# Beyond Encryption: Reducing the Value of Exposed Data

By Georgianna Shea, Cason Smith, and Sydney Johnson

## Executive Summary

Cybersecurity has long relied on barriers: firewalls, encryption, identity systems, access controls, and monitoring. Those controls primarily protect access to data, the environments where it resides, and the paths through which it moves. They share one limitation: They do not necessarily reduce the value of the data itself once an authorized or unauthorized party obtains a readable copy.

This paper examines vaultless, keyless pseudonymization as an additional data-layer control. It is not a replacement for encryption, zero trust, access control, monitoring, or post-quantum migration. It addresses a different problem: How do we protect sensitive data once it must be accessed, processed, or shared?

The basic idea of pseudonymization is simple. Sensitive fields such as names, Social Security numbers, payment card numbers, dates of birth, account numbers, or patient identifiers are changed before they move downstream. The changed values keep the format that existing systems expect, so databases, billing systems, analytics tools, partner systems, and AI workflows can keep working. But an unauthorized user, downstream partner, or attacker sees substitute values rather than the original sensitive data.

Substitution, tokenization, and pseudonymization have existed for years. "Vaultless" means the tested workflow did not depend on a central database that stores original values alongside their substitutes. "Keyless" means restoration did not depend on one conventional stored encryption key that unlocks the protected values. A single vault and a uniform key are attractive targets for hackers, but a vaultless, keyless architecture avoids this vulnerability. The engineering challenge associated with a vaultless and keyless approach is the need to replace values in real time, preserve the formats that downstream systems need, and control who can restore the originals without concentrating that authority in one place.

FDD's Transformative Cyber Innovation Lab (TCIL) assessed one implementation of vaultless, keyless pseudonymization, developed by Rixon Technology. The assessment was independent and limited to a defined set of technical claims and test conditions. TCIL used synthetic records, operated the platform directly with no vendor personnel present for hands-on testing, and tested practical claims: whether substitute values kept their format, whether authorized users could restore originals, whether unauthorized restoration failed and was logged, whether access could be revoked, whether policy rules governing access (such as role and geography requirements) were enforced, and whether intercepted substitutes revealed original values through the methods TCIL applied.

Within the tested configuration, the platform performed as represented. TCIL also reviewed findings about the Rixon platform from a third-party qualified security assessor who examined a deployment of the technology involving payment-card data. In that assessment, the partner firm received only pseudonymized data and never received the original cardholder information. The assessor determined that specified downstream systems were outside Payment Card Industry Data Security Standard (PCI DSS) scope for that deployment, reducing the partner's assessment burden. This determination was deployment-specific; another implementation would require its own assessment.[1]

These results show that the tested systems could replace sensitive values while allowing approved users to restore them under authorized and approved conditions. Before an agency relies on this approach for sensitive use, independent experts should examine how the values are transformed, inspect what the provider's systems retain, and test the approach in limited real-world pilots. Without this additional testing and analysis, solutions like vaultless, keyless pseudonymization will remain an underutilized answer to security and privacy challenges companies are struggling with today.

**1** Third-party qualified security assessor assessment of PCI DSS 4.0.1 cardholder-data-environment scope for the assessed Rixon deployment, on file with TCIL. The determination is deployment-specific and should not be generalized.

# Beyond Encryption: Reducing the Value of Exposed Data

## Reducing Risk by Replacing Original Data Values With Substitutes

Traditional cybersecurity asks organizations to keep attackers out, encrypt data when it moves, encrypt data when it is stored, and control who is allowed to see it. It is no longer enough by itself because modern data moves so much and resides in so many different places that breaches become much harder to prevent.

Sensitive data now moves through cloud systems, payment processors, healthcare systems, analytics tools, contractors, artificial intelligence workflows, backups, and international partners. Each movement creates another place where data may be copied, logged, indexed, cached, backed up, or exposed. Every outside system that receives raw sensitive data becomes another place where the originating organization can suffer harm if that data is exposed. If partners could work on tokens instead of raw data, it would reduce exposure risk if partners were breached. If database access alone did not reveal sensitive information, the risk of insider threats or stolen credentials would diminish.

Ransomware and cybercrime groups, for example, increasingly steal readable data and threaten to disclose the data to pressure organizations to pay.[2] Public exposure of names, medical records, payment data, or account identifiers may create regulatory, legal, reputational, and personal harm. If stolen downstream records contained substitutes rather than original identifiers, attackers would have less publication or resale leverage.

