DataStealth

Case Study: Securing Legacy Mainframe Data with DataStealth

Customer Profile: A financial services or telecommunications company operating IBM DB2 instances on a mainframe environment.

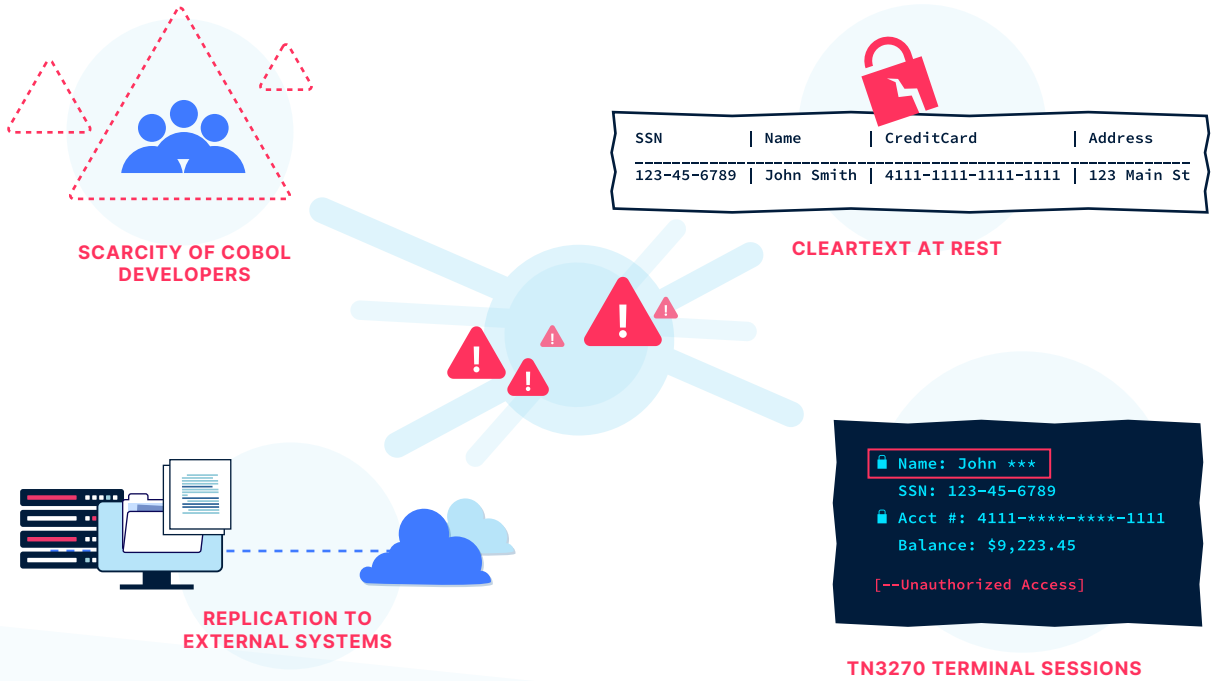
The Challenge: Unlocking Mainframe Data Securely

Organizations in sectors like financial services and telecommunications often rely on mainframes to store vast quantities of historical customer data. While this data is critical, it frequently resides in cleartext, posing a significant security risk.

The inherent complexity and age of mainframe application code (e.g. COBOL) and legacy database structures make modifications high-risk, expensive, and time-consuming. Adding to this issue is the scarcity of skilled mainframe developers.

A key concern is that organizations typically want to avoid installing agents directly on mainframes because doing so can impact performance, introduce operational risks, and is often unsupported or restricted. Furthermore, sensitive data is often replicated from the mainframe to other systems.

This replication increases the attack surface, particularly when data traverses trust boundaries or different security zones. These downstream systems may also incorporate external enrichment data, such as geolocation or behavioural information, which must be protected with the same rigour as the original mainframe data.



Accessing data on mainframes introduces unique security challenges due to legacy protocols like TN3270, which are still widely used for terminal-based sessions.

Supporting TN3270 terminal sessions is a critical yet complex requirement for protecting data on mainframes. These legacy access methods allow users to interact directly with sensitive data, often out of reach for modern security controls. Because TN3270 transmits data in real-time to user screens, protecting information at the point of display becomes essential.

