

## TECH BRIEF

# Exact Data Matching

## Accurate Data Classification Using Minimal Compute

### Exact Data Matching Defined

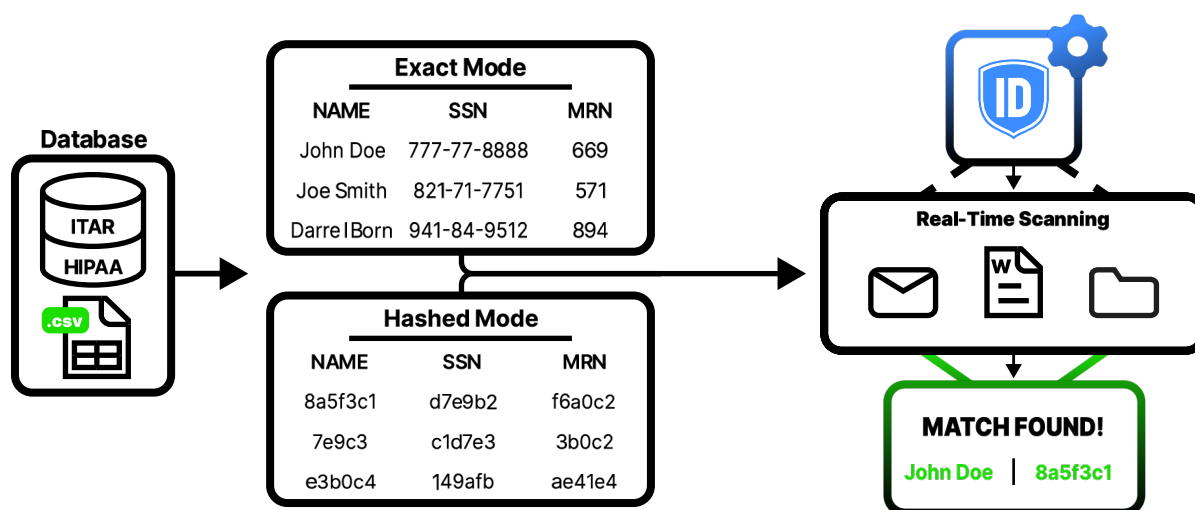
Exact Data Matching (EDM) is a sophisticated data protection technique that uses unique identifiers, via hash values, for specific pieces of sensitive information. These identifiers are used to detect and protect that information across various systems and processes within an organization. Unlike traditional pattern-based matching, which is prone to false positives, EDM focuses on exact matches, ensuring only designated data elements are flagged or acted upon, resulting in increased accuracy, achieving 100% precision.

The precision of EDM makes it an invaluable tool for security architects, IT directors and CISOs seeking to implement robust, efficient, and accurate data protection measures.

## How It Works

Two modes of operation are offered:

- exact mode, where the exact strings are searched for
- hashed mode, SHA256 hashes of the data to be found.



### 100% Classification Accuracy

Exact Data Matching (EDM) has emerged as a powerful tool in data protection strategies, offering 100% precision in identifying important data assets. By utilizing hash values instead of actual patient data, EDM substantially reduces the risk of medical data breaches. The precision of EDM in fraud detection translates directly into improved customer confidence in financial security. By creating unique hash values for ITAR-controlled information, EDM enables government agencies to accurately track and control the distribution of sensitive defense-related data.

## Key Features

### Improved Data Match Accuracy

Exact Data Matching delivers superior accuracy in data classification while handling large datasets of over 2 billion records. Unlike traditional pattern-matching approaches that often generate false positives and require constant manual verification, this matching system uses hashes to accurately identify and classify sensitive information, proprietary data, and compliance-related content. Accuracy levels can be fine-tuned to meet risk appetites by using combinations of variables; proximity, secondary and secondary count. Organizations can confidently automate their data classification workflows, eliminating resource waste associated with false positives and the security risks of false negatives, while reducing the operational overhead required for manual verification and error correction.



### Lower Operating Cost

Organizations achieve meaningful reductions in infrastructure costs by leveraging our solution that is 160 times more memory efficient and processing speed which is 28 times faster than competing solutions. These optimizations translate to lower cloud computing costs, an 88% yearly cost savings using general purpose AWS instances over the competition.



### EDM Worth the Investment

Exact Data Matching in a data security program can lead to significant financial savings for organizations across various industries. These savings stem from multiple areas, including reduced risk of data breaches, improved operational efficiency, and enhanced compliance management.

### Optimized Compute Resources

Exact Data Match demonstrates robust technical efficiency, processing data classification tasks with just 0.2GB of memory compared to competitors' 32GB requirement for scanning 50 million unique values—a 160-fold improvement. Optimizations in resource utilization and processing speed enable organizations to scale data classification operations without the need for expensive hardware upgrades or cloud computing resources.

### Handling Large Data sets

Indexes are used to optimize speed and memory usage. Exact Data Match handles a large single index of 2 billion records with the capability to extend by daisy chaining additional indices. This enables organizations to deploy data classification capabilities, efficiently process large-scale datasets without infrastructure bottlenecks, and achieve faster time-to-insight across their data ecosystem.

### Robust Data Privacy

Hashed mode delivers powerful data privacy protection by transforming sensitive data values into secure, irreversible hashes before comparison. This ensures the actual sensitive values like Social Security numbers, account details, or patient identifiers never leave your secure environment during the matching process. Unique cryptographic fingerprints of your sensitive data are created, enabling precise detection of data matches while maintaining zero visibility of the original values. The approach provides both ironclad security and highly accurate matching capabilities, effectively balancing robust data protection with reliable sensitive data discovery.

#### ABOUT INSPECT-DATA

Inspect-Data empowers your business with robust, real-time data classification, intelligence and analytics that seamlessly integrates with your technology stack.

#### VISIT INSPECT DATA

[inspect-data.com](https://inspect-data.com)

#### CONTACT

[contactus@inspect-data.com](mailto:contactus@inspect-data.com)