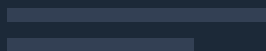
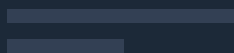
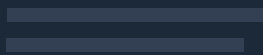
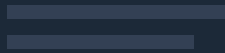
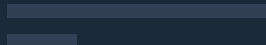


# Five Ways to Reduce Data Exposure



- 1 ☒ 
- 2 ☒ 
- 3 ☒ 
- 4 ☒ 
- 5 ☒ 

# Introduction

In the decade since Target's 70-million-customer data breach, the security community has continued to witness larger and more devastating breaches that have harmed businesses and individuals, resulting in loss of reputation, intellectual property, and privacy. Regulations that demand better protections for data continue to follow, and new laws in California, Virginia, Colorado, Connecticut, and Utah will take effect in 2023.

Cyber insurance questionnaires now ask policy applicants to estimate the amount of regulated data they store. Smart security executives expect those questionnaires to get more exacting over time and don't want to be caught flat-footed when insurers ask how those estimates were created and whether there were any legitimate efforts to quantify the amount of regulated data being processed and stored.

As data protection teams start to identify and quantify the sensitive data they store, they usually find more than they expect — and often in unexpected places. So how can you best protect sensitive and regulated data? And how can you extend this protection across all use cases, including internal and external third-party handling?

# Five ways to reduce data exposure

When you find sensitive or regulated data that's unprotected, there are several options available to remediate the potential exposure:

- Delete the data.
- Move it somewhere that's locked down correctly.
- Lock it down in place.
- Lock it down in flight by blocking and/or encrypting the data.
- Obfuscate data.

Each of these options presents technical and organizational challenges; what may be easier from a technical standpoint could be more challenging organizationally (and vice versa).

## Deleting data

Deleting data might be easy to automate, especially for structured data, but arriving at a retention policy can be challenging. Many organizations hand this off to a records and information management (RIM) team, but such a group is not always in place. Stakeholders from lines of business, legal and compliance, and IT usually need to get involved in the process, and coming to an agreement on what should be deleted takes time.

Deletion is irreversible, so it's understandably a tough decision to make. Once you do arrive at a retention policy, though, it's important to adhere to it and demonstrate that you're doing so. For automation to be successful, it must be able to identify the right data, delete it securely, and report on progress. This is true for local handling as well as contractually dictated data handling by third parties.



## Moving data

From a technical perspective, moving or quarantining data can be more challenging than deleting it, but it's usually an easier decision for organizations to make. The stakes aren't quite as high as deletion because you can always move the data back. Organizations sometimes want to leave a stub with instructions and/or a link to the data's new home, and companies may need to do some additional leg work if data is moving to a different kind of data store that requires different client software or configuration to access it, or requires different access controls to configure. It is worth mentioning that in some cases, the volume of data might preclude moving or quarantining data. Internet service providers, for example, might have trouble quarantining network flow data, but that is certainly a special case.

## Locking data down in place

Locking data down in place requires sophisticated automation, but from an organizational perspective, it's the easiest to accept. Access controls and encryption are common technical solutions to this problem and excellent commercial options exist to support enterprise security teams hoping to lock their data in place. Compliance and regulatory obligations will guide the type and strength of these functional security protections.

Stakeholders agree that sensitive data shouldn't be broadly exposed and that employees shouldn't have access to data they don't need, especially if that data is sensitive or regulated.

Collaboration settings, rules, and requirements change too frequently and are too complex for humans to review manually, so this option requires automation that correctly analyzes and optimizes access controls. With the right automation tools, this option is the least intrusive, as it does not require deletion or change to data content or location, nor does it require end users to learn or change their behavior.

## Locking data down in flight; blocking and/or encrypting

Many organizations look to restrict the flow of sensitive data, keeping it inside a logical “perimeter,” using data loss prevention (DLP) solutions that block data “in-flight,” and/or encrypting data. These technologies started with endpoint- and perimeter-enforcement mechanisms, but have evolved to leverage label-based enforcement, in which a sensitivity label is applied to files that need special handling. DLP systems will then restrict which files can be printed, copied to a USB, sent via email, and/or encrypt files based on their labels, so they can only be used by authorized systems and accounts.

DLP requires both technical sophistication and significant business involvement; not only is DLP labeling and blocking technology required, but organizations find that automation is required to apply the correct labels, even after training employees. Blocking or encrypting data can add friction to workflows, and labeling technologies are limited to specific file types.

From a business perspective, a clear labeling taxonomy and handling rules need to be created, and then end users need to be trained on how the process works and what to expect. If not implemented carefully, employees can become frustrated as workflows change and become disrupted. For an example of labeling taxonomy, please see [Appendix 1](#).

## Obfuscation

Some organizations consider obfuscating or tokenizing sensitive or regulated components within databases or documents. With an obfuscation strategy, most of the data is unchanged; the sensitive and regulated parts are hidden and available only to authorized users. This option requires both organizational consideration and sophisticated automation. Though most of the document remains readable and available for existing workflows, organizations need to decide what parts should be obfuscated, and who should be able to view them, when, and how. Automation will be needed to identify data that should be obfuscated, obfuscate it, and reveal it to the right people securely when appropriate. De-obfuscation usually requires an additional client on an endpoint.

It is worth mentioning that the strength of obfuscation can and will vary between organizational settings. In some less-sensitive cases, a weak obfuscation is fine, perhaps to hide some sensitive data from an internal work group. In other cases, however, much stronger means are demanded, which leads many teams to use encryption solutions, such as homomorphic algorithms. These are used to provide strong cryptography without stopping users from analyzing the data toward whatever mission is being worked.

