# Hype Cycle for Data Security, 2023

Published 14 July 2023 - ID G00792194 - 128 min read

By Analyst(s): Brian Lowans

Initiatives: Security of Applications and Data; Meet Daily Cybersecurity Needs

Security and risk management leaders should adopt innovations like data security posture management and data security platforms and prepare for the impacts of quantum computing and AI. Review these and other Hype Cycle entries to support business goals and mitigate data security and privacy risks.

# Analysis

## What You Need to Know

Cloud service providers (CSPs) have transformed how data analytics and data pipelines can be innovated. That these pipelines can now be deployed dynamically poses a new challenge for data security teams, which have to adapt. This results in increased data storage, backups and subsets of data, extractions and combinations of data formats. In turn, there are huge challenges relating to data observability and data lineage when it comes to discovering sensitive data and creating consistent categorizations and classifications. As data is stored, accessed and processed across different CSP architectures that span different geographic jurisdictions, complex data residency, security and privacy risks arise. [1,2]

Maintaining consistent data security is difficult because so many products provide siloed security controls, use proprietary data classification, act on specific repositories or processing steps, and do not integrate with each other. This restricts organizations' ability to identify and deploy adequate, and consistent, data security controls while balancing the business need to access data throughout its life cycle. However, sovereign data strategies is a new entry on the Hype Cycle that will support data security governance, privacy impact assessment, financial data risk assessment (FinDRA) and data risk assessment.

Technologies need to evolve to support both data and analytics categorizations for quality, accuracy, ethics and life cycle, and data security categorizations for confidentiality, integrity, availability and privacy. This will enable data security controls to support business access requirements by integrating data discovery, classification, and augmented data catalog/metadata management.

Increasingly, technologies are using generative AI techniques to increase performance, and privacy laws are changing guidance as a result. Data security strategies and technologies also need to adapt to the risks of access to data by generative AI.

All cryptographic technologies will need to evolve to cope with the future threat of quantum computing, which is increasing the need for innovative technologies such as:

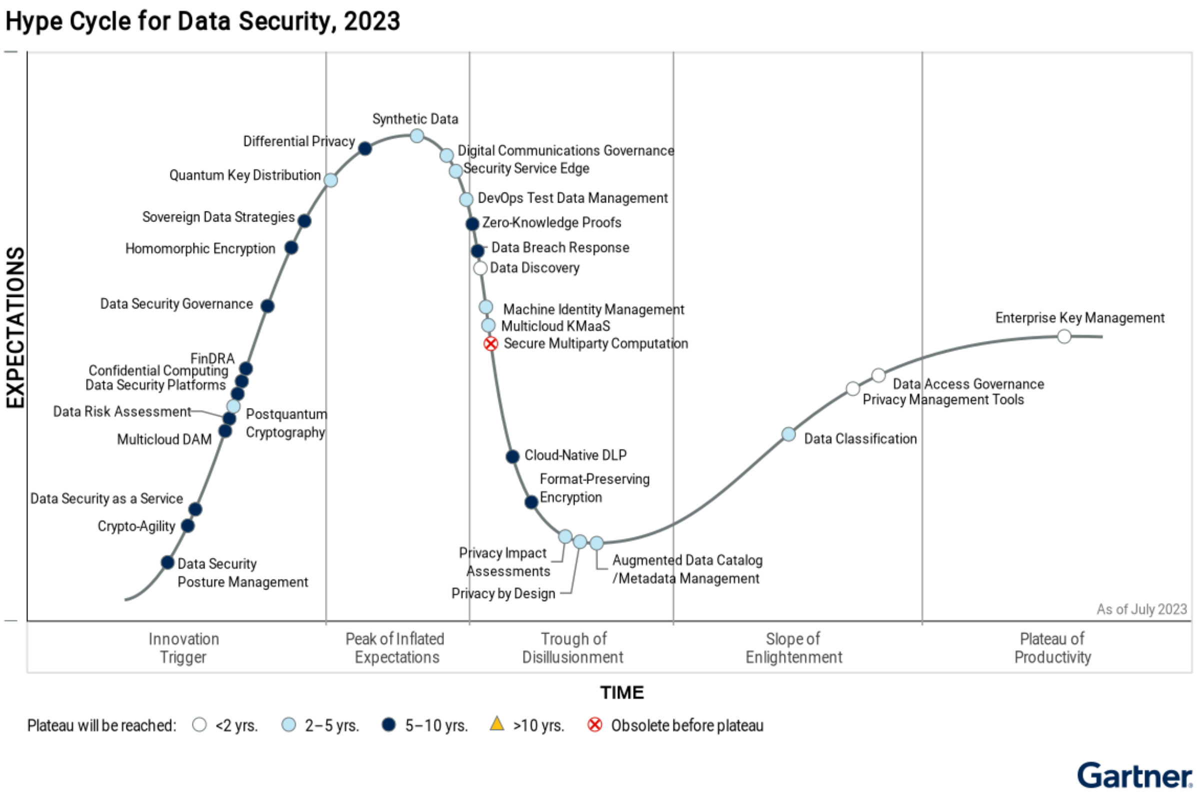
- Crypto-agility
- Postquantum cryptography
- Quantum key distribution

## The Hype Cycle

This Hype Cycle covers many aspects of data security that security and risk management leaders must review in relation to their risk appetite and where data is stored, processed and accessed. They include:

- **Data security governance, privacy and risk:** See the Hype Cycle entries for data security governance, data risk assessment, privacy impact assessments, data breach response, privacy by design, sovereign data strategies and FinDRA.
- **Data discovery, categorization, and classification of structured and unstructured data:** See the entries for data security posture management, data discovery, data classification and augmented data catalog/metadata management.
- **Data processing and analytics across endpoint, application or storage layers:** See the entries for DevOps test data management, digital communications governance and privacy management tools.
- **Anonymization, pseudonymization, privacy-enhanced technologies and other data protection techniques:** See the entries for crypto-agility, postquantum cryptography, confidential computing, homomorphic encryption, quantum key distribution, differential privacy, synthetic data, zero-knowledge proofs, multicloud KMaaS, secure multiparty computation, format-preserving encryption, enterprise key management and machine identity management.
- **Monitoring access, activity, alerting and auditing of user activity with data:** See multicloud DAM (database activity monitoring), security service edge, cloud-native DLP (data loss prevention) and data access governance.
- **Multicloud platforms with multifunctional data security functionalities:** See the entries for data security as a service and data security platforms.

Figure 1: Hype Cycle for Data Security, 2023



## The Priority Matrix

It is important to look for technologies that integrate multiple security controls to simplify orchestration of these controls. Some choices will support transformational or high levels of business benefit, resulting in reduced operational complexity and cost. For example, the evolution of data security platforms will be highly beneficial. Some vendors are developing techniques that support data security governance through focused data or privacy risk assessments, and the emergence of data security posture management will transform the implementation of data risk assessments across data security technologies. The combination of financial and security risk management will drive data security investment decisions supportive of desired business outcomes.

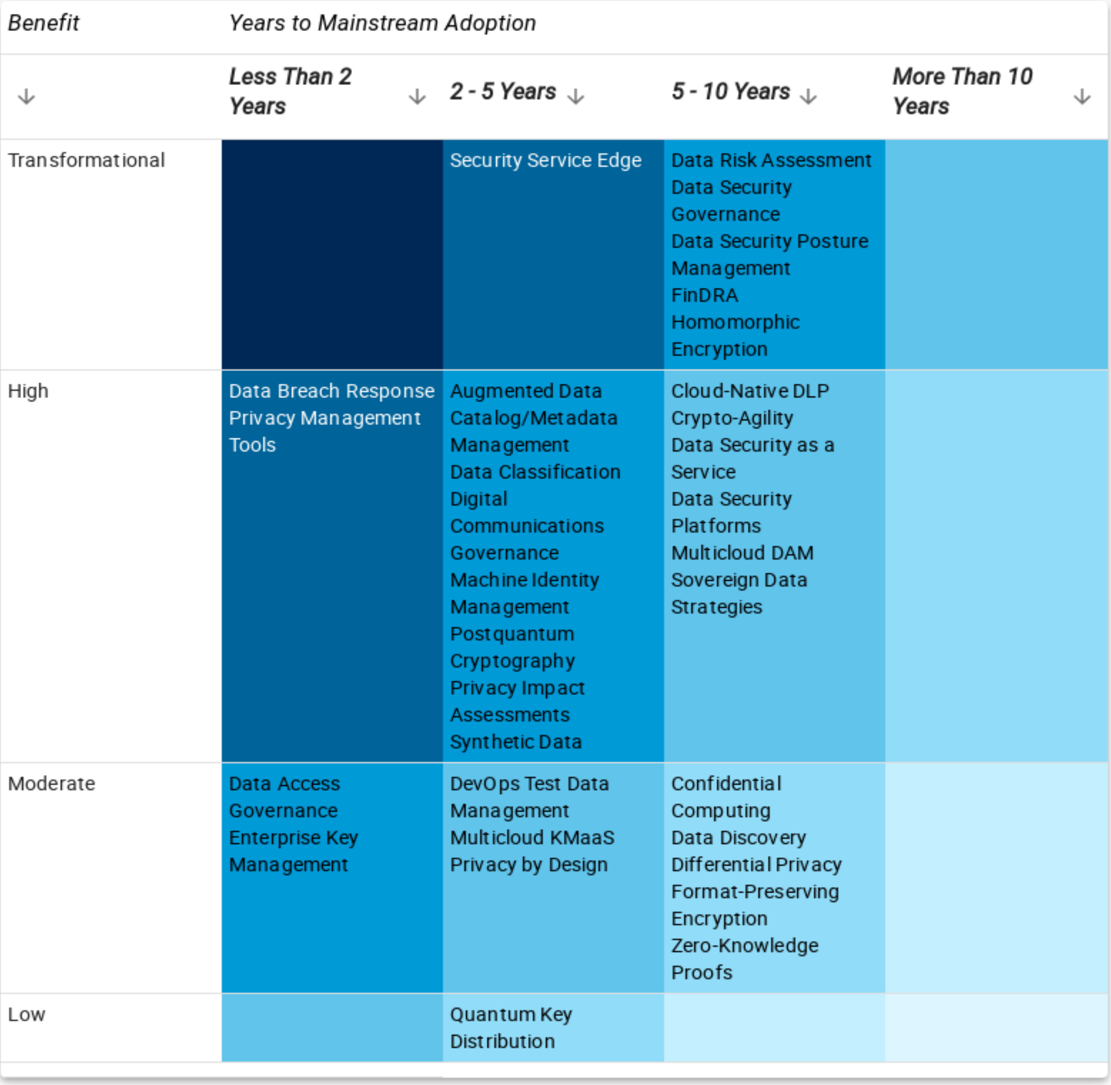
In the 2022 Gartner Shifting Cybersecurity Operating Model Survey, the desire to achieve and maintain consistency in cybersecurity policies is identified as the main reason for centralizing cybersecurity risk decision making by assigning it to an enterprise security steering committee. [3] Many organizations are also centralizing enterprise cybersecurity decision rights in a central cybersecurity team and/or steering committee, while making resource owners accountable for risk decisions. These developments are increasing the need for consistency of product functionality, which is achieved through convergence that reduces the number of management consoles and the inherent complexity of, and the number of staff required for, effective policy orchestration.

Convergence is observable in the following respects:

- Data discovery and classification and data security posture management are the subject of innovation and integration across most data governance and security markets, and integration with augmented data cataloging will follow.
- Anonymization and pseudonymization technologies are integrating within enterprise key management and multicloud KMaaS offerings, and will likely use synthetic data.
- Integration of data monitoring and protection techniques is emerging across technologies such as data security as a service, data security platforms and multicloud DAM.
- Data security governance, data risk assessments, FinDRA, privacy impact assessments, data breach response technology, sovereign data strategies and data security posture management are increasingly needed to implement consistent policies, especially through their data residency impacts and as new privacy laws emerge. Convergence of these technologies will make processes more effective.

Table 1: Priority Matrix for Data Security, 2023

(Enlarged table in Appendix)



| *Benefit* | *Years to Mainstream Adoption* | | | |
|---|---|---|---|---|
| ↓ | *Less Than 2 Years* ↓ | *2 - 5 Years* ↓ | *5 - 10 Years* ↓ | *More Than 10 Years* ↓ |
| Transformational | | Security Service Edge | Data Risk Assessment<br>Data Security Governance<br>Data Security Posture Management<br>FinDRA<br>Homomorphic Encryption | |
| High | Data Breach Response<br>Privacy Management Tools | Augmented Data Catalog/Metadata Management<br>Data Classification<br>Digital Communications Governance<br>Machine Identity Management<br>Postquantum Cryptography<br>Privacy Impact Assessments<br>Synthetic Data | Cloud-Native DLP<br>Crypto-Agility<br>Data Security as a Service<br>Data Security Platforms<br>Multicloud DAM<br>Sovereign Data Strategies | |
| Moderate | Data Access Governance<br>Enterprise Key Management | DevOps Test Data Management<br>Multicloud KMaaS<br>Privacy by Design | Confidential Computing<br>Data Discovery<br>Differential Privacy<br>Format-Preserving Encryption<br>Zero-Knowledge Proofs | |
| Low | | Quantum Key Distribution | | |

Source: Gartner (July 2023)

## Off the Hype Cycle

- Augmented data cataloging and MMS has been renamed augmented data catalog/metadata management.

- Blockchain for data security no longer appears on the Hype Cycle because it has become obsolete before reaching the Plateau of Productivity.

- CASBs (cloud access security brokers) have been absorbed into security service edge.

- Cloud data protection gateways no longer appear on the Hype Cycle because they have reached the Plateau of Productivity.

- CSP-native DLP has been renamed cloud-native DLP.

- Data discovery and management has been renamed data discovery.

- Secure instant communications has been replaced by digital communications governance.

- TLS decryption platform no longer appears on the Hype Cycle because it has become obsolete before reaching the Plateau of Productivity.