Not only can data be stolen from its original owner, but it may also be exposed when other parties decrypt or otherwise reveal the data in the course of necessary usage. Employees need to process claims. Clinicians need patient records. Banks need to detect fraud. Analysts need to link records. The organization that collected the data may remain legally accountable for retaining control over the data even after operational control has spread across many systems and parties.

Privacy laws create an additional data-control problem. The European Union's General Data Protection Regulation (GDPR) provides a right to erasure, often referred to as the "right to be forgotten," that can require organizations to delete personal data when specified conditions apply.[3] That becomes harder when sensitive values have been copied across applications, partners, logs, and analytics systems. AI systems also increase the number of places sensitive information can travel. Data may enter prompts, logs, retrieval systems, embeddings, training sets, and automated workflows. Once it enters those systems, it can be difficult to trace or remove. If, however, organizations had ways to revoke access to data, they might be able to reduce downstream identifiability without requiring every copy to be found and deleted.

Quantum computing creates a final challenge. Today, nation-state adversaries can collect encrypted data and store it for possible future decryption — what is known as "hack now, decrypt later" (HNDL). Post-quantum migration remains essential, but migration takes time.[4] If captured traffic or stored datasets contain tokens rather than original source values, future decryption may reveal less useful data.

The policy question is therefore not whether encryption still matters. It does. The question is how to continue to safeguard sensitive information after encryption and perimeter controls no longer fully protect it. Reducing the number

**2** U.S. Federal Bureau of Investigation, Internet Crime Complaint Center, "Federal Bureau of Investigation Internet Crime Report 2023," 2023. (https://www.ic3.gov/annualreport/reports/2023_ic3report.pdf); C. David Hylender, Philippe Langlois, Alex Pinto, and Suzanne Widup, "2024 Data Breach Investigations Report," *Verizon*, 2024. (https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf)

**3** Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), *Official Journal of the European Union*. (https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng); European Data Protection Board, "Guidelines 01/2025 on Pseudonymisation," January 2025. (https://www.edpb.europa.eu/public-consultations/guidelines-012025-on-pseudonymisation_en); European Data Protection Board, "2025 Coordinated Enforcement Action: Implementation of the Right to erasure by controllers," February 10, 2026. (https://www.edpb.europa.eu/system/files/documents/2026-02/edpb_cef-report_2025_right-to-erasure_en.pdf)

**4** Cybersecurity and Infrastructure Security Agency, National Security Agency, and National Institute of Standards and Technology, "Quantum-Readiness: Migration to Post-Quantum Cryptography," August 17, 2023. (https://www.cisa.gov/sites/default/files/2023-08/Quantum%20Readiness_Final_CLEAR_508c%20%283%29.pdf); U.S. Department of Commerce, National Institute of Standards and Technology, FIPS 203, "Module-Lattice-Based Key-Encapsulation Mechanism Standard," August 13, 2024. (https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.203.pdf); U.S. Department of Commerce, National Institute of Standards and Technology, FIPS 204, "Module-Lattice-Based Digital Signature Standard," August 13, 2024. (https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.204.pdf); U.S. Department of Commerce, National Institute of Standards and Technology, FIPS 205, "Stateless Hash-Based Digital Signature Standard," August 13, 2024. (https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.205.pdf)

of systems that receive the original source value can make it easier to limit exposure and, where appropriate, render distributed representations unusable without having to locate and remove every readable copy.

## What Vaultless, Keyless Pseudonymization Does

Pseudonymization means replacing or processing identifying information so the data can no longer be attributed to a specific person without separately protected additional information. It is not the same as anonymization. Anonymization is intended to be permanent. Pseudonymization is reversible when the authorized party has the required reconstruction information and permission to use it.

Vaultless, keyless pseudonymization is a specific architectural approach to that broader idea. In the workflow TCIL assessed, sensitive values were replaced with format-preserving substitutes without a central vault of originals and without a single conventional stored encryption key observed in the tested transformation pathway.

Format preservation matters because many systems will reject data that has the wrong shape. In this paper, these substitutes are also called "tokens." A token may preserve the original field's format, but format preservation is a property of the particular tokenization approach, not of every token. A ZIP code must look like a ZIP code. A payment card number must have the right number of digits. A phone number must look like a phone number. A patient identifier may have to match a field length or naming pattern. If the substitute breaks those rules, the receiving system breaks too. Realistic substitutes may also make it less immediately apparent to an unauthorized user that particular fields have been transformed, although TCIL did not separately test this as a security property. The choice depends on the customer's operational requirements and risk model.