This requires role-based enforced dynamic masking, and the ability to apply policies inline, all without disrupting the user experience. The broader challenge lies in achieving unified visibility and control across both database replication and terminal access so that consistent, robust data protection can be enforced at every point where data leaves the mainframe.

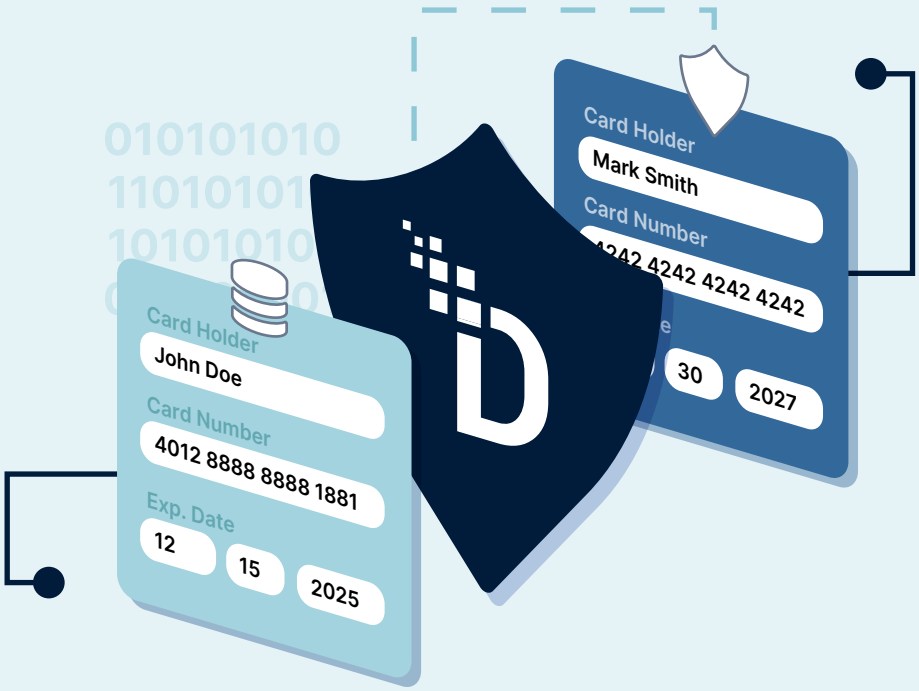
The Solution: DataStealth's Agentless Mainframe Protection

The company implemented **DataStealth's Data Discovery and Classification (DDC)** and **Data Tokenization (DT)** to address these multifaceted challenges without requiring intrusive code changes to its legacy mainframe environment.

Agentless Mainframe Interaction

DataStealth's approach is rooted in agentless mainframe interaction.

The solution was deployed inline, enabling communication with the mainframe using native protocols such as TN3270 for terminal access and relevant database protocols for DB2. This eliminated the need for high-risk software installation or code alteration on the mainframe itself.



In-Place Tokenization

A key component of the solution was in-place tokenization, which allows sensitive data to be protected at rest and in use, without changing its format or location. This approach ensures that data remains usable for business processes and analytics while reducing the risk of exposure.
By tokenizing the data where it resides, the organization avoided the complexity and risk of moving or duplicating sensitive information, aligning directly with the goal of protecting data without disrupting operational workflows.

DataStealth endpoints were configured to read data from the mainframe DB2. Vaulted tokenization was used to protect sensitive data: the original values are replaced with tokens stored directly in the database, with no mathematical relationship to the original data. This approach ensures that sensitive data is never exposed in clear text, supports format preservation, and is inherently quantum-resistant due to the absence of any computational linkage between the token and the real value.

An alternative approach was considered: keeping the data in clear text on the mainframe and applying protection dynamically during access, replication, or transmission; using methods like dynamic masking or generating test or synthetic data on the fly. While this option offered flexibility and avoided changes to the original data, it was not chosen due to the added complexity, the risk that protection might not always be applied, and a stronger need to keep the data securely protected at rest.