## On the Rise

### Data Security Posture Management

**Analysis By:** Brian Lowans, Andrew Bales, Joerg Fritsch

**Benefit Rating:** Transformational

**Market Penetration:** Less than 1% of target audience

**Maturity:** Embryonic

#### Definition:

Data security posture management (DSPM) discovers previously unknown data across cloud service providers (CSPs) and categorizes and classifies unstructured and structured data. As data rapidly proliferates, DSPM analyzes data maps and data flows to assess who has access to data to determine the data security posture and exposure to privacy and security risks. DSPM forms the basis of a data risk assessment (DRA) and evaluation of the implementation of data security governance (DSG) policies.

#### Why This Is Important

As data proliferates across the cloud, organizations must identify privacy and security risks with a single product. DSPM will transform how they identify business risks that result from data residency, privacy, and security risks. Risks multiply because data locations and content are unknown, undiscovered or unidentified. Data sensitivity, data lineage, infrastructure configurations and access privileges must be analyzed. This has led to rapid growth in the availability and maturation of technology that can operate across a dynamic landscape.

#### Business Impact

DSPM uniquely discovers shadow data by creating and analyzing a data map and data flow to identify data locations and user access to data. This will create critical insights to previously unassessed business risks. DSPM then enables data security posture to be applied consistently across previously independent data security controls. This allows organizations to mitigate these business risks despite the speed, complexity, dynamics and scale of data deployments. This is a unique combination of properties provided via a single console.

#### Drivers

- Dynamic changes to data pipelines and services across CSPs are creating a variety of shadow data repositories because the data is unknown, undiscovered or unidentified. The shadow data will create risk because of its geographic location, or because it is potentially misconfigured or connected via pipelines with inappropriate user access privileges.
- Creating a data map of user access against specific datasets has been a complex process in the past because traditional data security and IAM products are siloed in the way they operate.
- To achieve consistent analysis, organizations need to map and track the evolution and data lineage across structured, semistructured and unstructured formats, and across all potential data locations and shadow data. This is an emerging driver.
- The growth of regulations that require a DRA has created the need for tools that can assess DSG policies.
- There is a need to protect data against exposure (for example, via cloud misconfigurations, excessive access privileges or data residency risks that arise due to geographic locations and access pathways to data). These infrastructure risks are also driving integration or partnerships with cloud native application protection platforms (CNAPP).
- Organizations are looking to identify security gaps and undue exposure using a combination of data observability features, such as real-time visibility into data flows, risk and compliance with data security controls.

### Obstacles

- DSPM products today are mostly provided by early stage vendors, which may deter some organizations from deploying them. Given that there are several early stage vendors in this category, DSPM vendors are facing competitive pressure.
- Each DSPM product has differing abilities to map user access privileges, to identify, discover and track data across the hybrid IT architecture and to identify associated risks. Organizations face difficulties in achieving consistent product capabilities, or integration with selected security controls to identify data vulnerabilities and enable remediation.

- Currently, DSPM products will integrate with a limited set of third-party security products, leading to difficulty in orchestrating the output analytics across those products. There is no standard way to remediate issues identified, and approaches vary.

### User Recommendations

- Start with DSG to establish which datasets need specific DSG policies to be assessed via DSPM.
- Compare the DSPM products to establish their ability to find shadow data repositories and map the various datasets across your architecture. Evaluate their ability to integrate with the existing cloud security stack.
- Treat investments as tactical, and limit contracts to one to two years, as the market matures and further consolidates.
- Evaluate and compare the response options among different DSPM tools to create a data map. Then, detect changes to the data map, especially when undue exposure, security risks or compliance risks are encountered.
- Assess how quickly a DSPM product can be implemented to establish an assessment and, therefore, to establish the frequency required to update those assessments, given the evolution of your data portfolios.
- Identify any third-party data security products that can integrate with DSPM and investigate the ability of those products to use DSPM functionality.

### Sample Vendors

BigID; Concentric AI; Cyera; Dig Security; Flow Security; IBM; Laminar; Securiti; Sentra; Symmetry

### Gartner Recommended Reading

Use the Data Security Governance Framework to Balance Business Needs and Risks

Use a Data Security Steering Committee to Realize Data Security Governance Objectives

A Data Risk Assessment Is the Foundation of Data Security Governance

Innovation Insight: Data Security Posture Management

### Crypto-Agility

**Analysis By:** Mark Horvath

**Benefit Rating:** High

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

#### Definition:

Crypto-agility is the capability to transparently swap out encryption algorithms and related artifacts in an application, and replace them with newer, different and, presumably, safer algorithms. Because quantum computing poses an increasing threat to existing cryptography, Gartner expects crypto-agility to be a significant differentiator for technology vendors.

#### Why This Is Important

As quantum computing matures, asymmetric encryption faces the threat of being damaged or broken during the next five to seven years. Algorithms threatened are hard-wired into many applications, and most organizations lack a clear idea of the cryptography they're using.

Standard network/PKI protocols (e.g., Diffie-Hellman key exchange and TLS) will need quantum-safe deployments.

The National Institute of Standards and Technology (NIST) is standardizing quantum-safe alternatives for widespread use.

### Business Impact

- Applications/communications using encryption or PKI need to be inventoried. Data about algorithms, key sizes, expiration dates and uses must be gathered into a metadata database.
- Suitable replacements must be identified for applications and use cases.
- The new algorithms are not drop-in replacements for cryptography, so alternatives must be tested and replacement policies must be generated.
- Vendor cryptography products must be identified, and vendors will need to provide a replacement schedule.

### Drivers

- The development of quantum computers capable of breaking cryptography through Shor's or Grover's algorithms is estimated to be five to seven years away. This is much shorter than the typical life span of sensitive information found in most organizations.
- Keybreaking with Shor's algorithm goes linearly with key size but is limited by the number of available qubits; once it's been achieved, larger keys sizes will fall quickly, leading to a short runway to implement changes.
- Most organizations that have inventoried their cryptographic metadata have found they already have considerable technical debt with respect to PKI (e.g., expired certs, deprecated algorithms, short keys). Cleaning that up will substantially lower the organization's risk profile.
- New algorithms have additional uses that can foster novel business cases (e.g., stateful signatures, homomorphic encryption).
- Government organizations are beginning to require a quantum-safe encryption strategy for vendors selling to them (e.g., U.S. NSM-10, EO-14028). This includes all products or code that is included in the final product, including open-source software (OSS) and other vendor's products. Like with a software bill of materials (SBOM), this significantly expands the scope of these orders to include many vendors not otherwise selling directly to the government.
- Cryptography in vendor products will need to be identified, and vendors will need to provide a schedule for potential replacement.

#### Obstacles

- The time scale for change (five to seven years) is significantly beyond the expected tenure of many senior executives, leading to a “someone else's problem” mentality, leaving some decision makers to prioritize shorter-term projects.
- Many organizations don't know how to inventory their cryptographic usage, leaving them to choose between a vendor to do the work, or spinning up an internal program.
- The NIST standardization of current, approved algorithms has not been completed, leaving clients with nonstandard OSS versions of the algorithms.
- Many organizations lack the expertise needed to lead a cryptographic project of this magnitude.
- There is still some market confusion about what “quantum-safe” means, sometimes leading organizations to evaluate relatively useless technologies that talk about “quantum” but don't solve the problem.

#### User Recommendations

- Start looking at your cryptographic inventory sooner, rather than later. This will help scope the project and identify key systems and data, allowing a controlled rollover to quantum-safe encryption.
- Experiment with the OSS version of the NIST standardization candidates to determine what effect this will have on your infrastructure (e.g., lattice encryption is generally slower, and ciphertext sizes are larger than existing algorithms).
- Ask vendors what their plans are for replacing algorithms, and when quantum-safe versions of their products will be available.
- Explore off-brand uses for some of the new algorithms (e.g., homomorphic encryption).

#### Sample Vendors

Cryptomathic; DigiCert; Entrust; IBM; IronCore Labs; ISARA; Sandbox AQ; Thales; Utimaco; Venafi

#### Gartner Recommended Reading

Preparing for the Quantum World With Crypto-Agility

Infographic: How Use Cases Are Developed and Executed on a Quantum Computer

Cool Vendors in Quantum Computing

## Data Security as a Service

**Analysis By:** Joerg Fritsch, Brian Lowans

**Benefit Rating:** High

**Market Penetration:** Less than 1% of target audience

**Maturity:** Embryonic

### Definition:

Data security as a service (DSaaS) provides data security and protection capabilities as APIs. Organizations hand over their data to the service provider, which protects, transforms and shares it back to them or with third parties, while achieving the required compliance and secrecy goals.

### Why This Is Important

With the exception of payment processing, data security controls are generally managed and applied by the end-user organizations directly. Data security as a service (DSaaS) provides cloud-based, automated data security and protection capabilities such as encryption and masking. Data security and protection provided by cloud-based services is a paradigm change applying the appropriate controls for multiple use cases, ecosystem partners or jurisdictions in a scalable and flexible model with impressive performance.

### Business Impact

DSaaS makes complex or expensive data security controls accessible to mainstream organizations. It enables clients to shorten the deployment times of data security controls, bringing them into a position to match the speed of cloud and DevOps initiatives. It achieves this by making the required data security controls and data transformations readily available through, for example, cloud-based APIs. Customers can, in theory, start right away, without the need for in-depth expertise.

Drivers

- Data security controls and data security architectures are frequently complex, and the customer is loaded with both hardware and software — putting thorough, scalable and agile data security out of reach of most organizations. At the same time, most organizations must rearchitect data security controls to balance privacy regulations and a constantly evolving threat landscape with an increased need to share data internally and externally. For example, this may be required for artificial intelligence and machine learning use cases, or monetization of data with ever-changing ecosystem partners. This tension will lead to accelerated adoption of DSaaS.

- Startup DSaaS vendors have new, proprietary intellectual property that could potentially mitigate lack of trust in the DSaaS provider. Alternatively, selected DSaaS are using privacy-enhancing computation to achieve this.

- Mature clients believe that if data can flow securely among individuals, organizations and governments, wiser decisions can be made and better outcomes can be delivered, both for business and society as a whole. However, an increased amount of data-related regulations and associated legal, security and privacy risks are blocking this data sharing. DSaaS will be instrumental in solving this challenge.

### Obstacles

- Organizations involved with data security are often tentative toward cloud and SaaS adoption. Data security controls (for example, encryption and tokenization) applied on-premises before load are seen as compensation for a lack of trust in the cloud service provider. However, it is exactly these controls that DSaaS delivers as cloud-based services.

- Companies may refrain from DSaaS adoption if the required data security controls cannot fully be matched with DSaaS offerings. DSaaS solutions that exclude certain data types represent a common limitation.

- Privacy regulation uncertainty for controls that are used to secure personal data may significantly delay adoption of DSaaS in some geographies.

- DSaaS offerings are not interoperable. Each assumes to be the authority for its data. At a minimum, vendors will need interoperable capabilities to support external data sharing between customers that do not use the same DSaaS.

- Current DSaaS offerings focus on data encryption, data tokenization, data masking and sharing use cases. They do not include behavioral-based controls and authorization.

### User Recommendations

- Validate if the integration approaches offered by DSaaS (e.g., API-based approaches and proprietary connectors or agents) are viable for your environments and will lead to the desired user experience.

- Question and plan for DSaaS constraints like API request limits, delivered event limits, network latency and storage quotas.

- Ensure the DSaaS provider's approach is compliant with your applicable privacy regulations. Question the security practices and posture of the DSaaS (for example, use shared assessments and certifications such as ISO 27001 to aid in evaluation).

### Sample Vendors

Basis Theory; Google; InCountry; Inpher; Skyflow; Spring Labs; VectorZero; Very Good Security

### Gartner Recommended Reading

The State of Privacy and Personal Data Protection, 2020-2022

## Data Risk Assessment

**Analysis By:** Brian Lowans, Joerg Fritsch

**Benefit Rating:** Transformational

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Adolescent

### Definition:

A data risk assessment (DRA) should review the implementation of data security governance policies across all security products to identify any gaps or inconsistencies. These policies must balance business needs for access to data with risks caused by inappropriate access, data residency, compliance, privacy and breaches.

### Why This Is Important

A DRA should be used to regularly assess how business risks arise from gaps and inconsistencies in the controls applied by data security, privacy and identity management products. A DRA is fundamental to the successful implementation of data security governance (DSG).

### Business Impact

A DRA can:

- Enable the mitigation of business risks that have financial impacts. It focuses attention on the data risks an organization should prioritize based on its risk appetite.
- Support the business risk prioritization process to clarify how far each risk will be mitigated, given the budget and the impacts on business outcomes. Prioritization will typically focus on a financial DRA (FinDRA) and the economic impact to the business in order to decide how large the security budget should be.

Drivers

- Business stakeholders who form a data security steering committee (DSSC) to implement DSG will use a DRA to assess how successfully business risks are managed.

- It is important to use DRAs to identify gaps and inconsistencies in how data security policies are implemented by data security, privacy and identity and access management (IAM) products, and provide informed recommendations on changes to controls or new products.

- Every decision to mitigate data-related business risks will benefit from a DRA to establish how each risk might evolve. Such decisions also require a data categorization and classification process that can identify data in both structured and unstructured formats and use security, privacy and business metadata.

- Creating data maps can help analyze the impacts on data lineage and observability, which is important to identify risks, as data evolves during its life cycle in different formats, combinations, data residency and shadow data across hybrid IT architectures. This can also be supported through data security posture management (DSPM), which helps map data through its life cycle.

- Analyzing a data map also enables the assessment of the data access privileges provided to each user or machine account against the data in scope. Any gaps or inconsistencies in privileged access to data can then be identified in relation to DSG policies, and a data risk register can be created.

- There is a critical need to establish business support for DRA that connects data risks to business outcomes, so that data access requirements for project teams can be identified and assessed.

### Obstacles

- A DRA can only succeed if business leaders support the need for data security controls that may impact the business outcomes of each project or service.
- It is challenging to identify all user account privileges for each dataset associated with each project or service.
- Each data security product is siloed, because it typically applies only specific controls across the data flow path to specific repositories, and they neither integrate nor share policy enforcement.
- Each product deploys a proprietary data discovery, categorization and classification tool, each of which understands data differently and does not integrate with the others.
- The variety and geographic distribution of endpoints creates problems when identifying and analyzing user accounts' access to data repositories.
- Identifying all data security, privacy, IAM and application products that overlay the data flow from the data repository to the endpoints may require support from business leaders.