The following table illustrates the use of synthetic values. The point is not that pseudonymized values must always look realistic. Rather, the fields and substitute values can be customized for each customer based on operational, regulatory, and system requirements, including the need to preserve specific formats, validation rules, field structures, or other characteristics required by downstream systems.

*Table 1: An Illustrative Patient Record*

| Field | Original | Pseudonymized | Why it matters |
|---|---|---|---|
| Patient Name | Jane Smith | Kara Orent | Changed. Consistent across systems and usable as a record-linking value. |
| Date of Birth | 04/12/1978 | 09/27/1953 | Changed. Valid date format preserved. |
| Social Security Number | 523-44-7821 | 871-06-3349 | Changed. NNN-NN-NNNN format preserved. |
| Phone Number | (703) 555-0182 | (415) 555-0647 | Changed. Phone-number structure preserved. |
| Home Address | 142 Oak Lane, Arlington, VA 22201 | 389 Pine Court, Salem, PA 27103 | Changed. Address-like structure preserved. |
| Insurance Member ID | BCB-4471839-02 | BCB-8823051-07 | Changed. Alphanumeric format preserved. |
| Diagnosis Code | E11.9 | E11.9 | Unchanged where operationally required and appropriate. |
| Visit Date | 03/14/2025 | 03/14/2025 | Unchanged where operationally required and appropriate. |

# Beyond Encryption: Reducing the Value of Exposed Data

## How Rixon Enables Vaultless and Keyless Reconstruction of Original Data Values

Rixon has given the distributed required variables, or DRVs, to the inputs its platform uses to reconstruct original data values from substitutes. DRVs are security-critical and key-like in the everyday sense because they are needed to restore the original value. But they are not conventional encryption keys.

A cryptographic algorithm uses a conventional encryption key to encrypt plaintext into ciphertext or decrypt ciphertext back into plaintext. In the Rixon workflow TCIL assessed, DRVs functioned as separate reconstruction inputs held by different parties. They had to be combined under customer-defined policy conditions before reversal could occur. The difference is that a conventional key supports mathematical decryption, while DRVs participate in multiparty, policy-controlled reconstruction.

Restoration can be limited by role, purpose, time, location, department, service provider, consent status, or other conditions the organization defines. That means the system is not merely asking whether someone has a secret. It is asking whether this requester is allowed to restore this value under these conditions.

## What TCIL Tested

TCIL's assessment was not a formal cryptographic proof, a production certification, a legal compliance determination, or a contest between pseudonymization and encryption. TCIL tested a narrower and more practical question: after sensitive values are pseudonymized, can an unauthorized person who obtains the downstream data recover the originals or bypass the rules that govern restoration?

TCIL evaluated this question because federal agencies, critical infrastructure operators, defense contractors, healthcare organizations, and financial institutions increasingly need to share and process sensitive data across distributed environments. Every downstream copy expands breach, insider, ransomware, AI, compliance, and HNDL risk. The pilot asked whether one implementation could preserve operational use while reducing downstream exposure of source values.

### *How the assessment worked*

TCIL used 500 synthetic records containing names, Social Security numbers, dates of birth, payment card numbers, medical record identifiers, addresses, and phone numbers. TCIL operated the platform directly without vendor personnel present for hands-on testing. TCIL defined the test cases and acceptance criteria, then tested the behaviors described below.

The tests asked practical questions. Could protected fields keep the format downstream systems as expected? Could an authorized request restore the correct original value? Would an unauthorized request fail and leave a record? Could access be revoked after data had already been shared? Would role-based restrictions and geofencing work as configured? Would restoration attempts be logged? In a simulated interception scenario, could TCIL infer or recover original source values from the exposed tokens using the methods applied?

### *What TCIL found*

Within the tested configuration, protected fields kept their required formats. Fifty authorized restoration attempts returned the correct original values. Fifty unauthorized restoration attempts were denied and logged. Access revocation took effect immediately in the tested environment. Role restrictions, geofencing, and audit logging operated as configured.

In the interception simulation, TCIL did not recover source values from intercepted tokens through visual inspection, format inference, or statistical frequency analysis. In an insider-style test, database-level access without the required reconstruction authority did not produce original values, and the denied attempts were logged.