Format Preservation

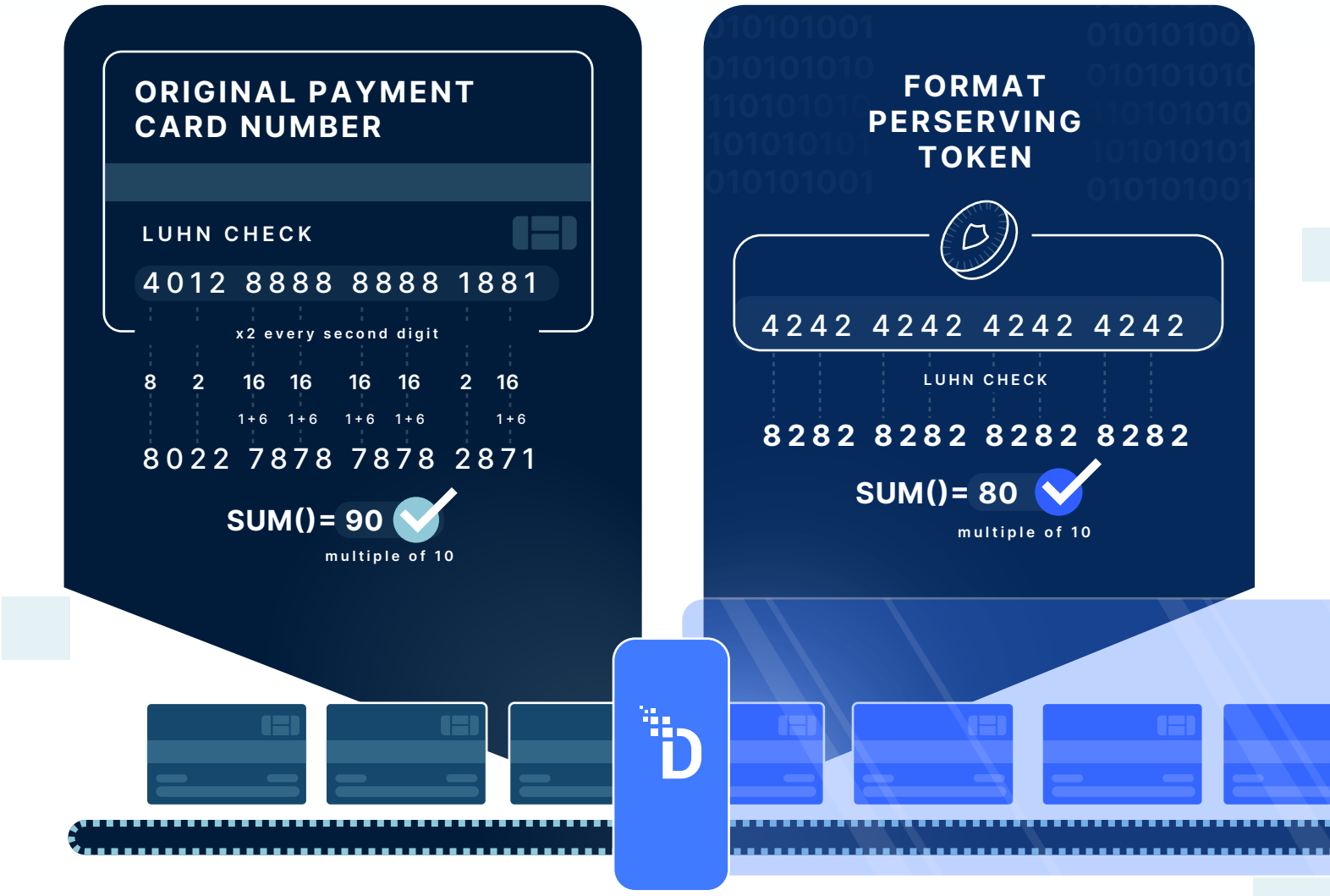
Format preservation during tokenization was crucial.

This is vital for legacy systems like mainframe DB2, where altering schemas or data formats can be exceptionally risky or entirely unsupported.

Many mainframe databases employ data integrity checks, such as validation constraints or Luhn checks for credit card numbers, before committing updates.

DataStealth's vaulted tokenization method generates tokens that preserve the original data's structure and pass these business logic rules, ensuring that system workflows continue uninterrupted.

For instance, a tokenized credit card number would still pass a Luhn check, preventing the system from rejecting it or failing in downstream processes.