### User Recommendations

- Work with the DSSC to analyze how each business project processes data, associated business outcomes and impacts of business risks, and implement a DRA to identify how those business risks could emerge.
- Create a data map and risk analysis for each project to establish who has certain privileges to access various datasets, and establish how gaps and inconsistencies might create business incidents.
- Use tools to categorize and classify data consistently and provide data security posture management to help identify which data risks are not mitigated and how they might still create business risks.
- Use a DRA to establish and communicate data risks in the language of business risk in order to achieve business understanding of the level of achievable risk mitigation.
- Identify data risks creating gaps or inconsistencies in the data security policies and unsuccessful mitigation of business risks. Communicate these findings to the DSSC to gain business support for changes to staffing and budgets.

### Gartner Recommended Reading

A Data Risk Assessment Is the Foundation of Data Security Governance

Use a Data Security Steering Committee to Realize Data Security Governance Objectives

4 Critical Steps to Accelerate the Adoption of Data Security Governance

3 Steps to Effectively Capture and Communicate the Business Value of Cybersecurity Initiatives

Innovation Insight: Data Security Posture Management

## Multicloud DAM

**Analysis By:** Brian Lowans, Joerg Fritsch

**Benefit Rating:** High

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Adolescent

### Definition:

Multicloud database activity monitoring (DAM) provides centralized and consistent security across a variety of relational and NoSQL databases hosted on cloud service providers (CSPs). Multicloud DAM sets up and manages database user privileges, monitors user behavior with data, and provides database activity logs. DAM helps mitigate critical data residency, data security and privacy risks.

### Why This Is Important

Multicloud DAM provides real-time analytics and audit logs of user account activities with data, while accessing relational and NoSQL databases across different CSPs. It enables real-time management and policy enforcement of user privileges in relation to specific datasets according to role, attributes, responsibility and data residency. It mitigates privacy and security risks resulting from user and administrator activities, database vulnerabilities, and poor segregation of duties (SOD).

### Business Impact

Multicloud DAM functionality is critical to enforce data security policies that will enable regulatory compliance with data protection, financial, health and privacy laws. By monitoring and auditing user activity with data, it helps prevent malicious activity and manages appropriate business use of data, as well as the audit records that will be essential for incident responses.

### Drivers

- An increasing number and variety of databases are accessed across CSPs, or directly installed and managed by an organization. This is increasing the need for a single product to protect the data consistently.

- Multicloud DAM enables consistent, policy-driven user privileges against specific datasets and across database platforms. They can also be role-based or attribute-based providing data context against user privileges and activity or behavior. Other tools, such as identity and access management (IAM), security information and event management (SIEM), and user and entity behavior analytics (UEBA), do not.

- Multicloud DAM uniquely monitors user activities with data that will support compliance with a variety of data protection and privacy laws. This includes enforcement of SOD and appropriate business user access to data, vulnerability management, user activity monitoring, and a forensic audit record of all SQL activities. Some products also provide data protection techniques, such as encryption, tokenization, masking or confidential computing.

- Depending on the vendor architecture, DAM analyzes database logs or intercepts specific communication paths to the database to analyze and/or modify database access requests and stored audit logs. The DAM may also integrate with the cloud database provider to manage user data access privileges.

- Only a few CSP-managed database as a service (DBaaS) services offer native audit log capabilities. However, they typically do not record database activity with data context, which increases security and compliance risks. Functionality will also differ across each CSP offering. This drives the need for a multicloud DAM.

### Obstacles

- CSPs do not allow vendors to install agent software on the infrastructure layer. This prevents multicloud DAM from being able to block direct database access by administrators or CSP staff.

- When multicloud DAM is installed as a proxy service or gateway, it is possible to monitor the database logs, which enables direct monitoring and auditing of all SQL activity. However, it does not enable the blocking of SQL activity or direct access to the database.

- If multicloud DAM only intercepts SQL statements at the application layer, it will not monitor database logs. This results in a lack of data residency control and an inability to prevent access by administrators or CSP, unless encryption or tokenization is also applied.

- If multicloud DAM monitors only database logs, or if the proxy or gateway interfaces directly with the database below the application layer, then SQL analysis may not enable direct identification of the user account.

### User Recommendations

- Create policies to monitor or prevent access by administrators, application users and CSP staff by combining DAM with encryption or tokenization. Policies need to reflect appropriate access to each dataset by role, attribute, business purpose and data residency.

- Apply multicloud DAM to provide consistent data security governance policies for monitoring access to databases across disparate CSP environments, and to maintain data protection and privacy across geographic jurisdictions.

- Review the vendor product capabilities to turn on and access the full log history of all relevant databases with sufficient data storage to enable full analysis of their content. On average, this may require approximately one year of log data.

### Sample Vendors

Cyral; DataSunrise; Datiphy; IBM; Immuta; Imperva; Oracle; Privacera; Satori Cyber; SecuPi

### Gartner Recommended Reading

Securing Your AI Data Pipeline

2023 Strategic Roadmap for Data Security Platform Adoption

## Confidential Computing

**Analysis By:** Mark Horvath, Bart Willemsen

**Benefit Rating:** Moderate

**Market Penetration:** Less than 1% of target audience

**Maturity:** Emerging

### Definition:

Confidential computing is a security mechanism that executes code in a hardware-based trusted execution environment (TEE), also called an enclave. Enclaves isolate and protect code and data from the host system (plus the host system's owners), and may also provide code integrity and attestation.

### Why This Is Important

As privacy concerns and fines increase:

- Confidential computing combines a chip-level TEE with conventional key management and cryptographic protocols to enable unreadable computation. This enables a variety of projects where cooperation between different groups is critical, without sharing data or IP.
- The ongoing adoption of public cloud computing and the increased availability and viability of enclave technology allow data to be used in the cloud in a more trusted manner.
- Cross-border transfers are a complex, key component to many businesses, addressed directly by confidential computing.

### Business Impact

Impacts include:

- Confidential computing may mitigate one of the major barriers to cloud adoption for highly regulated businesses, sensitive data workloads, or any organization concerned about unauthorized third-party access to data in use in the public cloud. This includes potential access by the infrastructure provider.

- Confidential computing allows a level of data confidentiality and privacy controls between competitors, data processors and data analysts that is very difficult to achieve with traditional cryptographic methods.

Drivers

- Cloud adoption is increasing alongside ongoing concerns regarding potential access to personal data by cloud service providers (CSPs).
- Global data residency restrictions are ongoing, with a need to segment content away from even the CSP with a level of independent assurance.
- Competitive concerns — not just around personal data, but also intellectual property — are spurring the adoption of confidential computing. This includes the need for confidentiality and protection against any third-party access, protection of the method of processing (including algorithmic functions) and protection of the data itself.
- Confidential computing has been mentioned as a viable protection mechanism by several authorities and standards bodies for these specific use cases. Correct implementation will help keep regulatory scrutiny at bay.
- Hyperscaler cloud providers are increasingly offering options that allow virtualized confidential computing, which allows apps to run without recoding or refactoring.

Obstacles

- Complexity of the tech and lack of trained staff or understanding of best implementation methods may hinder adoption and/or weaken deployment (e.g., key management/handling is done incorrectly, unaddressed side channel vulnerabilities).
- Trust is slow to build and quick to evaporate, especially when confidential computing is paired with occasional hardware vulnerabilities.
- Some forms of confidential computing are not usually plug-and-play, and are currently mostly reserved for high-risk use cases such as machine learning. Varying by vendor and technology, you may require a high level of effort but see only marginal security improvement over more pedestrian controls like Transport Layer Security (TLS), multifactor authentication (MFA), and customer-controlled key management services.
- Offerings directly from CSPs vary greatly in robustness, performance and reliability. Not all named confidential computing offers similar protection.
- Confidential computing that leverages cloud-native key management (KM) may be at risk of inadequate privacy because the CSP manages and has access to the keys. Therefore, using third-party KMaaS becomes more important.
- Confidential computing currently only integrates with the client's deployed technology. It is rare to find SaaS or BPaaS vendors offering integration to confidential computing — hence reducing protection choices.

### User Recommendations

- Design (or duplicate) a sample application using one of the available abstraction mechanisms and deploy it into an instance with an enclave. Perform processing on datasets that represent the kinds and amounts of sensitive information you expect in real production workloads. This way, you can determine whether confidential computing affects application performance and seek ways to minimize negative results.
- Examine confidential computing for projects in which multiple parties, who might not necessarily trust each other, need to process (but not access) sensitive data in a way that all parties benefit from the common results. None of the parties should control the TEE in this scenario.
- Look for vendors that enable integration to a broader enterprise key management system and complementary encryption and PEC techniques.

### Sample Vendors

Alibaba Cloud; Anjuna Security; Fortanix; Google; IBM; Intel; Microsoft; VectorZero

### Gartner Recommended Reading

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

Achieving Data Security Through Privacy-Enhanced Computation Techniques

Solution Criteria for Cloud Integrated IaaS and PaaS

Securing the Data and Advanced Analytics Pipeline

How to Make Cloud More Secure Than Your Own Data Center

## Data Security Platforms

**Analysis By:** Joerg Fritsch, Brian Lowans

**Benefit Rating:** High

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

### Definition:

Data security platforms (DSPs) aggregate data protection requirements across data types, storage silos and ecosystems, starting with data discovery and classification. DSPs typically protect data by using policy-based authorization controls — for example, database APIs, dynamic data masking, format-preserving encryption (FPE) or tokenization. Select DSPs also undertake data activity monitoring or perform data risk assessments.

### Why This Is Important

Traditionally, data security has been delivered by disparate products, which has resulted in operational inefficiencies and an inability to support, for example, data risk assessments, open data, commercial data, and internal innovations and collaboration involving data. Especially in cloud-based data stores, a DSP reduces integration cost, manual work and friction by connecting previously disparate data security controls and capabilities.

### Business Impact

A DSP significantly increases visibility of, and control over, data and its broad usage — for example, in relation to unknown behaviors, not just narrower, privacy-related compliance goals. It therefore puts organizations in a position to truly secure their data. The increase in visibility and control enables secure data flows between individuals, organizations and governments. Consequently, wiser decisions can be made and better outcomes can be delivered for both business and society as a whole.

### Drivers

DSPs aggregate individually mature technologies, but the integration of these technologies is immature and still emerging. Three trends relating to data security, privacy and advanced analytics are driving DSP adoption:

- Enterprise agendas for strengthened data security and privacy are increasingly in competition in terms of their approach to DevSecOps, open data regulations and advanced analytics.
- DSP supports data products. “Data products” are top of mind for mature clients. This presents a paradigm shift, where the goal is to support higher utilization of “secure data” by making it easier for a diverse set of internal and external consumers to share and use data.

- Organizational data is increasingly distributed across different service and trust boundaries, frequently outside traditional on-premises data centers. Data is likely to be processed and stored in public cloud services of all types — infrastructure-based, platform-based and SaaS. This situation requires organizations to manage their data security far more effectively.

### Obstacles

- In the process of creating DSPs, vendors are making acquisitions, forging partnerships and introducing new capabilities — for example, data security posture management (DSPM). This makes it harder for customers to reach sound purchasing decisions, due to the risk that key vendors and products may change their identity or focus.
- Newer DSP products with large coverage (or broad spectrum) are focusing solely on cloud-based data stores and have no offering for on-premises data stores.
- Integration of formerly disparate technologies is partially immature and still emerging — as is the integration with data catalogs.

### User Recommendations

- Prioritize DSP investments by assessing DSP for all new data security projects — for example, the transition to data mesh architectures and cloud-based data lakes.
- Choose products that can provide a well-integrated, broad spectrum of controls combining data stewardship, policies and late binding access controls. Popular late binding access controls used by DSPs are cryptographic transforms, such as tokenization and format preserving encryption (FPE), dynamic data masking (DDM), and proprietary connectors and agents.
- Prioritize a broad spectrum DSP that implements data security controls required for compliance and risk mitigation in your IT environment. For example, select a DSP that uses tokenization as standardized access control rather than a DSP that uses diverse access controls depending on the data store.
- Be prepared to have coverage gaps or partially resolved data security issues that remain for a longer time as DSPs are still evolving.

### Sample Vendors

comforte AG; IBM; Immuta; Microsoft; Prime Factors; Privacera, Satori; SecuPi; TrustLogix; Velotix

### Gartner Recommended Reading

Market Guide for Data Masking

A Data Risk Assessment Is the Foundation of Data Security Governance

Market Guide for Data Loss Prevention

Use Enterprise Key Management to Provide Stronger Data Security and Privacy

## Postquantum Cryptography

**Analysis By:** Mark Horvath, Matthew Brisse

**Benefit Rating:** High

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Embryonic

### Definition:

Postquantum cryptography (PQC), also called quantum-safe cryptography, are algorithms designed to secure against both classical and quantum-computing attacks. PQC will replace existing asymmetric encryption, which will weaken over the next decade, deprecating existing classical encryption methodologies and processes.

### Why This Is Important

PQC offers organizations a higher level of cryptographic protection, which will remain strong as quantum computers enter the mainstream.

Existing asymmetric algorithms like Diffie-Hellman, RSA and ECC are vulnerable and will be unsafe to use by the end of the current decade, requiring replacement for common cryptographic functions, such as digital signatures, public key encryption and key exchanges.

### Business Impact

PQC has the following impacts:

- With the advent of stronger quantum computers, existing asymmetric algorithms must be replaced with quantum-safe ones. This includes all network, file and data encryption, IAM, as well as any other uses of asymmetric cryptography.
- There are no drop-in alternatives for existing cryptographic algorithms, leading to discovery, categorization and reimplementation efforts.
- As new algorithms have different performance characteristics, current applications must be retested and, in some cases, rewritten.

