TERADATA AND DATAGUISE ALLIANCE

Privacy Data Compliance Starts with Knowing Your Sensitive Data

Security and risk pros can't expect to adequately protect customer, employee, and sensitive corporate data and IP if they don't know what data exists, where it resides, how valuable it is to the firm, and who can use it.

– Forrester Research, *Rethinking Data Discovery and Classification Strategies*

GDPR Is Really Important… But So Are Hundreds of Other Regulations

While the European Union's General Data Protection Regulation (GDPR) has received a tremendous amount of attention over the past year as it set the tone for data protection around the world for years to come, numerous other regulatory measures at the national and regional level are equally as important—like the California Consumer Privacy Act (CCPA), New York Cyber Law, China Cyber Security Law, Payment Card Industry Data Security Standard (PCI), the Health Insurance Portability and Accountability Act (HIPAA), and many more. Each of these regulations place restrictions on how personal data is handled and collected.

In the United States, while hundreds of state and federal laws about digital commerce or communication may apply to an enterprise, no one overarching, cohesive law like GDPR exists. It is the most comprehensive to date, impacting organizations around the world who market to, sell to, or do business with prospects and customers in the EU. Any organization who stores EU subject data must comply with GDPR (including companies in all verticals in the U.S. and countries outside the EU). The same is true for emerging privacy data regulations being enacted across the country.

While new regulations are more stringent, the end goal remains the same as it's been for many years: protecting the privacy and confidentiality of sensitive personal or corporate information.

In the case of GDPR, the objective of the May 2018 effective legislation is to strengthen the security and privacy of individuals' data across all EU Member States. It gives every EU subject the right to know and decide how their personal data is being used, stored, transferred, and deleted. GDPR also applies to all organizations that are either a data controller and/or processor of personal information.

GDPR is here to stay, and legislation like GDPR will likely be adopted by more governments outside the EU in the near future. Complying with existing regulations or being prepared to comply with them begins with knowing your data.

Do You Really Know All Personal Data in Your Enterprise?

As organizations were preparing for GDPR, many discovered they didn't really know where all the personal data lived within their repositories, databases, and archives, whether on premises or in the cloud. They weren't able to size their own exposure—not knowing what they didn't know.

Yet, legally, they had to find and identify all of it in order to implement policies, practices, and processes to comply with GDPR. Whether preparing for GDPR or other regulations, it's extremely challenging to scan all the places in your organization where personal data could be found.

To understand what's involved and choose prudently whether to undertake the journey to compliance manually or work with a service provider, first recognize that the personal data your organization holds may be voluminous, varied, vague, and vast.

Uncovering Sensitive Data in Hidden Nooks and Crannies

While sensitive data may be a small subset of your entire library, bits of it will likely be found in unexpected locations where nobody would ever think to look, and where sensitive data is not supposed to be.

Essentially, sensitive personal information may be in one or more of hundreds or thousands of data repositories in your organization, on premises or already in the cloud. It could be in archive file shares from many years past. Often, personal information you'll be legally required to protect is exposed, sitting in unstructured data repositories like Hadoop, making it very hard to analyze and detect. It could exist in comments and notes fields in databases, in saved text and chat conversations, or even in recorded phone calls with your employees. Perhaps someone in your organization made a copy of some data for a separate purpose and, once that was done, left it unattended and forgotten.

The sheer volume and spread of locations for personal information means it would be extremely time-consuming and error prone to try to search manually and find it all, making it easy to miss significant data. Not to mention, manual sensitive data discovery is neither scalable nor repeatable.

It's imperative to have a process for detection through all your file shares, databases, and repositories to verify the known sensitive data you already expect to find—and to detect sensitive data in the surprising, unknown locations you may never have considered or expected, so it can also receive proper attention and not be in danger of accidental reuse, propagation, or sharing. You must create standard and custom values to include all types of sensitive data covered by GDPR and other regulations, and policies with those attributes, to build look-up tables and search for all sensitive data and report on them for action.

Here are three Teradata-centric examples to give you a sense of all the places sensitive data can hide:

1. Teradata Database and Files
2. Teradata On-Premises and Teradata IntelliCloud™
3. Teradata On-Premises and Teradata IntelliCloud and Files Servers/Other Data Stores (like Hadoop)

Five Considerations for Automated Personal Data Discovery

Organizations recognize that the internal cost, effort, time, and risk involved in trying to detect all personal data manually is substantial and inefficient. If yours is better served by purchasing an automated detection solution, there are some capabilities the service provider must have that are vital to help establish regulatory compliance.

You'll notice many services that perform masking or encryption, but only after the sensitive data has been detected. The critical key is that the service you choose flawlessly detects all personal, sensitive data—and can do so extremely fast, in a matter of hours or days, not weeks or months.