# Which options are right for you?

For most organizations, the best options are the ones they can accomplish technically and get their organization on board with. And, of course, the compliance and regulatory pressures that might exist will have a huge impact on the data security strategy an organization selects. In some cases, the choice might be dictated by an external entity. Nevertheless, we will assume that the enterprise security and IT operations teams have some selection leverage in our discussion below.

Furthermore, for larger third-party data protection, contractual and service level agreement (SLA) constraints will be in place, hopefully consistent with the options included in this paper. A tough decision for most enterprise and IT teams is how to handle smaller third parties who might not have the ability to implement strong controls. This remains a gap in many protection strategies.

Nevertheless, as a first step, we recommend that the security and IT teams first assess how well they know your data. How accurate is the inventory of sensitive and regulated information? Who creates it, has access to it, and uses it? If these questions are hard to answer, then the team should consider conducting a **risk assessment** with an outside vendor or party that can help answer these questions for a meaningful portion of the data.

It is important to note that small- and medium-sized businesses will have the advantage of being able to refer to data in a simple, monolithic manner. But as an organization grows in size, this is a more complex task. Large companies, for example, will have hundreds or even thousands of different situations involving data, so they might have to divide and conquer to address sensitive and non-sensitive data in internal and external use cases, for a variety of business unit functions.

Regardless of the case being examined, a risk assessment will help quantify the relevant technical capabilities. How well can the teams handling data identify regulated data, and data that's exposed, stale, or in the wrong places? Can the team automate changes to access controls or labels, or encrypt, move, or delete data? Additionally, some assessments (such as the one offered by Varonis) can provide a critical findings report customized to your organization's needs, regulations, and configurations.

Next, it is recommended to take an inventory of the policies your organization has already created. Does your organization have a retention policy, an access control policy, and/or a data taxonomy and handling policy? Even if there are no formal policies, you may be able to identify “no brainer” policies to enforce, such as, “Regulated data shouldn’t be publicly accessible, or accessible to the entire organization.”

Consider creating a matrix of what your organization can both agree to do and what you can automate.

Data protection options	Org agreement?	Technical capability?
Delete	✓ X	✓ X
Quarantine	✓ X	✓ X
Lock data down in place	✓ X	✓ X
Lock data down in flight	✓ X	✓ X
Obfuscate	✓ X	✓ X

As data breaches continue and regulations demand more controls, organizations will need to take more of the above actions. To decide which methods you’re able to accomplish, you’ll need to know your data, your technology, and your organization’s appetite for tough decisions. Automation that solves problems completely will only become more crucial as data continues to grow, and manual remediation efforts continue to be ineffective and unsustainable.



## Detective controls

Few, if any, of the preventive options mentioned above can be accomplished without understanding data usage. You wouldn't want to delete sensitive data that's actively in use, and you wouldn't want to cut off access to data from those who need it to do their jobs. Again, this task is much more difficult as organizations increase in size, scope, and complexity. Larger organizations, for example, might include a RIM team – and this at least helps with certain tasks such as storage and/or deletion of business records.

If you decide to collect and aggregate data usage, it makes sense to use this information to increase your detective capabilities as well. Just as a credit card company uses a record of transactions to build a profile of normal charge behavior, a record of data transactions can help build a profile or determine normal data behavior. With the right automation, you can detect unwanted changes to preventive controls, as well as insider threats, ransomware, and more advanced cyber adversaries.

## Appendix 1

# Labeling taxonomy and enforcement

Organizations dream of using labels for many data-handling tasks, including data protection, retention, and geographic flows. Unfortunately, the more labels an organization has, the more confusing it can become for users. This is especially intense for larger organizations.

In addition to keeping the number of labels reasonable (five or fewer is common at the top level), it helps to have clear, simple names and descriptions for each label. If you're new to labeling taxonomy, it may make sense to start with a basic structure, and DLP controls on restricted documents.

As a first step, some organizations choose to implement information-only labels that do not introduce any enforcement until users become more proficient with labeling workflows.

### Public

Approved for public consumption

#### Examples

- Press releases
- Marketing materials

### Internal

Not intended for public consumption, but can be shared externally based on business need

#### Examples

- Internal directories
- Org charts
- Internal communication

### Confidential

Sensitive data that should not be shared externally

#### Examples

- Contracts and account data
- Security reports
- Forecasts
- Employee reviews

### Restricted

Very sensitive business data that could cause harm if accessed by unauthorized parties

#### Examples

- PII or PHI
- Employee or customer information
- Passwords
- Source code and intellectual property
- Pre-announced financial reports

#### DLP protection

- Encryption
- Endpoint controls
- Perimeter controls



# Ready to experience the Varonis difference?

Reduce your risk without taking any. Contact our team to learn what will be covered in your **free** data risk assessment.

[Contact us](#)

---

## About Varonis

Varonis is a pioneer in data security and analytics, fighting a different battle than conventional cybersecurity companies. Varonis focuses on protecting enterprise data: sensitive files and emails; confidential customer, patient, and employee data; financial records; strategic and product plans; and other intellectual property.

The Varonis Data Security Platform detects cyber threats from both internal and external actors by analyzing data, account activity, and user behavior; prevents and limits disaster by locking down sensitive and stale data; and efficiently sustains a secure state with automation.

Varonis products address additional important use cases including data protection, data governance, Zero Trust, compliance, data privacy, classification, and threat detection and response. Varonis started operations in 2005 and has customers spanning leading firms in the financial services, public, health-care, industrial, insurance, technology, consumer and retail, energy and utilities, construction and engineering, and education sectors.