# Beyond Encryption: Reducing the Value of Exposed Data

### *What TCIL did not prove*

TCIL did not prove that vaultless, keyless pseudonymization is a replacement for other security controls. Encryption, access control, zero trust, monitoring, and post-quantum migration remain necessary. Vaultless, keyless pseudonymization is an added data-layer control designed to reduce the value of exposed sensitive fields when those other controls are bypassed or when data must move through downstream systems.

TCIL's objective was not to determine whether the system could ever be penetrated. Systems can be compromised in many ways, including insider threats, poor cyber hygiene, zero-day vulnerabilities, implementation flaws, or attacks against supporting infrastructure. Instead, the assessment focused on what an adversary would actually obtain after a successful compromise: whether the exposed data would contain usable sensitive source values or protected representations that could not be readily reconstructed without the required authorization and reconstruction components.

Separate experts still need to examine whether the transformation can be reversed through mathematical analysis, whether information can leak through a system's behavior, whether the implementation has weaknesses, and what happens if someone compromises a party that controls restoration. TCIL's tests showed that the system performed the operational functions assessed under the stated conditions; they did not resolve those broader security questions.

## What the Findings Mean

The findings matter because they address concrete operational problems. The table below states the practical significance of each result in plain language.

| Finding | Why it matters |
|---|---|
| Format preservation worked | Existing systems may continue to process records without redesigning every database, form, application programming interface, or workflow. |
| Authorized restoration worked | Legitimate users can still obtain the original value when policy permits it. |
| Unauthorized restoration failed and was logged | Unauthorized attempts to restore original identities can be blocked and recorded, so downstream users and administrators do not automatically gain access to them. |
| Revocation worked in the tested environment | An organization can withdraw a party's ability to re-identify records already shared downstream. |
| Role and location controls worked | Restoration can be tied to policy conditions such as job function, geography, purpose, or time. |
| Intercepted tokens did not reveal originals through TCIL's applied methods | A stolen or intercepted dataset may have reduced identity value when reconstruction authority is unavailable. |
| Qualified security assessor scope finding | In the specific deployment reviewed, the assessor determined that a downstream partner receiving only pseudonymized payment data did not have to treat its systems as part of the environment handling original cardholder data. Another deployment would require its own assessment. |

## Policy and Procurement Implications

It is reasonable to ask why vaultless, keyless pseudonymization is not already standard if the idea is useful. The answer is that replacing one value with another is not the hard part. The hard part is making the model work in live operations without recreating the weakness it is supposed to remove.

## Beyond Encryption: Reducing the Value of Exposed Data

A workable system must generate substitutes in real time, preserve the format existing software expects, avoid a central vault that becomes a target, avoid a single conventional stored key that unlocks everything, divide reconstruction authority across separate parties, enforce policy before restoration, log what happens, support revocation, and do all of this without forcing authorized users to change how they work.

Miss any of those pieces and the system either breaks operations or recreates a single point of failure. That is why TCIL did not treat the vendor's claim as self-proving. TCIL tested the operational behaviors that could be tested in a bounded environment and recommends the deeper review needed for high-assurance use.

At the moment, however, federal agencies and private organizations are apt to ignore this type of technology simply because it does not fit neatly into existing categories for encryption, tokenization, de-identification, or access control.[5] The answer is not to create a new category for every emerging technology but to establish a consistent way to evaluate what the technology does, what risks it reduces, what new dependencies it introduces, and what evidence is required before it can be trusted in practice. The recommendations that follow are intended to provide that structure.

First, the National Institute of Standards and Technology (NIST) should commission an independent cryptanalytic review of the transformation architecture, and the National Security Agency (NSA) should conduct a parallel classified review for national security use cases.

Second, provider-side technical inspection should determine whether original values, token-to-source mappings, reconstruction material, or other sensitive artifacts appear in memory, logs, retry queues, telemetry, backups, or other locations not visible through user-facing tests.

Third, NIST should give organizations a clear way to evaluate whether these systems protect data as claimed. It should first determine whether existing guidance adequately addresses architectures that restore pseudonymized values without a central vault or conventional format-preserving encryption. Where guidance is missing, NIST should define the minimum-security properties, evidence, controls over restoration, audit records, and tests needed to support a vendor's claims. The National Cybersecurity Center of Excellence could then develop example architectures and use cases showing how organizations can deploy and assess these controls.[6]

Fourth, agencies should conduct bounded pilots in low-to-moderate-risk environments before considering high-assurance use. Those pilots should evaluate not merely whether transformation and reconstruction function correctly but also whether sensitive source values remain unavailable following compromise of the systems through which protected data travels.