### Drivers

- Existing asymmetric encryption algorithms will become vulnerable to quantum decryption attacks by the end of the decade, potentially requiring reencryption of all data where the risk of exposure of the symmetric keys or tokens is considered important.
- Governments around the world are preparing and issuing mandates and legal frameworks requiring government agencies and enterprises to start devising PQC strategies. For example, in the U.S., Quantum Computing Cybersecurity bill requires owners and operators of national security systems and organizations supplying to the U.S. government to start using postquantum algorithms.
- “Harvest now and decrypt later” attacks are an ongoing concern, leading to the urgency to implement PQC security measures sooner rather than later.
- Secondary uses of new encryption (e.g., homomorphic encryption, stateful signatures, etc.) will offer new business opportunities beyond data protection.
- Once PQC is adopted by an organization, data should be secure for the foreseeable future.

Obstacles

- Most organizations don't know how cryptography functions within their organization, where keys and algorithms are used, or how secrets are stored and managed. Swapping them out for new algorithms will be challenging.

- Encrypted file sizes and digital signatures for new algorithms are typically much bigger than existing equivalents, necessitating hardware and network infrastructure improvements.

- New PQC algorithms will require new standards. The current set of PQC candidates' standards are expected to be released in late 2023 or early 2024, while fresh algorithm development will continue for the rest of the decade, affecting hardware, firmware, software and credentials used along with supported algorithms.

- Most vendors are typically unprepared when it's time to upgrade the cryptography and often require some pushing from their clients to recognize the demand.

- Some very important protocols lack built-in crypto-agility. For instance, no one is developing plans on how to incorporate new algorithms into WS-Security, which is used to safeguard SOAP APIs (a crucial type of API for all financial transactions).

### User Recommendations

- Build a cryptographic metadata database of all in-use cryptographic algorithms.
- Develop crypto policies for easing the transition to new algorithms.
- Perform an exercise for data identifying the expected end-of-life targets in the short, medium and long-term time scales, and create a key life cycle policy to reflect risks to asymmetric and symmetric crypto keys.
- Create a transition phase plan identifying which algorithms are suitable for particular use cases.
- Implement transitional crypto policies for when algorithms should be replaced and which new algorithms should be used in each use case.
- Implement crypto-agile application development and stage to production after extensive testing.
- Vet and test new PQC algorithms to understand their characteristics and uses.
- Implement crypto-agility initiatives with an object-based approach to address future changes in PQC algorithm updates and replacement.
- Prioritize business impact potential when selecting potential PQC use cases.

### Sample Vendors

Amazon; Google; IBM; ISARA; Microsoft; Qrypt; Quantum Xchange; SandboxAQ

### Gartner Recommended Reading

Preparing for the Quantum World With Crypto-Agility

Emerging Tech: How to Make Money From Quantum Computing

Emerging Tech: Critical Insights on Quantum Computing

Infographic: How Use Cases Are Developed and Executed on a Quantum Computer

## FinDRA

**Analysis By:** Brian Lowans

**Benefit Rating:** Transformational

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

### Definition:

A financial data risk assessment (FinDRA) prioritizes financial investment decisions in data security and privacy. The aim of a FinDRA assessment is to mitigate business risks to a level that is acceptable to balance business outcomes. FinDRA achieves this prioritization by analyzing the financial impacts of business risks identified through data security governance (DSG) and data security and privacy risk assessments.

### Why This Is Important

FinDRA enables organizations to translate the language of data security and privacy risks into the language of business risks and business outcomes.

### Business Impact

FinDRA creates the basis for establishing a budget for investment in data security, by focusing on how data security and privacy risks affect business outcomes in financial terms.

### Drivers

- Organizations are exploiting a variety of data and analytics assets, leading to prolific growth of data and business risks. However, they rarely analyze the financial impacts of data security or privacy risks that can result from investment decisions.

- FinDRA is creating an opportunity to understand how financial impacts may emerge from the opportunity costs, waste and risks associated with each dataset.

- Organizations need the ability to assess how data security and privacy risks associated with security incidents can create financial impacts to a business. Examples of data security and privacy risks associated with security incidents include data breaches, privacy enforcement, noncompliance and even accidental processing. This assessment capability requires a process to translate the issues of data risks into the language of business and financial impacts.

### Obstacles

- Most organizations continue to separate the decision processes and responsibilities for data monetization investments from the investment decisions associated with data security and privacy.
- Security perspectives are normally not included in business decisions to create, collect, use or share data. Therefore, they do not have the opportunity to translate the impacts of disparate security and privacy risks into business outcomes.
- Business leaders do not understand the language of data security and privacy risks and will frequently prioritize business access to data over data-security-led controls.

### User Recommendations

- Establish a DSG framework to identify all essential datasets and associated data owners, and identify and prioritize the mitigation of business risks through a set of data security policies.
- Conduct a data risk assessment (DRA) to identify how data security or privacy technologies apply security controls to each dataset. Use the DRA to analyze gaps and inconsistencies in data security policy implementation, and to gauge policy effectiveness when mitigating business risks.
- Work with business and financial leaders to evaluate how the business risks for each dataset affect business outcomes that may be measured in financial terms. Examples of these metrics can be revenue, ROI and opportunity cost.
- Use FinDRA to establish business support and to identify which business risks should be prioritized for data security investment.
- Use FinDRA to help reprioritize the data security budget as the risk assessments and IT architecture evolve.

### Gartner Recommended Reading

Use the Data Security Governance Framework to Balance Business Needs and Risks

4 Critical Steps to Accelerate the Adoption of Data Security Governance

A Data Risk Assessment Is the Foundation of Data Security Governance

Develop a Financial Risk Assessment for Data Using Infonomics

## Data Security Governance

**Analysis By:** Brian Lowans, Andrew Bales, Joerg Fritsch, Bart Willemsen

**Benefit Rating:** Transformational

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Adolescent

### Definition:

Data security governance (DSG) enables the assessment and prioritization of business risks, caused by data security, privacy, and compliance issues. This enables organizations to establish data security policies that support business outcomes and balance business needs against associated business risks. These risks arise from security, data residency and privacy issues, as data is processed across ecosystems or shared with partners.

### Why This Is Important

DSG enables the assessment, prioritization and mitigation of business risks caused by security, privacy, and other compliance issues, as data proliferates across on-premises and multicloud architectures. DSG establishes a balance between business priorities and risk mitigation through data security policies that can be applied across the whole IT architecture.

### Business Impact

DSG offers a balanced approach to define how data is accessed and used to support business performance objectives and client experience, while enforcing appropriate data security and privacy controls to mitigate risks. DSG requires collaboration among chief information security officers (CISOs), chief data and analytics officers (CDAOs), and business leaders, through a data security steering committee (DSSC). This would help break down communication barriers and contribute toward business outcomes.

### Drivers

- It is essential to use DSG as a continuous process to manage, assess and prioritize business risks, and create focused data security policies that can mitigate those risks.
- Data security policies are needed to guide the implementation of consistent data access privileges security controls across a portfolio of datasets.
- DSG must be leveraged to address internal and external requirements to manage user access privileges to each dataset in terms of privacy, confidentiality, integrity, availability, business purpose, and life cycle risks.
- Organizations need to develop processes to create and orchestrate data security policies across multiple independent data security and identity access management (IAM) products, to minimize data security policy gaps and inconsistencies.
- No single product mitigates business risk sufficiently, emphasizing the need for centralized creation and coordination of data security policies.
- It is essential to leverage adequate privacy impact assessments (PIA) through DSG to mitigate data residency and sovereignty risks.

### Obstacles

- Business stakeholders have fragmented responsibilities for managing data. Unless they create data security policies together with DSG, they will fail to balance business outcomes and risk mitigation.
- The deployment of data security, IAM and application products are purchased and managed by different leaders.
- Each product applies independent security controls, as IAM products do not control access to data. Data security products often operate on either unstructured or structured data, apply controls to specific platforms, and use custom data discovery technology. This reduces the effectiveness of DSG because it is not possible to deploy consistent data security policies.
- The security team must orchestrate data security policies manually across the portfolio of available security product controls. This also requires regular data risk assessments (DRA) to assess gaps and inconsistencies that need to be reported as stronger business risks, or to support new policies or product deployments.

### User Recommendations

- Use DSG to create and manage consistent data security policies across your portfolio of datasets, according to the level of business risks defined.
- Use DSG to analyze business risks and their impacts due to specific security monetization choices, by using infonomics to evaluate the financial impacts on business outcomes.
- Use principles such as Gartner's financial data risk assessment (FinDRA) to establish prioritization of security investment options.
- Ensure cooperation and collaboration between the CDAO and the CISO, to reduce redundancy and waste in evaluating data management and security.
- Apply data security policies across all data security, IAM and application management products that interact with each dataset.
- Consider leveraging the DSSC to reach out to your CIO, CDAO or risk officer to extend your DSG operating model with connected governance. This would help with the most complex, cross-enterprise and geographic risk governance programs.

### Gartner Recommended Reading

Use the Data Security Governance Framework to Balance Business Needs and Risks

4 Critical Steps to Accelerate the Adoption of Data Security Governance

Use a Data Security Steering Committee to Realize Data Security Governance Objectives

A Data Risk Assessment Is the Foundation of Data Security Governance

3 Steps to Effectively Capture and Communicate the Business Value of Cybersecurity Initiatives

## Homomorphic Encryption

**Analysis By:** Mark Horvath, Bart Willemsen

**Benefit Rating:** Transformational

**Market Penetration:** Less than 1% of target audience

**Maturity:** Emerging

### Definition:

Homomorphic encryption (HE) uses algorithms to enable computations with encrypted data. Partial HE (PHE) supports only limited use cases, such as subtraction and addition, but with little performance impact. Fully homomorphic encryption (FHE) supports a wider range of repeatable and arbitrary mathematical operations; however, it worsens performance.

### Why This Is Important

HE offers an unparalleled advance in privacy and confidential data processing, although this is largely at the database level. Benefits include the ability to:

- Perform analytics on data while in an encrypted state, so that the processor never sees the data in the clear, yet delivers accurate results.
- Share and pool data among competitors.
- Share all or part of users' data, while protecting their privacy.
- Systems based on lattice encryption, which are quantum-safe.

### Business Impact

Even in restricted form (PHE), HE enables businesses to use data, send it to others for processing and return accurate results, without fear it will be lost, compromised or stolen. Any data intercepted by a malicious actor is encrypted and unreadable, even by the coming generation of quantum computers.

Applications include:

- Encrypted search
- Data analytics
- Machine learning (ML) model training
- Multiparty computing
- Securing, long-term record storage, without concerns about unauthorized decryption

### Drivers

- The enhanced enforcement of data residency restrictions worldwide is forcing organizations to protect data in use, rather than only when it is in transit or at rest.
- Globally maturing privacy and data protection legislative frameworks demand that more-precise attention be paid to sensitive data. As a result, data pooling, sharing and cross-entity analysis use cases increasingly benefit from forward-looking and sustainable technologies, such as HE.
- Aside from primarily financial use cases (e.g., cross-entity fraud analytics), other industries can benefit as well. One example is the healthcare industry, where analysis of sensitive data across various entities happens often with data protected while in use.
- Solving issues of trust and cooperation with secure multiparty computation (sMPC) will benefit internal and external protection of data.
- The oncoming availability of quantum computing (QC), as highlighted by NIST and the Canadian Forum for Digital Infrastructure Resilience, threatens to compromise the confidentiality of almost all data. This includes digital communication previously considered protected by conventional cryptography. For example, there are signals that malicious actors may retain exfiltrated encrypted data in expectation of the ability to decrypt it years later and re-engage with victims for extortion and ransom demands. Timely adoption of HE in data protection will sustainably protect data, even when previously compromised in (conventionally) encrypted form.

### Obstacles

- The application of various forms of HE to daily use cases leads to a degree of complexity, slows operations and requires highly specialized staff.
- The market's unfamiliarity with this technology stands in the way of speedy adoption.
- Although PHE can be a Turing-complete implementation, which means an arbitrary set of instructions could be executed, no vendor has a robust implementation that exploits this capability.
- Some scenarios will never be a good match for HE — for example, those that require security in components beyond analytics and processing, such as production databases and proprietary algorithms.

### User Recommendations

- Brainstorm opportunities with your technical and executive teams. For example, come up with a list of five to 10 use cases for HE to improve the adoption of core solutions.
- Treat potential HE projects as experiments, keeping in mind the early stage of the technology's development and the significantly not-real-time nature of HE products. Consider these experiments proofs of concept (PoCs) to build experience, until the technology matures.
- Continue with existing security controls. HE does not necessarily negate the need for other security controls, observance of data residency requirements or access control.
- Assess the core benefits of using HE in combination with other quantum-safe or privacy-enhancing computation techniques.
- Integrate in-use protection via forms of HE into messaging and third-party analytics services.
- Assess the merits of piloting HE by using a vendor's solution, which could offer functionality without the time investment associated with a custom solution.

### Sample Vendors

CryptoLab; Duality; Enveil; IBM; Inpher; IXUP; LiveRamp; Lorica; Ziroh Labs

### Gartner Recommended Reading

Emerging Technologies: Homomorphic Encryption for Data Sharing With Privacy

Emerging Technologies and Trends Impact Radar: Security

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

What Executives Need to Do to Support the Responsible Use of AI

Achieving Data Security Through Privacy-Enhanced Computation Techniques

## Sovereign Data Strategies

**Analysis By:** Andrew White

**Benefit Rating:** High

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

### Definition:

Sovereign data strategies are state-controlled efforts to control data about its citizens and economy. Such strategies include regulations for privacy, security, access, use, retention and sharing, as well as processing and persistence.

### Why This Is Important

Sovereign data strategies will impact every organization that does business across notable sovereign jurisdictions. They represent the coordinated approach from a sovereign entity to control and regulate how data is used in and across its economy and society. As such, these strategies and associated policies will impact data privacy, access, security, sharing, analysis, processing, storage and value.

### Business Impact

Sovereign data strategies will impact where data about citizens and businesses can be stored, accessed, processed and used. Tracking such policies is hard and increasingly a challenge. The impact on organizations that trade or operate across sovereign boundaries can be as excessive as requiring you to replicate your entire operational business to exist and stay within certain sovereign boundaries. All storage and processing of such data would need to remain local and could not take advantage of the scale and synergy of global public cloud services.

