SPIRION™

eBOOK

Future-proofing sensitive data privacy and compliance

How contextual data classification gives organizations greater compliance agility

It's been an uphill battle for data owners to regain control over how their personal information can be used. Organizations looking to leverage this data have their hands full trying to keep up with constantly evolving data privacy laws, rules, and expectations.

While data is key to business insights, innovation, and transformation, business leaders across the spectrum view the increase in domestic and international regulations as something of a deterrent — or at least an impediment — for business growth.

Already, California, Virginia, Colorado, and Utah all have new or expanding data privacy regulations that will go into effect during 2023. These new regulations are only the beginning, as state and foreign governments are similarly enacting more stringent rights-based regulations.

Every new state law, federal regulation, or international standard adds complexity to an already complicated and fragmented regulatory environment. Many organizations feel they're unable to fully leverage the value of their data because they are handicapped by overwhelming data privacy and security concerns.

NEW DATA PRIVACY LAWS

Domestic	
	California Privacy Rights Act (CPRA)
	Colorado Privacy Act (CPA)
	Utah Consumer Privacy Act (UCPA)
	Virginia Consumer Data Privacy Act (VCDPA)
International	
	European Data Protection Board (EDPB)
	Chinese Personal Information Protection Law (PIPL)

By 2023, 75% of the world will have its personal data covered under some kind of privacy law with built-in subject rights requests and consent. The key will be whether privacy management programs can be automated.[1]

The constantly evolving nature of personal data and how it's being used undoubtedly make compliance mandates a moving target. However, investing the time and resources to build a privacy-forward strategy for your organization today will help you be better prepared for what the future holds.

To protect sensitive data, it must be located, then classified according to its level of sensitivity. With this classification in place, it is then possible to operationalize sensitive data management to ensure that only authorized people can gain access and that the data is always handled in full compliance with all relevant regulations and frameworks.

Organizations should look to next-generation data classification systems—capable of auto-classifying sensitive data based on content and context—to strike a balance between maintaining data privacy and unleashing the full potential and value of that information. Such systems make using and protecting sensitive data more efficient and scalable. These systems also offer a sustainable way to consistently meet new privacy standards because they can be dynamically modified to meet new or changing regulations.

New privacy reality is crushing old privacy strategies

Post-pandemic, the right to control one's personal data has never been more top of mind for consumers everywhere. Today's consumers want to exercise control over how their information is collected and used. They are also advocating for stricter laws that intensify the demand for stronger privacy and security measures — including requiring businesses to respond to consumers' requests for their personal data to be "forgotten."

In 2020 alone, Microsoft received nearly 3 million Data Subject Access Requests (DSARs), a submission by an individual to a corporation to find out what information of theirs has been collected and how it is being used and stored.[2] DSARs are a right afforded to consumers as per the California Consumer Privacy Act (CCPA/CPRA) and other regulations like the EU General Data Protection Regulation (GDPR).

This new-found empowerment is a boon to consumers but carries a significant burden for organizations. Failing to meet data security and privacy regulations can result in punitive fines and fees, legal liability, and irreversible damage to consumers' trust in a brand. Despite these obvious and significant risks, however, many organizations still rely on legacy data classification tools and manual processes that can't keep pace with the explosive growth in data production and consumption.

More cyberattacks mean more regulations — and more headaches for organizations

How sensitive data is collected and used is just one factor organizations must consider when ingesting and managing information. Data privacy is also a significant component of a broader data protection and governance strategy, critical to today's environment where cyberattacks continue to proliferate in both scale and sophistication.

In 2020 alone, Microsoft received nearly 3 million Data Subject Access Requests (DSARs).

Challenges in balancing data privacy, security, and business growth:

- Approximately 37% of organizations across the globe were the victims of ransomware attacks in 2021.[3]
- By 2024, over 80% of organizations will face modern privacy and data protection requirements that will result in steep penalties and reputational damage if not met.[4]
- 84% of respondents — enterprise data owners and their down-stream customers — reported that data privacy and security requirements will limit access to data at their organizations over the next 24 months.[5]

As a result, many organizations attempt to double-down on manual data privacy operations, which typically falls on the shoulders of employees who aren't experts in data privacy and have other business priorities to consider. This time-consuming process delivers inconsistent, error-ladened results that serve to further lock away valuable insights while increasing risk exposure. Worse, others forgo enhanced privacy efforts to better capitalize on important data intelligence, and risk significant penalties and damage to the brand's reputation when a privacy violation inevitably occurs.