Finally, federal procurement language should require vendors making data-layer protection, tokenization, pseudonymization, or quantum-resistance claims to disclose what their architecture does and does not do. Agencies should ask whether raw data enters the vendor's infrastructure, whether any token-to-source mapping exists, how reconstruction authority is stored and distributed, what standards or independent validation the product satisfies, what audit evidence is produced, and what limits apply to the vendor's security claims.

---

**5** Barbara Guttman, Simson Garfinkel, Joseph Near, Aref N. Dajani, and Phyllis Singer, *National Institute of Standards and Technology*, "Special Report 800-188 De-Identifying Government Datasets: Techniques and Governance," September 2023. (https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-188.pdf); "Payment Card Industry Data Security Standard Version 4.0.1," *Payment Card Industry Security Standards Council*, June 2024. (https://docs-prv.pcisecuritystandards.org/PCI%20DSS/Standard/PCI-DSS-v4_0_1.pdf); U.S. Department of Health and Human Services, "Guidance on De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule," November 26, 2012. (https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf)

**6** Morris Dworkin and Nicky Mouha, "Special Paper 800-38G Rev. 1 (Second Public Draft) Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption," *National Institute of Standards and Technology,* February 2025. (https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-38Gr1.2pd.pdf); TCIL identified no NIST or NSA standard specifically governing the vaultless, non-Format Preserving Encryption (non-FPE), multiparty architecture assessed here.

# Beyond Encryption: Reducing the Value of Exposed Data

## Conclusion

Federal cybersecurity policy has a data-layer gap. Encryption, access control, monitoring, zero trust, and post-quantum migration remain essential, but they do not fully control what happens to the value of sensitive data once it is decrypted, copied, shared, exfiltrated, or ingested into AI workflows. The implementation TCIL assessed offers a practical example of a different model: transform sensitive fields before they spread, preserve enough format for operations to continue, and make reconstruction a governed event rather than a default downstream capability.

Vaultless, keyless pseudonymization is ready for independent review, bounded pilots, clear standards, and procurement language that forces vendors to prove data-layer protection claims rather than merely assert them.

# Beyond Encryption: Reducing the Value of Exposed Data



## About the Authors

Dr. Georgianna "George" Shea is chief technologist at the Center on Cyber and Technology Innovation and its Transformative Cyber Innovation Lab at FDD. With more than 30 years across federal and commercial roles, including deep roots in the intelligence community (IC) through HUMINT support to the Defense Intelligence Agency, her work has informed congressional briefings, the Department of Defense, and the 2024 PCAST Report to the President on Cyber-Physical Resilience. She serves on multiple private sector boards, holds a Doctor of Computer Science in Information Assurance, a CISSP, and is a U.S. Army veteran.



Cason Smith is a student at the University of South Carolina pursuing a bachelor's degree in information technology with a minor in data science. He is passionate about national security and the resilience of critical infrastructure, with a particular focus on how emerging technologies shape both opportunities and risks. Cason worked as an intern at FDD under Dr. Shea's direction. His cybersecurity research focuses on the intersection of emerging technology, policy, and national security.



Sydney Johnson is a senior at Embry-Riddle Aeronautical University pursuing a B.S. in mechanical engineering with an emphasis in propulsion. She has a significant background in information technology. She is a recipient of the Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship, leads her capstone research group on auxiliary power units for Honeywell, and serves as a teaching assistant for Embry-Riddle's cornerstone engineering courses.

***About the Foundation for Defense of Democracies***

*FDD is a Washington, DC-based, nonpartisan 501(c)(3) research institute focusing on national security and foreign policy.*

***About FDD's Transformative Cyber Innovation Lab***

*TCIL finds and nurtures technologically feasible, testable pilot projects which begin to solve some of the hardest cyber problems afflicting the national security industrial base and the United States. TCIL's mission is to help shorten the lag between idea and piloting and between piloting and the adoption of potential solutions to the thorniest of cyber problems.*

*For more information, visit:* https://www.fdd.org/projects/transformative-cyber-innovation-lab

---

*FDD values diversity of opinion and the independent views of its scholars, fellows, and board members. The views of the authors do not necessarily reflect the views of FDD, its staff, or its advisors.*