### Drivers

We won’t analyze the politics driving sovereign motivations, but we can comment on the resulting policy or implied actions referenced by the strategies:

- Many sovereign data strategies refer to failures in public markets, public data sharing efforts, intellectual property use, and state or citizen personal data protection. The EU is a good example (see EU Data Strategy) as it seeks to create internal EU data markets to force (and reward) private firms to share data concerning the health of EU citizens. This data is to help solve health issues that otherwise would not be solved, according to the EU guidelines. Thus, the value in data will be exposed and increasingly shared.

- The U.S. (see Federal Data Strategy), China and several other states fall back on security concerns, ethics and the general well-being of the state as reasons for their sundry data strategies. Data privacy and data security are two very large drivers in this category; driving improved economic growth through better data sharing across public and private sector organizations is also referenced a lot. Together all sovereign strategies may use the same words but the motivations behind the policies and the focus tend to conflict and overlap, all at the same time.

- In 2023, there is growing recognition that geopolitical risk is elevating what started out as a set of policies and roles focused on data, to something far more visible. The shift is effectively from sovereign data strategies to sovereign digital strategies. Sovereign digital strategies describe how each seeks to govern and control access to their respective digital economy. As such, what started out as important to the CDAO and maybe the CIO is now important to the CEO and board of directors (see Choose the Optimal Corporate Structure to Cope With Geopolitical Risks).

### Obstacles

- You may not always be able to order up a copy of your local sovereign data or digital strategy. Individual policies are observable even if the coordinated effort by the sovereign across all policies is not visible.

- Responding one policy at a time may well consume all your resources such that you run out of time and money should a clear sovereign strategy appear, and a coordinated response then becomes clear.

- Inability of CDAOs and other executive leaders, including CISO, CIO, and COO/CEO, along with Legal and Risk, can lead to inefficient and slow responses to the growing challenges.

### User Recommendations

- Ensure your data and analytics (D&A) and digital strategy takes account of the constraints for those regions you operate in.
- Adopt a risk stance. Focus on mitigating risks as you develop global business strategies or seek to grow new businesses or services across sovereign boundaries. For example, invest tactically in privacy, security, and data sharing or processing efforts only when (and if) you have confidence in the reliability and enforcement of your targeted sovereign data strategy.
- Consider scenario planning to emulate costs and operational changes needed for extreme situations. Consider, for example, the need to air gap your entire IT landscape, or how you may advise the CEO/CFO on how data will be governed should parts of the business have to be jettisoned in worse-case scenarios.
- Don't panic or rush, as benefits may outweigh risks in a couple of years once sovereign data strategies stabilize and become more robust. At the end of the day, what was meant to be a global business environment and global cloud will become a fractured and distinct set of business environments and clouds.

### Gartner Recommended Reading

Choose the Optimal Corporate Structure to Cope With Geopolitical Risks

Trends 2023: Rise and Risks From EU, U.S., China and Other Sovereign Data Strategies and Policies

Overcome the Challenges of Cross-Border Data Communication Over Internet With China

How New Privacy Laws in California and China Mirror the GDPR

Client Questions Video: How to Deal With EU-U.S. Personal Data Transfers (With Privacy Shield Gone)?

## At the Peak

### Quantum Key Distribution

**Analysis By:** Mark Horvath, Brian Lowans

**Benefit Rating:** Low

**Market Penetration:** Less than 1% of target audience

**Maturity:** Embryonic

**Definition:**

Quantum key distribution (QKD) is a tamper-evident communication method that implements a cryptographic protocol for transporting keys based on quantum mechanics. It enables two parties to produce a shared secret known only to them, which then can be used to encrypt and decrypt messages. These keys are created and exchanged in such a way that the system can detect the interference of any third party trying to gain knowledge of the key, and terminate the link.

#### Why This Is Important

QKD is an important, early-stage technology for creating, moving and preserving the entanglement of two or more quantum particles. QKD claims to provide a tamper-evident channel for data exchange by maintaining the quantum entanglement of particles (usually photons) as they transfer data (e.g., keys) between systems. The nature of entanglement is such that any disturbance will result in automatic channel collapse, destroying the key and preventing the transmission of further data.

#### Business Impact

- QKD's main property, tamper evidence, is needed by companies that need to exchange high-value encryption keys in as secure a way as possible.
- QKD provides a demonstrably secure channel that allows high-value data to be transmitted without compromising the integrity of the data.
- A quantum random number generator (QRND) can generate cryptographically better keys than classical sources of entropy.
- Governments and military departments have a need for extremely secure methods of key exchange, which are becoming increasingly commercialized.

### Drivers

- As quantum computers become more realistic, organizations are looking to move to “quantum-safe” technologies, sparking renewed interest in QKD.
- QKD is a crucial, foundational technology for quantum networks and quantum information science.
- QKD is the foundational technology for the next generation of secure satellite-based networking.

### Obstacles

- While QKD is sometimes advertised as “quantum safe,” keys or other data transmitted through a QKD channel will then rely upon traditional key management life cycle technologies.
- QKD is a channel with relatively low bandwidth, when compared to classical cryptographic channels like Transport Layer Security (TLS).
- Currently, QKD channels cannot be boosted and cannot use repeaters without breaking the entanglement, making large scale routing difficult. The technology is currently limited to a few hundred kilometers when routed through existing fiber optics.
- Existing classical systems for key exchange typically provide a high enough level of trust for most purposes, especially when using quantum-safe cryptographic keys.

### User Recommendations

Security and risk management leaders evaluating QKD should:

- Evaluate if QKD is needed as part of their plan for moving to postquantum networking.
- Identify high-value or sensitive data that could benefit from tamper-evident key exchange.

### Sample Vendors

AUREA Technology; ID Quantique; iXblue; MagiQ Technologies; Qasky; QNu Labs; QuantumCTek; Thorlabs; Toshiba

### Gartner Recommended Reading

## Differential Privacy

**Analysis By:** Bart Willemsen

**Benefit Rating:** Moderate

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

### Definition:

Differential privacy is an approach to using or sharing data while withholding or distorting certain elements about individual records in the dataset. It uses exact mathematical algorithms that randomly insert noise into the data, and add parameters for distinguishability, closeness and diversity of outcomes at each query. When applied correctly, this prevents the disclosure of identifiable information while ensuring that the resulting analysis does not significantly change informationwise.

### Why This Is Important

Concerns continue to exist about privacy and the use of personal data in algorithms to serve content or personalize recommendations. As regulatory measures are employed to prevent unauthorized use of personal data, businesses are looking for ways to protect personally identifiable information while still using the data. One technology that can be deployed to accomplish this is differential privacy.

### Business Impact

Business data holds value and much of it is personal data. Regulations that constrain the use of personal data are increasing, and the liability for misusing personal data can be substantial. Businesses need to ensure their reputation reflects a company that protects customer data. There are many techniques to address problems in preserving privacy when training AI models. Differential privacy ensures the privacy of individual rows of data while supporting meaningful analysis of aggregate data.

Drivers

- Differential privacy helps to not only reduce risk but also unlock data for AI that was previously too difficult to access.
- Businesses need to uncover value from data without crossing the boundaries of ethical or regulatory restrictions on the use of personal data.
- It is increasingly likely that more restrictive regulations will be enacted, including on the use of personal data in training of algorithms and on how algorithms handle personal data in turn.
- The risk from sophisticated, state-sponsored bad actors that target theft of personal information to facilitate fraudulent actions, remains on the rise.
- Business reputations and trust can be significantly damaged by information breach or misuse.
- Exposure is not limited to datasets in control of the business, as malicious actors can increasingly combine data sources to reidentify individuals even if the data used by the business is anonymized.
- With differential privacy, source data is not altered because the answer to each query is treated “on the fly,” protecting the data in use while retaining source data integrity.
- With differential privacy, information value is maintained in a controllable manner via a privacy budget, delivering the desired level of anonymity.
- Some providers have started to add collaborative differential privacy capabilities in their offerings for further privacy protections.

### Obstacles

- Solutions that reference the use of differential privacy are not always comparable or equally easily implemented.

- Privacy protection solutions use a variety of techniques and they vary in effectiveness. Organizations often lack a framework to consistently determine the appropriate approach based on use-case requirements, technology maturity and fit.

- Most tools cover anonymity in different degrees and focus on the extent to which reidentification can occur. Other deployments add measures to diversity and closeness of outcome, apart from reidentification protection. This can cause confusion in comparison.

- Lack of familiarity with differential privacy — and the skilled staff to effectively deploy and manage it — hinders adoption. This is exacerbated by how jurisdictions define and determine “anonymous” versus “pseudonymous” data differently.

- Lack of transparency around setting of the privacy budget (the extent to which controls are implemented) undermines trust, whereas increased transparency could elevate trust.

### User Recommendations

- Explore the use of differential privacy techniques to decrease the likelihood of sensitive data exposure.

- Use a privacy impact or data protection impact assessment to establish whether additional means are necessary and relevant to the use case.

- Compare differential privacy with other privacy-enhancing computation techniques when operating in high-performance environments that require a high level of precision in analytics models.

- Prioritize differential privacy techniques if you’re operating in a highly regulated industry, such as financial services or healthcare.

- Explore differential privacy techniques when using data across regions where privacy regulations may vary, and always be transparent about where you have set the privacy budget.

### Sample Vendors

Immuta; LeapYear; LiveRamp; PHEMI Systems; Privitar; Tumult Labs

### Gartner Recommended Reading

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

## Synthetic Data

**Analysis By:** Arun Chandrasekaran, Anthony Mullen, Alys Woodward

**Benefit Rating:** High

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

### Definition:

Synthetic data is a class of data that is artificially generated rather than obtained from direct observations of the real world. Synthetic data is used as a proxy for real data in a wide variety of use cases including data anonymization, AI and machine learning development, data sharing and data monetization.

### Why This Is Important

A major problem with AI development today is the burden involved in obtaining real-world data and labeling it. This time-consuming and expensive task can be remedied with synthetic data. Additionally, for specific use-cases like training models for autonomous vehicles, collecting real data for 100% coverage of edge cases is practically impossible. Furthermore, synthetic data can be generated without personally identifiable information (PII) or protected health information (PHI), making it a valuable technology for privacy preservation.

### Business Impact

Adoption is increasing across various industries. Gartner predicts a massive increase in adoption as synthetic data:

- Avoids using PII when training machine learning (ML) models via synthetic variations of original data or synthetic replacement of parts of data.
- Reduces cost and saves time in ML development.
- Improves ML performance as more training data leads to better outcomes.

- Enables organizations to pursue new use cases for which very little real data is available.
- Is capable of addressing fairness issues more efficiently.

Drivers

- In healthcare and finance, buyer interest is growing as synthetic tabular data can be used to preserve privacy in AI training data.
- To meet increasing demand for synthetic data for natural language automation training, especially for chatbots and speech applications, new and existing vendors are bringing offerings to market. This is expanding the vendor landscape and driving synthetic data adoption.
- Synthetic data applications have expanded beyond automotive and computer vision use cases to include data monetization, external analytics support, platform evaluation and the development of test data.
- Increasing adoption of AI simulation techniques is accelerating synthetic data.
- There is an expansion to other data types. While tabular, image, video, text and speech applications are common, R&D labs are expanding the concept of synthetic data to graphs. Synthetically generated graphs will resemble, but not overlap the original. As organizations begin to use graph technology more, we expect this method to mature and drive adoption.
- The explosion of innovation in AI foundation models is boosting synthetic data creation. These models are becoming more accessible and more accurate.

### Obstacles

- Synthetic data can have bias problems, miss natural anomalies, be complicated to develop, or not contribute any new information to existing, real-world data.

- Data quality is tied to the model that develops the data.

- Synthetic data generation methodologies lack standardization.

- Completeness and realism are highly subjective with synthetic data.

- Buyers are still confused over when and how to use the technology due to lack of skills.

- Synthetic data can still reveal a lot of sensitive details about an organization, so security is a concern. An ML model could be reverse-engineered via active learning. With active learning, a learning algorithm can interactively query a user (or other information sources) to label new data points with the desired outputs, meaning learning algorithms can actively query the user or teacher for labels.

- If fringe or edge cases are not part of the seed dataset, they will not be synthetized. This means the handling of such borderline cases must be carefully accommodated.

- There may be a level of user skepticism as data may be perceived to be “inferior” or “fake.”

### User Recommendations

- Identify areas in your organization where data is missing, incomplete or expensive to obtain, and is thus currently blocking AI initiatives. In regulated industries, such as healthcare or finance, exercise caution and adhere to rules.

- Use synthetic variations of the original data, or synthetic replacement of parts of data, when personal data is required but data privacy is a requirement.

- Educate internal stakeholders through training programs on the benefits and limitations of synthetic data and institute guardrails to mitigate challenges such as user skepticism and inadequate data validation.

- Measure and communicate the business value, success and failure stories of synthetic data initiatives.

### Sample Vendors

Anonos (Statice); Datagen; Diveplane; Gretel; Hazy; MOSTLY AI; Neuromation; Rendered.ai; Tonic.ai; YData

### Gartner Recommended Reading

Innovation Insight for Synthetic Data

Innovation Insight for Generative AI

Data Science and Machine Learning Trends You Can't Ignore

Cool Vendors in Data-Centric AI

Case Study: Enable Business-Led Innovation with Synthetic Data (Fidelity International)

## Digital Communications Governance

**Analysis By:** Michael Hoeck

**Benefit Rating:** High

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Early mainstream

### Definition:

Digital communications governance (DCG) solutions provide methods to monitor and enforce corporate governance and regulatory compliance across a growing number of communications tools available to employees. For the various communications tools in use, DCG solutions enable consistent policy management, enforcement and reporting capabilities such as data retention, surveillance, supervision, behavioral analytics, auditing, and e-discovery.

### Why This Is Important

DCG solutions are critical to an organizations' efforts to meet a growing number of regulatory and organizational compliance requirements. They facilitate the collection of multiple communication channels to retain them, accurately classify the data, consistently apply retention policies, improve timely response to audits and e-discovery requests, surface organization insights, and help effectuate data use policies.