The most prudent organizations implement next-generation data discovery, classification, and remediation solutions that utilize automation to fully protect and enhance the business value of their data.

Why do traditional classification methods fail?

Traditional data classification tools that use manual processes to determine how data is collected, who uses it, and how it must be restricted are ill-equipped to handle the requirements of modern business where the rules and regulations as well as cyberthreats are constantly evolving. Today's data is complicated: it has various ranges of sensitivity, is governed by an increasing patchwork of regulations, resides on-premises, in the cloud, and increasingly on endpoints like employee laptops—and there's a lot of it. In fact, around 1.145 trillion MB of new data is created every single day.

The amount of data created since the start of digital storage will be less than half the amount created between 2020 and 2025.[6]

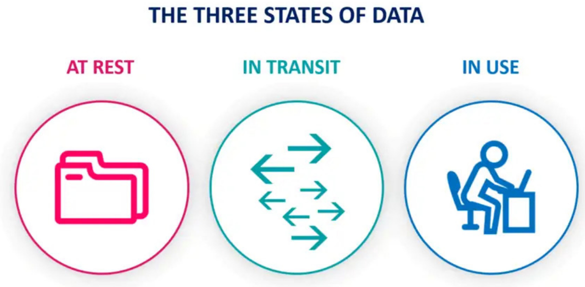
PROBLEMS WITH MANUAL DATA CLASSIFICATION METHODS

Inaccurate	Classification duties fall on a team without the proper skill set.
Subjective	Two people may classify the same data in different ways.
Inconsistent	Human error leads to inconsistent classifications.
Tedious	Manual classification is frustrating and interferes with user workflows.
Static	Classifications don't evolve with changing regulations or security requirements — nor reflect dynamic changes in data.
Insufficient	Manual classification only addresses new data as it is created; it doesn't address data-at-rest.
Inappropriate	The policies and practices in place don't solve the right problem.
Political	Departmental politics take precedence over effective data security.

According to Gartner, many organizations have failed or stalled data classification deployments because they attempted to assign classification labels to content without first understanding their data, its characteristics, and its usage pattern.[7] Understanding these factors is critical when considering the launch of a data classification program.

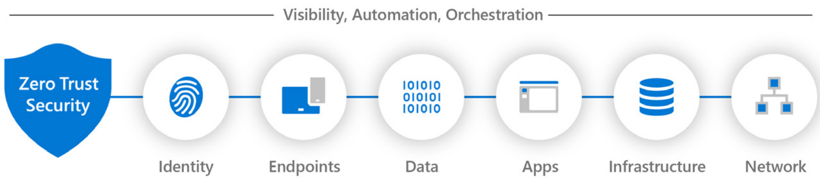
Further, most tools can't handle classification for both data at rest and in transit, meaning that classifications are at increased risk of being lost or misapplied whenever a file is moved, copied, accessed or otherwise used. Today's enterprise must go beyond simply separating and organizing data into relevant groups ("classes") based on their shared characteristics, such as their level of sensitivity and the risks they present, or the compliance regulations that protect them. The lack of persistent classification throughout the data lifecycle significantly increases the possibility of unauthorized access or exposure of personally identifiable information (PII) or other highly sensitive data.

VISUALIZE DATA AT REST, IN TRANSIT, IN USE[8]



As more organizations adopt Zero-Trust frameworks to bolster systems and data protection, they also need next-generation technologies capable of more granular data protection controls, greater visibility into data stores anywhere in their environment, and robust intelligent automation that can classify and update data files to comply with any and every new data privacy rule and regulation — without human intervention.

ZERO-TRUST FRAMEWORK[9]



Spirion is step one for strengthening sensitive data privacy

As organizations synthesize their privacy-forward data protection and compliance strategies, Spirion is step one for demonstrating compliance with new and emerging regulations, today and in the future.

Today's enterprise use of data has expanded. In addition, new location-specific data privacy regulations, different cybersecurity frameworks like NIST, ISO 27001 and other guidelines necessitate advanced classification schemas that extend beyond standard sensitivity classes into other aspects of data that need to be categorized, such as the business processes where the data is being used, where the data is being stored and other context.