Controlled Replication

Controlled replication was a critical component in ensuring secure data movement across systems while maintaining compliance.

DataStealth endpoints intercepted replication flows from the mainframe DB2 to downstream systems—such as an Oracle database used for fraud detection—to enforce data protection policies in transit. This allowed the organization to leverage de-identified data for advanced analytics without exposing sensitive data in cleartext, ensuring that privacy and regulatory requirements were upheld throughout the replication process.

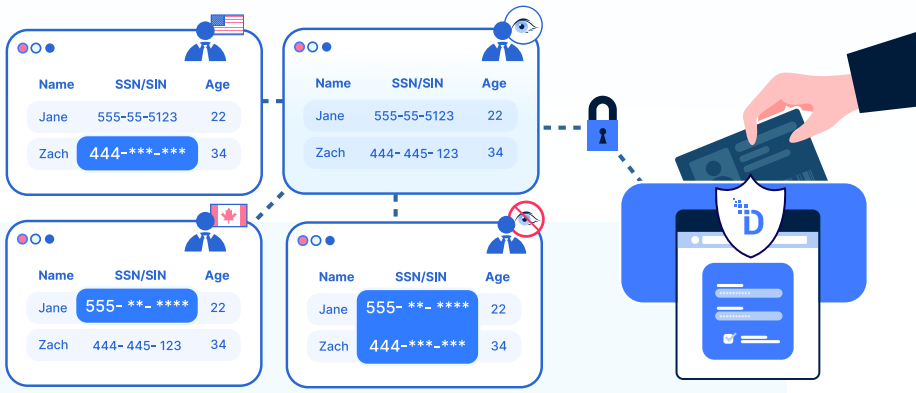
Data protection policies were applied to secure sensitive data throughout this process, even if the data had already been tokenized on the mainframe. This additional layer of protection helps ensure consistent enforcement of policies across environments. DataStealth uses vaulted tokenization, where the mapping between each original value and its token is stored in a secure, centralized vault. These mappings are unique to each vault, meaning tokens generated in one vault cannot be resolved in another. This separation is intentional; it enforces strict data boundaries between environments or security zones, reducing the risk of unauthorized re-identification or cross-environment data leakage.

Depending on the configuration and the destination system, several actions could be taken:

- **Pass-through:** If the target system shared access to the same vault, tokenized data could be passed through directly.
- **Controlled detokenization:** If the target system required access to the original cleartext values such as for processing, analytics, or integration with legacy applications then detokenization could be performed. However, this was done under strict controls, since exposing sensitive data, even temporarily, increases the risk of data leakage or misuse. Access was typically governed by fine-grained policies, audit logging, and time-limited access to minimize the exposure window and ensure compliance with data protection requirements.
- **Re-tokenization:** More securely, data could be detokenized from the source vault and then re-tokenized using a different vault associated with the target system. This preserved security boundaries while maintaining reversibility within each domain.

Vaulted tokenization differs from **vaultless tokenization**, which relies on mathematical algorithms to generate deterministic tokens without storing the original value. While vaultless methods for tokenization avoid the need to manage a centralized vault, they are not quantum-resistant and offer limited policy control. For example, they typically rely on algorithms to preserve certain characters or digits from the original data such as keeping the first six and last four digits of a credit card number which can increase the risk of data inference or re-identification. DataStealth's use of vaulted tokenization ensures strong separation between environments, meets PCI requirements and allows for policy-driven control over how and where tokens can be resolved.

This re-tokenization-based approach enables secure data transfer across trust boundaries, not just by converting tokens between vaults, but also by enforcing protection policies on any external enrichment data introduced during replication. This ensures that both original and newly added data remain consistently protected throughout the process.