### Business Impact

DCG solutions help manage regulatory and organizational use policies for the growing volumes and types of digital communications. They enrich communications data for behavioral analytics to surface insights, such as employee sentiment, misconduct risks, and industry-specific conduct assessments. DCG improves e-discovery efforts supporting legal hold, search, review and export requirements. It is used for consistent application and use of retention policies across various communication channels.

### Drivers

DCG solutions enable businesses and organizations to utilize information within digital communications data sources to:

- Comply with regulatory requirements for highly regulated industries such as financial services, healthcare and public sector.
- Consolidate, centralize and simplify access to the multiple channels of digital communications in use by employees, including text-, voice-, and video-based content.
- Assess effectiveness of compliant-use policies for digital communications tools.
- Identify behavioral attributes and business intelligence within digital communications, such as employee sentiment and conduct and risk assessment, based on the use of sentiment analysis and data models.
- Capture and retain mobile device communications, including SMS/Multimedia Messaging Service (MMS) and other mobile application messages, such as WhatsApp, WeChat, and Signal.
- Respond to public records requests, such as Freedom of Information Act (FOIA) and Public Records Act (PRA).
- Simplify e-discovery and access across multiple communication channels driven by litigation, audits, internal matters, and other investigations.
- Respond to subject rights requests of privacy regulations, such as General Data Protection Regulation (GDPR) and California Privacy Rights Act (CPRA).

### Obstacles

- DCG solutions may require use of multiple vendor offerings to connect the variety of communication channels with the appropriate archive offering.
- DCG requires an organizationwide strategy of supervising or surveilling employee communications, which may run into barriers of adoption based on a “big brother” perception.
- Executive sponsorship and stakeholder buy-in to data classification and related retention policies is critical for successful deployments, but may be difficult to obtain.
- Ensuring clear understanding of organizations’ data residency requirements for archived data and proper alignment to vendor solutions can be challenging.
- Transitioning or migrating existing data, including previous archives, to a new archive can be time-consuming, complex and costly.

### User Recommendations

- Mitigate potential compliance and regulatory violations by shifting from a reactive to a proactive posture using DCG solutions.
- Shortlist solutions that best align scope of digital communications sources to required use cases, such as data retention, e-discovery, supervision, surveillance, and business intelligence.
- Differentiate vendor solutions by assessing use of AI/ML as a critical component of their offering’s use of analytics to automate and accelerate business processes.
- Improve e-discovery and supervision/surveillance outcomes by using solutions that reduce false positives.
- Negotiate exit strategy terms upfront to create transparency and minimize future costs for data export/extraction processing.
- Focus attention on SLAs, which obligate SaaS/platform as a service (PaaS) vendors to support the defined performance levels as the data volume grows.
- Scope migration of legacy communication archives, including export from old, import to new and ongoing storage during the selection phase.

### Sample Vendors

CellTrust; Global Relay Communications; Mimecast Services; Movius; Proofpoint; Shield; Smarsh; TeleMessage; Theta Lake; Veritas Technologies

### Gartner Recommended Reading

Magic Quadrant for Enterprise Information Archiving

Critical Capabilities for Enterprise Information Archiving

## Security Service Edge

**Analysis By:** Charlie Winckless, John Watts

**Benefit Rating:** Transformational

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

### Definition:

Security service edge (SSE) secures access to the web, cloud services and private applications. Capabilities include adaptive access control, data security, visibility and control. Further capabilities include an advanced threat defense and acceptable use control enforced by network-based and API-based integrations. SSE is primarily delivered as a cloud-based service and may include on-premises or agent-based components.

### Why This Is Important

SSE improves organizational flexibility to secure the usage of web and cloud services, and remote work. SSE offerings are the convergence of security functions (at least, secure web gateways [SWGs], cloud access security brokers [CASBs] and zero trust network access [ZTNA]) to reduce complexity and improve user experience. They are delivered from the cloud. When organizations are pursuing a secure access service edge (SASE) architecture, SSE is paired with software-defined WAN (SD-WAN) to simplify networking and security operations.

### Business Impact

Hybrid work is continuing to drive the adoption of public cloud services, especially of SaaS applications. Both hybrid work and the adoption of public cloud services remain business enablers for most Gartner clients. SSE allows the organization to support anytime-anywhere workers by using a cloud-centric approach to enforce a security policy when accessing the web, cloud services and private applications. Simultaneously, SSE reduces the administrative complexity of running multiple products.

Drivers

- Organizations need to secure user, application and enterprise data that is distributed, decentralized and requires secure remote access.

- For many enterprises, a significant amount of critical data is now hosted in SaaS. Therefore, there is a need to perform data loss prevention (DLP) on data that is located in, going to, and leaving these SaaS platforms.

- SSE enables flexible and primarily cloud-based security for users and devices without being tied to on-premises network infrastructure and connectivity. The same security outcome is delivered to users regardless of their location or connectivity.

- Administrators can have enhanced visibility on user traffic and a single configuration and monitoring location for this traffic.

- SSE allows organizations to implement a posture based on identity and context at the edge.

- By consolidating vendors, organizations reduce complexity, costs and the number of vendors used to enforce security policy. Using a single SSE platform rather than multiple point offerings, they can both reduce complexity and reduce gaps in security coverage.

- Sensitive data inspection and malware inspection can be done in parallel across all channels of access. SSE allows doing both inspections in parallel, leading to a better performance and more consistent configuration than doing them separately.

- An adaptive access can take into account more input signals and be more consistently enforced, regardless of the application location or type.

- Organizations look for deeper security capabilities when building a SASE architecture compared to vendors that may have a minimal set of security features as part of their SD-WAN offering.

- Tight integrations that exist between discrete SD-WAN and SSE vendors allow interoperability without requiring a single-vendor approach.

Obstacles

- As the market is being formed by the convergence of capabilities, vendors may be strong in certain capabilities and weak in others. Vendors may also lack overall tight integration between SSE capabilities or with SD-WAN vendors.
- Not all vendors provide sufficiently sensitive data identification and protection to manage business risks.
- Some vendors have focused less on SaaS security and integrations. However, businesses increasingly need this visibility and protection.
- Being cloud-centric, SSE typically doesn't address every need supported by on-premises controls such as internal firewalling.
- Organizations are concerned about uptime or availability of services that they depend on for their business. This is compounded by weak SLAs from some vendors.
- Not all vendors provide all features locally in all geographies, resulting in performance or availability issues.
- Switching costs from incumbent vendors or timing of contract expirations prohibit near-term consolidation.
- Migrating from a VPN will increase costs.

### User Recommendations

- Exploit the converged market, consolidate vendors, and cut complexity and costs as contracts renew for SWGs, CASBs and VPNs by replacing them with a ZTNA approach.
- Approach SSE consolidation identifying which elements you may already have in place (for example, existing cloud-based CASB or SWG). Develop a shortlist of vendors based on your use cases regarding secure end-user requirements, the cloud services you use, and the data you need to protect.
- Inventory your equipment and contracts to implement a multiyear phaseout of on-premises perimeter and branch security hardware in favor of the cloud-based delivery of SSE.
- Global enterprises should validate that remote offices have acceptable performance and features with selected vendors. Vendor point of presence (POP) locations and service support are key.
- Actively engage with initiatives for branch office transformation, SD-WAN and Multiprotocol Label Switching (MPLS) offload to integrate cloud-based SSE into the scope of project planning.

### Sample Vendors

Broadcom; Cisco; Cloudflare; Forcepoint; iboss; Lookout; Netskope; Palo Alto Networks; Skyhigh Security; Zscaler

### Gartner Recommended Reading

2022 Strategic Roadmap for SASE Convergence

Magic Quadrant for Security Service Edge

Critical Capabilities for Security Service Edge

Adopt Security Service Edge (SSE) to Replace Stand-Alone SWG, CASB and ZTNA Products

## DevOps Test Data Management

**Analysis By:** Andrew Bales

**Benefit Rating:** Moderate

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Early mainstream

## Definition:

DevOps test data management is the process of providing DevOps teams with sanitized data to evaluate the performance, functionality and security of applications. It typically includes copying production data, anonymization or masking, and sometimes, virtualization. In some cases, specialized techniques, such as synthetic data generation, are appropriate. Given potential compliance and privacy issues, the efforts frequently involve members of application and data security teams.

## Why This Is Important

Test data management is inconsistently adopted across organizations, with many teams still copying production data for use in test environments. As organizations shift to DevOps and the pace of development increases, this traditional approach is increasingly at odds with requirements for efficiency, privacy and security, and even the increased complexity of modern applications. This opens organizations to a variety of legal, security and operational risks.

## Business Impact

Quick provisioning of test data helps ensure the pace of development isn't slowed. It's also increasingly important to remain compliant with the growing number of privacy mandates to which organizations are subject. This helps avoid fines and remediation and mitigation costs, along with the inevitable delays associated with audits and investigations. Finally, by providing application teams with anonymized or synthetic data, the risk of data breaches is reduced.

Drivers

- Test data management is generally viewed as a mature, relatively uncomplicated, practice. However, the reality is the combination of the increased pace of development from DevOps and a growing number of privacy mandates and constraints have stressed traditional approaches — prompting the use of virtualization as well as alternative masking and protection techniques, including synthetic data generation.

- More traditional test data management has been inconsistently adopted, with many organizations either simply using copies of production data in unsafe environments or generating “dummy data” (distinct from emerging synthetic data generation techniques) that doesn’t accurately reflect production data. The data privacy requirements and complexity issues noted have prompted organizations to revisit and update their processes with an eye toward scalability and automation. Updated technologies may also be a requirement. For example, requirements for speed and agility have created a need for data virtualization tools.

- Data protection is cited by most Gartner clients in inquiries regarding test data management. Privacy and data protection requirements mean it’s no longer safe to simply provide development teams with a copy of production data since this practice leaves organizations open to increased risk of regulatory violations, data breaches and other security issues.

- With modern applications relying on an increasing number of interconnected data stores (many of which, technologically speaking, are vastly more complex) and applications and APIs to function, testing has become more complex. Such complexity demands that tools support the ability to coordinate and synchronize changes across different data stores to ensure relational consistency while still addressing security and speed mandates.

### Obstacles

- In the absence of a strong culture of security, processes and technologies to protect sensitive information during development and testing will encounter friction. Conflicting needs for rapid development and privacy require attention to a mix of organizational and cultural issues to strike a balance across groups.
- Responsibility for test data management in organizations has been shared by application development and database administration. New technologies and processes may shift those responsibilities to include security, complicating organizational dynamics and potentially creating the need for additional resource allocation.
- Implementation can be a burden, especially where little or no data sensitivity classification has been done. This must be accomplished before teams can proceed with the required data transformation and masking. These efforts are typically combined with an analysis of data relationships so that relational integrity can be assured.

### User Recommendations

- Involve stakeholders — such as data management, privacy and security teams, compliance teams, etc. — to understand consumption patterns and needs — as appropriate.
- Document existing test data management practices so tools and processes can be evaluated against data protection mandates.
- Coordinate with other teams to avoid duplication of effort and tooling since data masking tools may also be used by analytics teams (e.g., to provide data for machine learning or other purposes).
- Evaluate data masking tooling by considering support for databases and other stores, data discovery capabilities, types of masking supported, and the ability to coordinate change to ensure consistency (e.g., key fields) across multiple sources.
- Evaluate data virtualization for DevOps use cases where frequent updates to test data are required. Virtualization can speed up the process of providing copies of safe data.
- Determine whether synthetic data generation is appropriate in cases where suitable data doesn’t exist or reidentification risks are high.

### Sample Vendors

BMC Software; Broadcom; DATPROF Delphix; Hazy; Informatica; K2View; Mage; OpenText; Solix Technologies

### Gartner Recommended Reading

Market Guide for Data Masking

Elevating Test Data Management for DevOps

3 Steps to Improve Test Data Management for Software Engineering

Innovation Insight for Synthetic Data

## Sliding into the Trough

### Zero-Knowledge Proofs

**Analysis By:** Mark Horvath, Bart Willemsen

**Benefit Rating:** Moderate

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

#### Definition:

Zero-knowledge proofs (ZKPs) are privacy-preserving messaging protocols that enable entities to prove that information available to either or both of them is correct, without the requirement to transmit or share the underlying (identifiable or otherwise sensitive) data. ZKPs enable entities to prove information validity without the requirement to transmit personal or confidential data.

#### Why This Is Important

Following increasingly imminent digital threats and legislative data protection requirements, security and risk management (SRM) leaders must support use cases that enable digital business while ensuring in-use protection. These protocols limit the requirement for mass decryption/encryption of data elements, which benefits the efficiency of work — including potential adoption of blockchain-based systems.

#### Business Impact

ZKPs are being applied for many use cases, especially in the context of authentication and transaction verification. Other use cases include payments, decentralized identity, custody management, anti-money-laundering (AML), know your customer (KYC), consumer identity and access management (IAM), age verification, etc. With the addition of ZKPs to blockchain platforms, SRM leaders can cover information security use cases that require confidentiality, integrity and availability (CIA). Some blockchain platforms have evolved to include this.

#### Drivers

- Traditional data protection techniques typically focus on data in motion (i.e., transport layer security) and data-at-rest encryption. Data-at-rest encryption, as commonly implemented, does not provide strong protection from data theft and privacy disclosures. It is unable to secure data in use and data sharing scenarios.

- New use cases and maturing privacy legislation worldwide present new privacy and cybersecurity concerns that require data-in-use protection. There are also scenarios where the data itself does not need to be shared. ZKPs enable such data-in-use protection.
- Concerns about data security in several scenarios, including collecting and retaining sensitive personal information, processing personal information in external environments such as the cloud and information sharing.
- Privacy violations (due to the exposure of sensitive information).
- Need for mitigation of sensitive data leakage and cyberattacks.

### Obstacles

- Even with a variety of web applications (e.g., ZKProof), ZKPs remain in an emerging state. They still require a common framework for applications to leverage.
- Only a limited number of practical implementations have emerged to date.
- The variety of methodologies and the multiplicity of approaches to data management inhibit adoption. ZKPs will need to scale at the rate of blockchain transactional volumes to be effective.
- ZKPs require integration into applications. Downstream applications, such as CRMs and databases, will need some modification.
- Some ZKPs, like ZK-SNARK, have a dependency on existing encryption/hashes (ECDSA in this case) as part of their implementation. This adds a potential complexity in upgrading them to quantum-safe protocols and limits available staff/experts.