Spirion is step one because sensitive data cannot be proven to be compliant with complex, evolving regulations unless it has first been discovered and classified accurately. Enhanced data privacy protection and compliance begin with Spirion's proven 98.5% accurate discovery of data anywhere and everywhere in your environment, and continues with context-based, automated data classification that persists throughout the file's entire lifecycle.

Working in the background without user intervention, Spirion eliminates the need for manual data classification and all the inadequacies and inefficiencies that come with it. In the process, it ensures that all data at rest gets auto-classified based on content and business context including—how the data was collected, the purpose for which it was collected, and the preferences around how it can be used—so that sensitive data ends up in the right location, subject to the right protections and policies.

HOW SPIRION DATA CLASSIFICATION SOLVES DATA PRIVACY ISSUES

Automated	Data is automatically labeled based on its purpose of collection, the process through which it was collected, and its privacy level. The labels are then federated across the entire IT environment.
Purposeful	Each piece of information is classified with context around the data's purpose. This includes dynamic classification labeling, which updates classification labels when data is modified or added.
Persistent	Data is automatically categorized and tagged at the file level based on sensitivity and existing information, security policies, and processes. Regardless of how many times data is moved or copied, metadata tags follow the data.
Contextual	Gives organizations more flexibility in how they can organize and granularly define their data to stay compliant, including: sensitivity, purpose for collecting the data, business processes, applicable regulatory guidelines, consumer preferences, and custom categories.
Visual	At-a-glance, user-friendly icons and classification markers for each piece of data make it easy for data governance teams to understand what data is sensitive based on visual indicators.

Context-rich data classification addresses today's data privacy challenges (and tomorrow's)

It's only natural that the process of separating and organizing data into relevant groups based on how sensitive they are or what kinds of risks they present should evolve as laws and regulations to protect those users evolve.

Unlike legacy technologies that apply blanket sensitivity labels based on loosely defined policies or the subjective preferences of individual data handlers, context-rich data classification adds layers of granularity and precision to classifications. This helps organizations better define their data to stay compliant with data privacy and security mandates that require documentation of the reason the data is being collected in the first place.

Granular, contextual classification categories offer more precision in how enterprises organize, define and control data to stay compliant according to: sensitivity, purpose for collecting the data, business processes, applicable regulatory guidelines, consumer preferences, and custom categories. Spirion dynamically updates tags and rules as data sensitivity evolves and can be easily adapted to suit virtually any data profile. So, if a law changes or a state creates a new data privacy law, as Utah just did, organizations can easily adapt to stay compliant. And when remediation is necessary, having data sorted and secured by class allows countermeasures to take effect instantaneously and run independently.

For many organizations, it's the missing link between effectively unlocking the full business potential of their enterprise data while minimizing non-compliance events that can result in litigation, multi-million-dollar penalties, reputational risks, loss of consumer trust, and other negative impacts.

SPIRION’S CONTEXT-RICH CLASSIFICATION CATEGORIES

Sensitivity	Sets the level of data confidentiality. Default categories include public, internal, confidential, or restricted. For example, an individual’s date of birth can be segmented away from their personal health information, the latter of which requires more security.
Process	Provides context around the business processes where data is stored and being used to ensure organizations have the individual user consent to use private data—a major stipulation of GDPR and CCPA/CPRA. It also provides context for why the data is being used.
Purpose	Provides context around what an organization is doing with data and the reasons they are holding onto it in order to comply with GDPR, CCPA/CPRA and other data privacy regulations. This allows organizations to transparently inform customers why their information is being collected and how it will be used, such as for data analysis, marketing campaigns, or improving customer experience.
Preference	Links data to customer preferences in terms of how they want their data used or shared. For example, customers may opt out of having sensitive data like protected health information or personally identifiable information being used for email solicitations or shared with third parties.
Regulatory	Ties data to regulations that govern it to help organizations automate compliance for CPRA, PCI DSS, GDPR, HIPAA, and others. For example, a classification that focuses on healthcare documentation can include HIPAA-specific metadata.
Custom	Classifies data according to user-specific needs to give organizations added flexibility in how they map to and follow specific security frameworks and regulatory standards. For example, the company can map to frameworks like NIST or ISO 27001 and implement the specific security protocols for each.