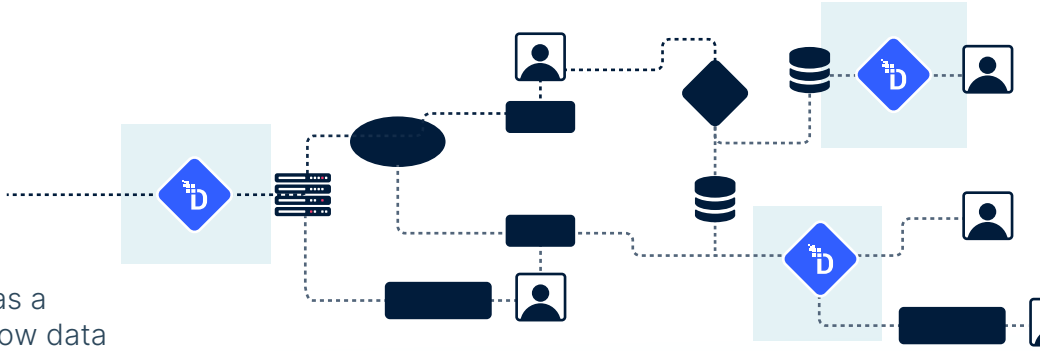
Terminal Access Control

For terminal access control, DataStealth endpoints managed terminal sessions by operating in the flow of TN3270 traffic.

Dynamic Data Masking (DDM) was selected to provide real-time protection of sensitive information without disrupting existing applications or workflows. Unlike static redaction or manual data filtering, DDM obfuscates sensitive data at the time of access, ensuring that only authorized users can view unmasked values, while others see masked content. This approach helps organizations comply with regulations like GDPR and significantly reduces the risk of data exposure or misuse.

Through integration with Identity and Access Management (IAM) systems such as Active Directory and Entra ID, DataStealth identifies users and enforces masking policies based on roles, attributes, and permissions. This enables dynamic, attribute-based masking or selective detokenization, ensuring each user sees only the data they're permitted to access. With support for both attribute-based and role-based access controls (ABAC and RBAC), DDM offers precise, adaptable policy enforcement across systems, maintaining security without compromising usability.

Holistic View of Data Flows



At the core of the solution was a complete understanding of how data moves across systems.

This involved mapping all data sources (mainframe databases, enrichment feeds), all consumers (terminal users, replicated databases), and the pathways between them.

Based on this map, consistent policies for data discovery, classification, and protection (tokenization or masking) were applied universally.

Vault Management

Vault management played a critical role in the security architecture.

Vaults securely manage the relationship between the original data and their corresponding tokens.

Depending on the configuration, the system utilized single or multiple vaults, particularly when data crossed security zones, allowing for detokenization from one vault and re-tokenization into another.

Tokens are generated and stored along with their corresponding original values in a secure vault. DataStealth supports multiple mechanisms for token generation, including the use of sequencers to ensure uniqueness and format control. Separately, a hash of the original data, optionally combined with salt, prefix, or suffix, is used to optimize performance by enabling fast lookups of previously tokenized values. The secure mapping maintained in the vault allows for authorized detokenization while ensuring that sensitive data is stored securely and access is tightly controlled.

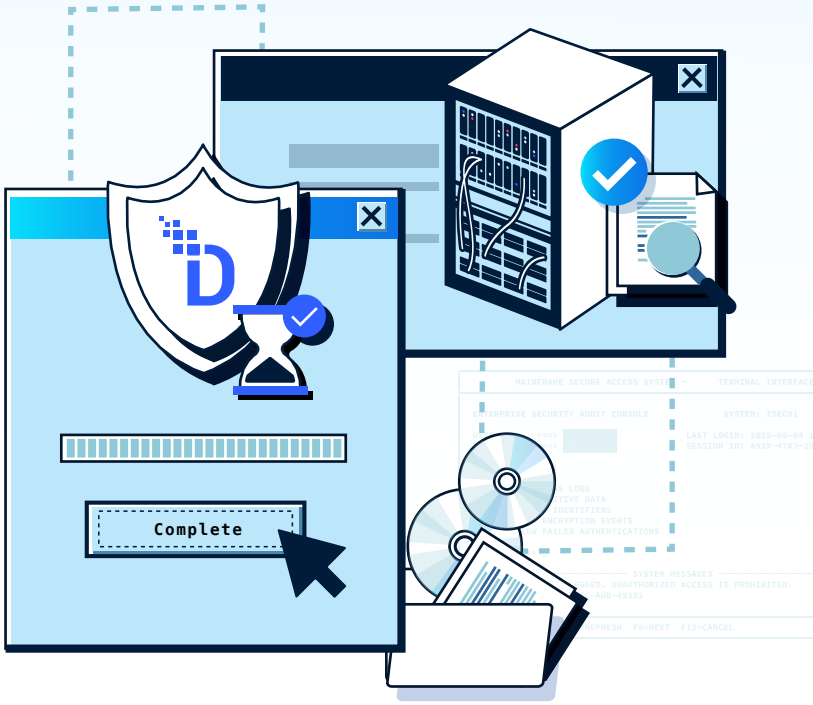
This vaulted tokenization approach is **distinct from encryption**, as tokens contain **no mathematical relationship** to the original data and no intrinsically valuable information. As a result, it offers strong security and is inherently **resistant to emerging threats such as quantum computing.**