### User Recommendations

- Work with SRM leaders to gain a deeper understanding of the nature of these controls, understanding that ZKP techniques are a paradigm shift.
- Be realistic with the current immaturity of ZKP solutions and approaches when evaluating ZKP benefits for privacy protection.
- Evaluate how ZKP controls may impact transaction authentication and, ultimately, consumers.
- Assess the impact on the broader information management strategy.

- Assess the architectural implications for using ZKP with different blockchains and distributed ledgers.

### Sample Vendors

DropSecure; Evernym; IBM; Ligero; Microsoft; Ping Identity; QEDIT; Sedicii; StarkWare

### Gartner Recommended Reading

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

Emerging Tech: Assess Zero-Knowledge Proof Technologies to Strengthen Competitive Advantage in Decentralized Ecosystems

Predicts 2022: Privacy Risk Expands

Top Strategic Technology Trends for 2022: Privacy-Enhancing Computation

## Data Breach Response

**Analysis By:** Nader Henein, Bernard Woo

**Benefit Rating:** High

**Market Penetration:** More than 50% of target audience

**Maturity:** Early mainstream

### Definition:

Data breach response, augmentation and the associated disclosure are the activities required to assess and notify regulatory authorities and, depending on the impact, the affected individuals when personal data is compromised. Disclosure is mandated by omnibus laws, such as the European Union's General Data Protection Regulation (GDPR), Australia's Notifiable Data Breaches (NDB) scheme, or subject- and region-specific laws, like the individual U.S. state breach notification legislation.

### Why This Is Important

Appropriate management of a breach impacting personal data can substantially reduce fines and potentially strengthen ties with affected consumers. It demonstrates that the organization is proactively taking ownership of the situation. However, delayed response, limited transparency and overly legal communications often elicit regulatory investigations, resulting in reputational damage and customer loss.

### Business Impact

Data breach response can have a critical impact on an organization's resilience. Breaches often create significant chaos as key executive team members pivot from preexisting priorities to address the reputational, regulatory and likely financial impacts of the breach. Further, newer legislation imposes statutory sentences on company directors for inadequate or negligent handling of personal data.

Drivers

- Modern privacy regulations have raised the bar for data breach notification. When personal data is impacted, disclosure to a supervisory authority within days of discovery is often required.
- In the U.S., all 50 states have breach notification laws in place and many states, such as New York and California, have amended their laws in the past two years. Amendments typically expand the data in the scope of the legislation and the responsibilities surrounding disclosure.
- Regulatory evolution illustrates the need for organizational commitment and resource allocation.
- Organizations must constantly align the technical and operational elements of incident response (IR) with new legal and regulatory requirements.
- Elevating the capacity to disclose a data breach to regulators and potentially affected individuals in an accelerated time frame is something many organizations still need to prepare for.
- Though many organizations are driven by fine avoidance, incidents are bound to happen, and a well-developed response program can pay back in dividends with fine reductions of over 50%.
- An emerging trend is rapid consumer mobilization following an incident. The impact of mass customer exodus, often led by social media, is expected to suppress regulatory fines. Also, it does not offer the organization the option of an appeal through the courts.

### Obstacles

- Establishing and testing a data breach program is an expense without an immediate return. It will pay off only if something goes wrong. This often causes the program to be deprioritized in place of more pressing or revenue-generating tasks.
- Data breach service retainers are not commonly available because of the variability and uncertainty of the type of breach, the data involved and the number of records that makes each breach scenario unique.
- Even with a strong program, the time to discover an incident can range from months to years — although it is improving over time.
- Tensions between the general counsel and chief information security officers (CISOs) over limiting information may become available through discovery following an incident. This could negatively impact the organization's capacity to effectively handle a breach.
- Data breach response requires a combination of technical acumens, such as forensics analysis of how the breach occurred, the number and type of records involved, and appropriate remediation. Data breach response must be paired with coordinated organizational processes.

### User Recommendations

- Assess whether an incident will trigger regulatory actions. Meet the threshold for a privacy violation.
- Record and maintain the details about incidents (not just violations), as some jurisdictions have stringent record-keeping requirements.
- View data breach response as a multidisciplinary process involving documented procedures and simulated drills, such as tabletop exercises. Doing so will ensure tasks are well-defined and responsibilities are clear. The process should also involve coordination and transparency between various teams and integration into the larger IR training to provide breach disclosure and rapid response requirements.
- Augment your organization's ability to address data breaches in an efficient and timely manner to fulfill regulatory and data-subject disclosure requirements.

### Sample Vendors

BigID; BreachRx; Canopy; OneTrust; RadarFirst; Securiti

### Gartner Recommended Reading

Toolkit: Cybersecurity Incident Response Plan

Toolkit: Security Incident Response Roundtable Scenario for Privacy

## Data Discovery

**Analysis By:** Michael Hoeck

**Benefit Rating:** Moderate

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

### Definition:

Data discovery solutions discover, analyze and classify structured and unstructured data to create actionable outcomes for security enforcement and data life cycle management. Using elements of metadata, content and contextual information, combined with expression- and machine-learning-based data models, data discovery solutions provide actionable guidance and processes to advance data management and security initiatives.

### Why This Is Important

Data discovery solutions improve organizations' ability to manage ever-expanding repositories of structured and unstructured data in on-premises, hybrid and cloud infrastructures. They increase visibility of disparate and unorganized sources of information. They enable compliance teams to improve insight into policy adherence and sensitive information, including personal data (PD); and enable security teams to improve visibility of sources of data access risk.

### Business Impact

Data discovery solutions can have the following business impacts:

- Accelerate the identification of sensitive data to improve the outcomes of an organization's security controls and privacy initiatives.
- Advance data life cycle management activities by assigning retention policies with data discovery categorization and classification results.

- Reduce business risk through advanced capabilities to eliminate and quarantine sensitive information, and identify data lineage and access permissions issues.

Drivers

- Organizations want to mitigate business risks associated with data processing activities (including data breach, data exfiltration, PD and intellectual property exposure, auditing and regulatory fines), identify sensitive data and implement effective data life cycle initiatives
- There is a need to minimize the blast radius of a cyberattack's access to sensitive information through data classification coupled with security controls, defensible deletion and minimization efforts.
- Organizations want to be able to align and monitor proper data access based on categorization and classification of all data.
- Retention policies can be difficult to establish, refine and consistently enforce without clear data inventory knowledge and awareness of potential sensitive data risk.
- The demands associated with the growing number and complexity of compliance and privacy regulations, such as the EU's General Data Protection Regulation (GDPR), the California Privacy Rights Act (CPRA), Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, and financial services compliance, have greatly increased interest in, and awareness of, data discovery software.
- The potential value of contextually enriched data is capturing the interest of data and analytics teams.

### Obstacles

- For its broad set of use cases, capabilities and benefits, funding and budget for data discovery solutions may require a collaborative effort across multiple departments, including security, privacy, compliance, legal and IT teams.
- Successful results of using data discovery software may be affected by a lack of data life cycle management policy buy-in or consensus from key internal constituencies, including executive sponsorship.
- Action-oriented retention policies are required to defensibly delete data identified by data discovery software.
- It can be challenging to get the organization to commit to aligning administrative costs for data discovery solutions and their management with an ongoing business program and investment, rather than a singular project-based activity.

### User Recommendations

- Use data discovery software to enable IT, security operations, privacy, compliance and line-of-business (LOB) teams to make better informed decisions regarding classification, data management and content migration.
- Use data discovery software to better grasp the risks of data footprints, including where data resides and who has access to it, and to expose another rich dataset through subsequent classification and content analysis to drive business decisions.
- Develop strong data life cycle management principles by establishing, updating and enforcing retention policies using the information gathered and remediation actions from data discovery software.
- Identify the potential risks of unknown data stored in structured database repositories often associated with applications that enable the storage of free-form text.
- Create data visualization maps to better identify the value of data and the risks to it, including the data owner, using data discovery software.

### Sample Vendors

ActiveNav; BigID; Concentric AI; Congruity360; Data443; DataGrail; Netwrix; Securiti.ai; Spirion; Varonis

### Gartner Recommended Reading

How to Succeed With Data Classification Using Modern Approaches

State of Privacy: The Privacy Tech Driving a New Age of Data Wealth

2022 Strategic Roadmap for Storage

Market Guide for E-Discovery Solutions

## Machine Identity Management

**Analysis By:** Erik Wahlstrom, Felix Gaehtgens

**Benefit Rating:** High

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Emerging

### Definition:

Machine identity management establishes identity, trust, observability and ownership of workloads and devices such as services, applications, scripts, containers, VMs, robotic process automation, IoT and mobile devices. It includes the management of the life cycle of machine identities, as well as policies and credentials such as secrets, keys and certificates that are used for trusted identification and authorization.

### Why This Is Important

- Organizations have more machines than humans and this expands the threat landscape.
- Organizations struggle to protect machine credentials such as secrets and certificates properly, leading to their leakage and abuse.
- Machine identity management needs a strong focus on observability, ownership and automation to be able to scale.
- Machines are used by many different business units within an organization, so a common set of standards and a well-aligned strategy to manage them is needed.

### Business Impact

Organizations have and use many machine identities. However, these identities are rarely managed properly, exposing them to risk. Proper machine identity management allows organizations to secure communications between workloads and/or devices, enabling current and new digital use cases. Machine identity management is critical to secure DevOps workflows by automating security through management of credentials used within the tool chain, and for machine-to-machine communication.

### Drivers

- Attacks against machine identities are on the rise, because hackers see them as “soft targets” that many organizations don’t govern and properly lockdown.
- Interest in secrets management solutions continues to drive Gartner client inquiries.
- New vendors have entered the space, resolving problems with the decentralization of workloads. Rather than storing and issuing all machine identities in a central place, new approaches discover, govern and control them in multiple places.
- Initial trust establishment — secret zero — remains a hard-to-solve issue and continues to cause frustration and security concerns. Many Gartner clients continue to store unprotected API keys used by workloads to authenticate to secrets managers.
- Stand-alone secrets management vendors are experiencing rapid growth through adoption of this technology. However, Gartner clients tell us that high pricing/licensing and reach into platforms are big obstacles to using one vendor for managing all secrets.
- There is an ongoing, but slow, convergence in the market. For example, IaaS/PaaS providers offer native capabilities, privileged access management (PAM) vendors add support for more workloads, and PKI and certificate management vendors extend into SSH and symmetric key management.
- Cloud infrastructure entitlement management (CIEM) tools (offered as stand-alone tools or part of an IaaS, identity governance and administration [IGA] or PAM suite) increase their focus — from low levels — on discovery and reporting of machine identities across systems.
- Integrations and hybrid or multicloud single-pane-of-glass functionality are becoming more important for organizations. The aim is to help increase organization efficiency and decrease response times.

### Obstacles

- Different machines, their identities and the credentials used are managed disparately, which requires a combination of different tools. Examples are tools that manage service accounts, secrets management and PKI. It is currently not feasible to manage all machine identities with only one technology.
- There are different definitions of machine identities, exacerbated by vendors providing their own interpretation and messaging. This makes selecting the right tooling difficult.
- Different business units have different needs and tool preferences. The establishment of a cross-team strategy is, therefore, critical to balance expectations around centralized governance and operations.

### User Recommendations

- Define your scope when managing different machine identities and how their management differs from IAM for humans.
- Establish ownership for machine identities using a fusion team. Also use the team to govern multiple tools and set expectations.
- Use a best-of-breed approach via multiple tools that can provide continuous observability (discovery, monitoring and behavior analysis) of machines.
- Provide tailored guidance to developers, I&O, DevOps and security teams by defining how the tools in the technology stacks should and shouldn't be used, and under what circumstances new tools or instances can be acquired and deployed. The lack of a single pane of glass increases the importance of best practices.
- Use emerging products and capabilities that address fragmentation of secrets by providing a governance and administration layer across environments via support of a bring your own vault (BYOV) approach coupled with deeper discovery.

### Sample Vendors

Aembit; Akeyless; Amazon Web Services; AppViewX; CyberArk Software; Entro Security; HashiCorp; Keyfactor; Microsoft; Venafi

### Gartner Recommended Reading

Managing Machine Identities, Secrets, Keys and Certificates

Innovation Insight: Secrets Management Tools

IAM Leaders' Guide to Privileged Access Management

### Multicloud KMaaS

**Analysis By:** Brian Lowans, Joerg Fritsch, Neil MacDonald

**Benefit Rating:** Moderate

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

#### Definition:

Multicloud key management as a service (KMaaS) solutions can be deployed across one or more private or public cloud service platforms (CSPs). These are deployed as SaaS and may use a hardware security module (HSM). KMaaS controls the keys across multiple CSPs and may also integrate to CSP-native bring your own key (BYOK) solutions to enable consistent data access control policies.

#### Why This Is Important

The deployment of data across cloud service platforms (CSP) and geographies, increases data residency, data security and privacy risks. CSP-native key management (KM) offerings do not integrate with each other. This creates inconsistent KM and increases these risks. However, multicloud KMaaS enables consistent KM through a single product that supports organizations in deploying multicloud services with strong risk mitigation.

#### Business Impact

Multicloud KMaaS supports business adoption of cloud services by enabling access to data more securely through consistent cryptography. KMaaS should be deployed to reduce exposure to data residency, security and privacy risks. This involves preventing inappropriate clear-text access by staff or CSPs to data or encryption keys at each location. Selecting a single multicloud KMaaS product instead of separate CSP-native KMs is important to simplify and apply consistent KM life cycle requirements.