Here are five factors for evaluating whether data detection services can find all personal information in your organization needed for GDPR compliance:

1. Credible

The vendor should have an extensive history of performing its detection service for customers concerned with regulatory compliance—especially PCI, HIPAA, and other existing data privacy mandates—with a good track record and results you can verify. Your organization should not be a test case for a new solution or service. GDPR is not a drill, and sensitive data detection is not a one-time operation, as it impacts any organization that stores EU subject data—regardless of where the organization is located. The GDPR compliance journey and other compliance efforts continue without a predictable end.

2. Comprehensive

The service must be sufficiently flexible and capable of finding all types of sensitive data in the myriad of data locations, traditional and in the cloud, that your organization uses. A comprehensive service detects and locates the personal information your organization is required to protect, everywhere, on any platform, regardless of data structure. Ideally in looking at the database search result, you should be provided confidence levels for all the temporary tables, archives, copies, and other locations that may have been forgotten.

3. Customizable

Ideally, the solution will be policy-based and offered with standard policies and pre-defined templates to detect, locate, and identify personal information at first start-up. Your organization should be able to include its unique policies to detect company-specific personal information fields, like account numbers. Customization is critical for your organization's regulatory compliance.

4. Comprehendible

The solution should be able to instantly present a clear view of all personal and sensitive data fully identified and prepared to be managed for legal compliance. Ideally, it should provide a simple C-level dashboard or graphical interface that your IT organization can drill down into to obtain more detailed information of its critical fields. From a single pane view, your organization should be able to understand what data is sensitive, where it is, whether it is protected or exposed, and who has accessed it—at the executive and operations levels.

5. Continuous

The service should alert you when new exposed personal information enters your organization. It should detect, locate, and identify it, so you can manage the data according to your compliance driven policies. Once your organization has confidence that the known, expected, hidden, and unhidden personal information is detected, located, and identified—and that you can detect new personal information as it comes in—you will be far better equipped to prepare the organization for compliance with regulations.

Teradata and Dataguise: Sensitive Data Detection and Powerful Masking Options

Teradata customers want to maximize the value of all their data, giving more users access to more data, but they also need to keep confidential information safe. Concerns about data security and regulatory compliance shouldn't keep them from leveraging the power of their data.

Dataguise helps maintain trust and compliance by delivering the industry's only solution for sensitive data detection in Teradata, as well as all other data sources or repositories (databases, Hadoop, structured, unstructured), plus powerful, flexible masking options and encryption for ensuring data privacy.

Dataguise is a leader in sensitive data governance, providing data-centric audit and protection (DCAP) solutions that discover sensitive data and secure it far faster than any other global sensitive data governance solution—and appearing as a key player in Gartner's DCAP Marketplace. Even the largest data sets can be analyzed in minutes or hours, compared to weeks and months with other solutions.

DgSecure by Dataguise precisely detects, protects, and audits sensitive data across the enterprise, on-premises, and in the cloud. Once sensitive data is found, DgSecure continues to protect and monitor it in real-time. Delivering a single dashboard view of sensitive data security, policies, access, and trends, DgSecure gives IT and business leaders the insights they need to maximize the value of information assets while managing risk and meeting PCI, HIPAA, GDPR, and other data privacy mandates.

DataGuise leverages Attribute Based Access Control (ABAC), the newest, most granular technology for achieving ultra-fine grain access control on multiple data structures. This functionality includes access to heterogenous and/or unstructured data sets queried and assembled for purpose-built use cases like GDPR data consent, right to forget, and other privacy by design processes, or highly focused use cases. Unlimited enterprise use-cases can be created and deployed for auditing, isolating, or restricting access to specific datasets.

Teradata organizes data and delivers insights on a massive scale. DgSecure ensures the delivery is secure. Typically a trade-off occurs between data security and compliance and data distribution, but DgSecure eliminates this gap. DgSecure helps safely and seamlessly unlock the benefits of Teradata by detecting, protecting, and auditing sensitive data assets in real time wherever they might be hiding—saving time, resources, and compliance costs. Management templates help implement privacy by design business processes and DgSecure helps make data auditable.

Teradata Security Services

Teradata Center for Enterprise Security has the consulting experts to help you meet any data compliance regulations using an extensive set of internally vetted and industry best practices. Teradata security experts can get DgSecure online and ready to find and secure your data in hours so you can leverage the tremendous power of Teradata and all your other data sources with complete confidence that it's secure:

- **Detect:** Locate and identify sensitive data, creating a complete picture of your sensitive data landscape. Understand the sensitive data's connections.
- **Protect:** Rather than encrypting everything, which is an expensive, high-risk process, DgSecure replaces sensitive data elements with fictitious content using one of many available data replacement options or creates a cipher that restricts access to select users.
- **Audit:** Measure risk overhead with metrics and graphics that track data protection status and data access.

To learn more, visit

Teradata.com/Partners/Dataguise
Dataguise.com/teradata
Teradata.com/Products/Information-Security/Resources

17095 Via Del Campo, San Diego, CA 92127   Teradata.com