Making the case for contextual data classification in modern data privacy and compliance strategies

Understanding what constitutes personal information and where it resides across an enterprise is fundamental to protecting that data. While not a simple process, it is a necessary one, and is the predicate for all forms of sensitive data protection.

Among the many benefits of context-rich data classification is a privacy-forward posture that significantly improves and proves compliance with rapidly expanding regulatory requirements. For example, among the most significant developments of 2021 were the passage of two rights-based data privacy laws at the state level, the Virginia Consumer Data Protection Act (VCDPA) and the Colorado Privacy Act (CPA). These laws represent an acknowledgement of the need to grant to individuals the right to have control over how others use their personal data. Complying with these stricter regulations requires data controllers and their processors to employ more advanced data classification schemas and to develop and maintain records of their processing activities, including:

- **the purpose(s) for processing**
- **descriptions of the categories of data subjects and categories of personal data**
- **the categories of recipients to whom the personal data have been or will be disclosed**
- **data retention and security measures.**

Under these regulations, when a consumer wants to know what data the organization is collecting about them as part of a data subject access request (DSAR), an organization must be able to disclose the purpose for collecting the data.

The high cost of non-compliance:

- Violations of the CCPA are subject to enforcement by the California attorney general's office, which can seek civil penalties of $2,500 for each violation or $7,500 for each intentional violation after notice and a 30-day opportunity to cure have been provided.[10]
- Infringements of GDPR requirements could result in a fine of up to €20 million, or 4% of the firm's worldwide annual revenue from the preceding financial year, whichever amount is higher.[11]

Spirion automates the process for rapid turnaround when such requests are made to uphold the rights of data subjects. By retrieving the set of documents with data about a given individual, context-rich classification meets the GDPR Art. 30 Record of Processing Activities (RoPA) mandates (and similar ones) by classifying the purpose for collecting the data and other details. As a result, Spirion helps you comply with regulations by demonstrating how data flows throughout your environment to prove its compliance and security. Context-rich data classification can bolster your organization's ability to respond to regulatory requirements and proxy protection efforts in a manner that is seen as material, demonstrable, and proven.

HOW CONTEXTUAL CLASSIFICATION ENABLES REGULATORY COMPLIANCE

Type	Regulation	How Data Classification Helps
State	CPRA, VCDPA, UCPA	Upholds the rights of data subjects, including satisfying data subject access requests by retrieving the set of documents with data about a given individual.
International	GDPR	Prevents unauthorized disclosure or modification. Context-rich classification meets the Record of Processing Activities (RoPA) by classifying the purpose for collecting the data and other details.
Industry	HIPAA Healthcare	Locates where all health records are stored to implement security controls for proper data protection.
Industry	PCI DSS Financial	Identifies and secures consumer financial information used in payment card transactions.
Framework	NIST	Under this security framework, data discovery and classification is step one. Categorizing data helps federal agencies properly architect and manage their IT systems.
Standard	ISO 27001 and 27002	Classify information according to value and sensitivity to meet requirements for preventing unauthorized disclosure or modification.

Spirion unburdens IT teams of time-consuming manual work, significantly enhances a company's ability to comply with ever-changing compliance requirements and improves risk-related decision-making and responses with continually updated data. In the event of a data breach, having compliant security measures in place based on accurate classification can even help to reduce the fines imposed by regulatory agencies.

ILLINOIS COLLEGE

CASE STUDY

Illinois College secures an agile framework for data privacy

CHALLENGE

When the Department of Education began requiring Title IV higher education institutions that process U.S. federal student aid to conduct audits to assess their compliance with the Gramm-Leach-Bliley Act (GLBA), Illinois College knew they needed help.

The GLBA governs how universities collect, store, use, and safeguard the bank account information, social security numbers, financial aid, tuition payments, and other sensitive financial data of students and their guardians. At the time, Illinois College was relying on manual methods to identify and secure their sensitive data. They knew it wasn't going to suffice to meet GLBA audit objectives and future regulatory requirements.

In preparing for their upcoming audit, Illinois College needed an automated approach to identify the location of their sensitive data to ensure its privacy. A Spirion-first approach to proven, accurate and automated data discovery and classification offered the reinforcements Illinois College needed to protect sensitive information across campus in time for its GLBA audit.

