in different ways, and the applications they use determine the makeup of the data. Protegrity Vaultless Tokenization protects those many data types: the structured data that's fed into transactional systems such as ATMs, CRM systems, and inventory management systems; and the unstructured data of emails, word processing documents, PDF files, photos, and many other formats.

Banks that need to protect credit-card numbers in data, for example, will typically choose tokenization, with randomly generated numbers (tokens) replacing the primary account numbers, while the process preserves the format and length of the data–simplifying its use in analytics. Health-care organizations also turn to tokenization to protect data in analytics. A hospital can tokenize data on, say, twenty platforms to observe HIPAA standards, and de-tokenize data on one analytics platform, no longer hiding the sensitive data elements to those who are approved to view it.

The need for data protection has never been greater. Just scan news headlines any day of the week and there's word of yet another corporate data breach. The insurance company Geico suffered a breach that exposed, for more than a month, customers' driver's license numbers. That headline followed news that the personal information of more than 533 million Facebook users were stolen and leaked on a popular hacker forum.

Facing strict data regulations from governments and heightened expectations for data privacy from individuals, organizations have to effectively protect data–but they also can't tuck it away and ignore its immense value in delivering business insights. Tokenization lets them safely put aside sensitive elements of data but still tap larger data sets to power analytics, AI-supported initiatives, containerization, and other applications that drive business.

Because of the encompassing security of tokenization–and the efficiency of PVT–data-driven organizations can check sensitive data at the door and retrieve it on the way out.

PROTEGRITY

Protegrity USA, Inc.
(Global Headquarters)
1165 E Wilmington Avenue
Suite 200
Salt Lake City, Utah 84106
1.203.326.7200
1.650.431.7000

Protegrity (Europe)
1 St Katherine's Way
London, E1W 1UN
+44 1494 857762

Protegrity (Asia Pacific)
1 Nanson Road, Level 3
Singapore 238909
+65 6904 6063

Email: info@protegrity.com

© 2021 Protegrity. All rights reserved. Protegrity protects the world's most sensitive data–whatever it is and wherever it resides at any given moment. Our platform frees businesses from the constraints typically associated with the access and fine-grained protection of sensitive data so they can have the confidence to create better customer experiences, make intelligent decisions, and fuel innovation. Data knows no boundaries, nor should data protection. Protegrity's technology is built for data ubiquity–we secure data everywhere and empower businesses to thrive.