Outcomes and Benefits

Effective Mainframe Data Protection

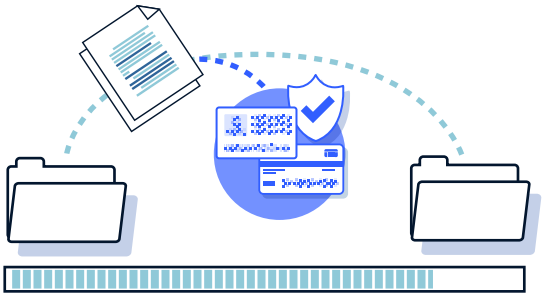
Sensitive data on legacy mainframe systems was secured without necessitating risky or complex modifications to existing applications or the underlying infrastructure.

This also eliminated the need to modify COBOL-based applications, an older programming language commonly used in mainframe systems, or any other mainframe application code, significantly reducing costs, operational risk, and the demand on scarce, specialized development resources.



Reduced Risk of Exposure

The solution minimized the exposure of cleartext data during replication, access, and transmission, significantly strengthening the company's data security.



Enhanced Compliance Support

By protecting sensitive data within legacy systems, DataStealth enabled the organization to meet stringent compliance requirements, including those related to privacy and security. The platform provided persistent data protection without requiring changes to existing infrastructure, allowing the organization to maintain operational efficiency while ensuring that sensitive data remained secure. This was especially critical for environments subject to strict regulatory oversight.

Facilitated Modernization Initiatives

DataStealth enabled secure data sharing between legacy mainframes and modern systems.

This capability acts as a crucial bridge for digital transformation efforts, allowing the company to innovate without compromising the security of its sensitive data.

Consistent data protection policies were applied across the hybrid environment, encompassing both legacy and modern infrastructure.

Controlled Data Access

Sensitive data accessed through legacy terminal sessions was protected using dynamic masking and selective detokenization, governed by user roles, attributes, and session context.

Because TN3270 protocols stream data directly to user screens, DataStealth operated inline to apply policies in real-time, masking or detokenizing fields as appropriate without modifying mainframe code. Integration with IAM systems (e.g., Active Directory) enabled fine-grained, identity-based control, ensuring users only saw data they were authorized to access.

This approach minimized the risk of unauthorized exposure while maintaining usability and performance, even in legacy environments.

About DataStealth

DataStealth (https://datastealth.io) **is a patented Data Security Platform (DSP)** and PCI Level 1 Service Provider focused on helping enterprises discover, classify and protect their sensitive data across any environment. Recognized as a leading DSP, we empower enterprises with innovative technologies to ensure compliance with regulatory and industry standards while delivering formidable protection against evolving threats.

Our DSP integrates data discovery, classification, and protection via vaulted tokenization and dynamic data masking in one platform solution. Unlike tools that only surface shadow data or flag risk, DataStealth sits directly in the flow of traffic, enforcing policy and eliminating exposure before threats materialize. We make sensitive data inaccessible to attackers by replacing it with secure, vault-referenced tokens that are quantum-resistant by design.

Deployment requires no code changes, no agents, or a need to rearchitect legacy systems. DataStealth operates transparently at the network layer, enabling seamless protection across hybrid, cloud, SaaS, and even mainframe environments.

Whether you're securing a global enterprise or reducing PCI scope across business units, DataStealth is purpose-built to scale with you, delivering enterprise-grade data protection with zero disruption

Learn more at https://datastealth.io