> "Spirion's automated approach gives us full visibility to better understand, control, and protect sensitive data without burdening our staff or risking human error. **Furthermore, by reducing our sensitive data footprint, we can better focus our limited resources and data security spend."**
>
> —**Marc Benner,** *Assistant Chief Information Officer*

SOLUTION

The first step to a radically improved data privacy solution for Illinois College was to automate their data privacy processes. This included scanning 28 terabytes of structured and unstructured data on servers and roughly 250 faculty, staff, and lab endpoints. Spirion combed the entire IT infrastructure for hard-to-find files to provide full visibility into Illinois College's sensitive data.

Once this had been done, the school used Spirion to automate its data classification process according to government compliance regulations and internal campus policies. This step allowed Illinois College to take critical remediation actions, including electronically shredding unnecessary information while securing the data they still needed and automatically quarantining files to a more secure location. Automated triggers notified the IT team of policy violations for immediate response.

OUTCOME

Implementing Spirion was Illinois College's first foray into an automated data governance process, and it proved to be a successful one at that. The school managed to meet GLBA rules on a tight deadline and fortified their data privacy management program to meet new compliance standards that will likely crop up in the future.

Unburdening their staff of tedious, labor-intensive and error-prone discovery and classification has given Illinois College full visibility into their sensitive data, while also affording them the flexibility to refocus their limited IT resources.

Automated, purposeful, and persistent data classification: a Spirion innovation

Robust data protection is essential to safeguarding sensitive, personal and regulated data from unauthorized access or compromise, while also protecting organizations from financial loss, reputational harm, consumer trust degradation, and brand erosion. Teams that adopt a privacy-forward mindset will have no trouble ensuring regulatory compliance, preventing costly data breaches from happening, and adapting to new laws and regulations in the future.

That's because data classification and data privacy aren't just related processes—they're inextricably linked. Classifying data accurately and thoroughly are prerequisite for ensuring it complies with location-specific data privacy regulations, different cybersecurity frameworks like NIST, and other evolving guidelines. And when inevitable issues and attacks arise, having data segregated and secured by class helps to catalyze the response and focus the defenses. In that way, Spirion isn't just a potent addition to upholding data privacy rules—it's the foundation that elevates everything else.

More than a data classification tool, Spirion provides a significant upgrade to data privacy and the tools it relies on by automatically moving data out of the wild and into the most secure location possible. As the innovators of context-rich data classification run by artificial intelligence, Spirion has made this transformative capability not just possible but accessible.

Stop struggling to get every piece of data classified correctly—and risking data privacy and compliance issues in the process. Spirion has the solution.

Talk to a Spirion data security and compliance expert for a demo today: expert@spirion.com

Spirion has relentlessly solved real data protection problems since 2006 with accurate, contextual discovery of structured and unstructured data; purposeful classification; automated real-time risk remediation; and powerful analytics and dashboards to give organizations greater visibility into their most at-risk data and assets. **Visit us at spirion.com.**

ENDNOTES

[1] "Gartner predicts privacy law changes, consolidation of cybersecurity services and ransomware laws for next 4 years," ZDNet, Oct. 20, 2021

[2] "California Consumer Privacy Act (CCPA) Notice for California Consumers." Microsoft, June, 2021.

[3] Dickson, Frank & Kissel, Chris. "IDC 2021 Ransomware Study - Where you are matters!" July 20, 2021.

[4] Willemsen, Bart. "Hype Cycle for Privacy." Gartner. July 13, 2021.

[5] "DataOps Dilemma: Survey Reveals Gap in the Data Supply Chain." Immuta. August, 2021.

[6] IDC, "Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts." IDC, March 24, 2021

[7] "How to Overcome Pitfalls in Data Classification Initiatives." Gartner, July 31, 2014

[8] https://www.sealpath.com/blog/protecting-the-three-states-of-data/

[9] https://docs.microsoft.com/en-us/security/zero-trust/

[10] "The California Consumer Privacy Act: Frequently Asked Questions." Bakerlaw.com. Accessed April 7, 2022.

[11] "What are the GDPR Fines?" GDPR.eu. Accessed April 7, 2022.

SPIRION™