### Drivers

- Organizations face fragmented KM policies across multicloud because each CSP offers its own native KM that does not integrate with other CSPs.
- Concerns about unauthorized access by national governments, CSPs or data processor staff continue to drive interest in cryptography to provide data protection and privacy.
- The risks of a data breach due to hacking, insider theft or unintended disclosure drive the need for strong, consistent KMaaS support across multicloud.
- A growing number of data protection techniques need KM support — including encryption, tokenization, secure multiparty computation, confidential computing and homomorphic encryption.
- Integration with confidential computing is enabling organizations to prevent unauthorized CSP or insider access to data and keys within these environments.
- Multicloud KMaaS can enable integration through customer-managed keys to CSP-native KM through BYOK or hold your own key (HYOK), or completely independent of the CSP.
- When BYOK integration is selected, the KM root of trust for the deployed encryption is still under the control of the CSPs. This is driving the selection of multicloud KMaaS solutions that create the root of trust independently of the CSP KM solutions.

### Obstacles

- CSPs that offer native KM solutions do not integrate with other CSPs, leading to inconsistent key life cycle management. Hence, organizations must choose between using multiple independent KM products and selecting a single multicloud KMaaS that integrates through BYOK.
- Each multicloud KMaaS product will have its own custom-made integrations to cloud services and its own custom-made selection of protection techniques.
- If a vendor provides both KMaaS and cryptography products, it rarely enables its cryptography product to integrate with KMaaS provided by another vendor. Therefore, multiple independent KMaaS purchases may be required.
- If multiple KMaaS products need to be used across multicloud, then the organization faces increased complexity for key life cycle management across these independent products.

### User Recommendations

- Investigate and identify risks around how data is stored, processed and accessed based on data security governance policies when choosing either CSP-native KM or multicloud KMaaS solutions.
- Protect all data deployed in public clouds at rest as a minimum. The granularity of protection and KM choices must achieve the accepted level of risk — leading to choices of BYOK, hold your own key (HYOK) or independent KMaaS to provide different levels of access control to keys and data.
- Ensure all cryptographic operations and access to keys meet data residency and compliance requirements.
- Ensure KM backup and life cycle management is performed consistently across all KMaaS deployed on-premises and within each CSP.
- Look for solutions that are compliant with standards for cryptography such as NIST FIPS 140-2, OASIS Key Management Interoperability Protocol (KMIP) or OASIS PKCS#11 and algorithms that will enable crypto-agility.

### Sample Vendors

Baffle; Entrust; Fortanix; OpenText; Protegrity; StorMagic; Thales; Titaniam; Utimaco

### Gartner Recommended Reading

Use Enterprise Key Management to Provide Stronger Data Security and Privacy

Select the Right Key Management as a Service to Mitigate Data Security and Privacy Risks in the Cloud

Preparing for the Quantum World With Crypto-Agility

Getting the Most Out of Your Investment in Encryption

## Secure Multiparty Computation

**Analysis By:** Joerg Fritsch, Bart Willemsen, Brian Lowans

**Benefit Rating:** Moderate

**Market Penetration:** 1% to 5% of target audience

**Maturity:** Emerging

### Definition:

Secure multiparty computation (SMPC) is a method of distributed computing and cryptography that enables entities (e.g., applications, individuals, organizations or devices) to work with data, while keeping data or encryption keys in a protected state. Specifically, SMPC allows multiple entities to share insights while keeping identifiable or otherwise sensitive data confidential from each other.

### Why This Is Important

Security and risk management (SRM) leaders struggle to achieve a balance between data security and privacy when processing (personal) data. This is further complicated by regulations and business objectives. Historically, data protection has focused on securing data at rest and in transit. However, SMPC-based methods introduce data protection in use, much like homomorphic encryption. It supports processing of data confidentially in analytics and business intelligence, using untrusted computing environments.

### Business Impact

Due to their reliance on data for artificial intelligence (AI)-based decision making and the sharing of insight from those decisions among multiple parties, SRM leaders need privacy-enhancing approaches to protect data amid an evolving landscape of maturing data protection regulations. SMPC supports the secure enablement of business, enabling organizations to uncover and exchange information, while addressing security and privacy concerns.

### Drivers

- Traditional data-at-rest encryption, as commonly implemented, does not provide strong protection against theft and data breaches. It is incapable of securing data in use and data-sharing scenarios.
- SMPC-enabled data security enables the protection of data while in use, providing SRM leaders with another data protection technique. This can be applied to new and existing use cases (e.g., multiparty information sharing).
- New use cases — such as big data analytics, AI or machine learning (ML) model training — present new privacy and cybersecurity concerns that require data-in-use protection.
- Concerns about data security in several scenarios, including collecting and retaining sensitive personal information, are driving SMPC adoption.
- SMPC helps ease fear of privacy law violations (due to the accidental exposure of sensitive information).
- SMPC supports the mitigation of sensitive data leakage, and the overall reduction and mitigation of cyberattacks.

### Obstacles

- Commercial SMPC implementations have not reached the end customer traction they could have had because implementations do not frequently match clients' needs. For example, commercial implementations may only be applicable to selected identifiers, or be only practical to protect smaller amounts of data, such as encryption keys.
- Low-end customer awareness or traction for products based on SMPC technologies outside certain niches (e.g., encryption key management or DSaaS for advanced analytics of numerical data).
- When compared with existing techniques (i.e., cryptography based on hardware-generated and stored keys), end customers could have potential issues with audits, like when their accreditation authority is not familiar with SMPC.
- If the obstacles are not addressed successfully to reignite end customer interest, SMPC will most likely head into obsolescence and will need to be removed from the Hype Cycle for data security.

### User Recommendations

- Work with developers/architects to establish a high-level position on SMPC relevance and a vision for future adoption, including proofs of concept (POCs).
- Evaluate use cases such as cloud computing, focusing on confidentiality with data in a cloud environment; privacy-enhancing (personal) data analytics initiatives; and cryptographic key protection, including encryption key management initiatives (i.e., for protection of data at rest). Also look at secure and private data mining for data and analytics use cases, including data lake security and blockchain security (e.g., wallet protection and/or quorum-based multisignature operations).

### Sample Vendors

Baffle; Cybernetica; Inpher; IXUP; LiveRamp; Nth Party; Ziroh Labs

### Gartner Recommended Reading

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

## Cloud-Native DLP

**Analysis By:** Andrew Bales, Ravisha Chugh

**Benefit Rating:** High

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Adolescent

### Definition:

Cloud-native data loss prevention (DLP) provides built-in DLP capabilities and is offered by some SaaS-delivered business applications and IaaS cloud service providers (CSPs). These solutions enable DLP control within their own environments without using third-party tools. Some cloud-native DLP mechanisms also offer content data classification and the ability to apply enterprise digital rights management (EDRM) protection.

### Why This Is Important

Cloud-native DLP offers greater visibility into and control over sensitive information processed and stored in public cloud and SaaS environments. Solutions can come with little or no additional cost if chosen from a CSP. Cloud-native DLP overcomes the limitations of legacy on-premises DLP and proxy tools to gain visibility into and control over cloud data, and provides a way to mitigate risk from sensitive data that may be inadvertently publicly exposed or used inappropriately.

### Business Impact

Choosing a cloud-native DLP solution can be beneficial for:

- Pursuing an organizationwide cloud-first strategy.
- Reducing cost, since the DLP functionality can be a sunk cost included in the license from the CSP for some cloud use cases.
- Increasing visibility and access into the stored and processed data.
- Consolidating the organization's data security vendor stack.

### Drivers

- CSPs themselves provide the most seamless integration of their DLP offerings with their native cloud offerings (such as IaaS, PaaS, SaaS).
- CSPs have the best access to — and the best opportunity to operate on — the data stored within their cloud offerings.
- Cloud-native solutions act as a good starting point for the implementation of DLP for organizations with a cloud-first strategy.
- Most of the cloud-native solutions include data classification and machine learning algorithms that can identify sensitive information.

### Obstacles

- Maintenance of several disparate and siloed cloud-native DLP technologies and policies can present difficulties to a cohesive DLP program. This is often because DLP policies cannot be automatically ported from one cloud-native DLP tool to another.
- Cloud-native DLP products are less sophisticated than enterprise DLP products and thus generally lack advanced detection capabilities.
- DLP policy duplication and overlap may happen when organizations use multiple cloud services. We have witnessed cloud-native DLP integrating with other third-party classification tools for consistency of data identification. However, a SSE solution is likely to provide a better, more cohesive outcome when the requirement is to apply consistent DLP across multiple cloud services.
- Enforcement and control of DLP policies can be nonuniform across different cloud implementations because of varying levels of maturity in product capabilities. As a result, precision and accuracy of DLP policies may vary.

### User Recommendations

- Favor cloud-native DLP if a significant portion of your sensitive data is processed and stored in a small number of public clouds.
- Test whether the built-in DLP policies are sufficient for identifying your sensitive information and fine-tune them to minimize false negatives and false positives.
- Create roles using the solution's role-based access control (RBAC) mechanism. This will permit DLP policy creation, maintenance and response.
- Evaluate the ability of cloud-native DLP to report policy violations to existing security information and event management (SIEM) and security orchestration, automation and response (SOAR) tools.
- Evaluate alternatives to outright blocking, such as redaction or encryption via EDRM.

### Sample Vendors

Amazon Web Services; Box; DoControl; Google; Microsoft; Netskope; Nightfall; Polymer; Skyhigh Security; Zscaler

### Gartner Recommended Reading

Market Guide for Data Loss Prevention

Getting DLP Right: 4 Elements of a Successful DLP Program

5 Steps to Successfully Implement Data Loss PreventionChoosing the Right Data Loss Prevention Architecture

5 Components for Securing Microsoft 365

## Format-Preserving Encryption

**Analysis By:** Brian Lowans, Joerg Fritsch

**Benefit Rating:** Moderate

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

### Definition:

Format-preserving encryption (FPE) protects data at rest and in use, and when accessed through applications while maintaining the original data length and format. It's used to protect fields in a variety of databases and document types on-premises and on public cloud services. FPE is an important anonymization technique to support data protection and privacy compliance requirements and reduces the risks of data residency, hacking or insider threats by controlling access to data.

### Why This Is Important

FPE can be used to protect data at the point of ingestion, storage in a database or access through data pipelines. It is deployed to protect data stored or processed across a variety of databases and select document types on-premises or on cloud service platforms (CSPs). However, it is still a blunt-force access control, and, when applied, it will protect data wherever it resides or is accessed.

### Business Impact

FPE is widely accepted to address privacy and financial compliance requirements and security threats without having to extensively modify databases or applications. It provides a strong, agile method to prevent unauthorized user access to data on-premises and in public CSPs. This helps organizations meet data protection and privacy regulations and data residency requirements to protect personal, health, credit card and financial data, and to adhere to data breach disclosure regulations.

### Drivers

- The national institute of standards and technology special publication (NIST SP) 800-38 standard for FPE using FF1 or FF3-1 mode is widely accepted to support privacy and financial regulations, even though modes FF2 and FF3 have suffered security flaws.
- Adoption is increasing due to the fast-growing need to provide data protection, data residency restrictions, and increasing number of privacy laws across the globe.
- Organizations increasingly want to analyze data, while keeping it anonymized. But some staff will need access to cleartext which is driving the need for FPE to provide business-friendly access controls that leverage augmented data cataloging.
- The ability to mix the implementation of FPE with data masking, and multicloud database activity monitoring (DAM) is also increasing its dynamic adoption for different use cases such as test and development.
- There is an increasing need to deploy FPE with multicloud key management as a service (KMaaS) to provide strong and consistent enterprise key management (EKM) policies to support international privacy and data residency requirements.

### Obstacles

- Encryption is frequently not coordinated across all data silos, and organizations struggle to track data flow across their architectures. This will result in clear-text access to sensitive data in other data stores that cannot be secured with FPE.
- Conflict of interest could be an issue; for example, database administrators (DBA) or application owners that have been loaded with security responsibilities without thinking about the proper segregation of duties (SOD) of DBAs from security controls.
- Encryption keys not managed by resilient life cycle best practices and EKM could lead to the loss of larger amounts of data if the encryption keys are lost.

### User Recommendations

- Ensure FPE is deployed and managed as part of EKM.
- Deploy FPE FF1 or FF3-1 to implement data security policy rules for user access in coordination with other security controls, according to Gartner's data security governance (DSG) framework.
- Review how FPE interacts with applications, establish whether this can be used to control which user identities are allowed to see the data in clear text, and try to leverage augmented data cataloging to support business access requirements.
- Evaluate the impact on the performance and functionality of applications accessing the database.
- Identify any impacts on application and database functionality, such as search and sorting.
- Monitor and audit all user and administrator access to sensitive data, even when FPE is deployed.
- Augment the data security strategy with DAM, where feasible, to monitor data movement and access behavior when accessed from a database.

### Sample Vendors

Baffle; Comforte; OpenText; Oracle; PKWARE; Prime Factors; Protegrity; SecuPi; Thales Group; Titaniam

### Gartner Recommended Reading

Use the Data Security Governance Framework to Balance Business Needs and Risks

Use Enterprise Key Management to Provide Stronger Data Security and Privacy

Select the Right Key Management as a Service to Mitigate Data Security and Privacy Risks in the Cloud

Preparing for the Quantum World With Crypto-Agility

Getting the Most Out of Your Investment in Encryption

## Privacy by Design

**Analysis By:** Bart Willemsen, Bernard Woo

**Benefit Rating:** Moderate

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

### Definition:

Privacy by design (PbD) is a set of principles about proactively creating a culture of privacy, by embedding it often and early in technology (e.g., application or customer interaction design), as well as into procedures and processes (e.g., through privacy impact assessments, data minimization and subsidiarity). There is no finite list of principles, yet PbD as a best practice is globally applicable to the basis of any privacy program.

### Why This Is Important

Privacy is one of the core tenants for organizations that are seeking to earn trust with their customers and drive increased revenue opportunities. In addition, the number of new or significantly revamped regulations continues to increase worldwide. Organizations can expect to operate more efficiently by adopting PbD and embedding privacy considerations throughout their processing activities.

### Business Impact

Privacy must be built-in. A proactive risk-based approach helps enhance consumer trust, prevent violations (such as costly data breaches) before they occur, and reduce the damage from them if they do (such as fines or brand damage). All technology design must account for the protection of any personal data at the core to mitigate privacy risk, which is at unprecedented heights with the current data volumes processed.

Drivers

- Systems must be designed so that the collection of privacy-sensitive data is transparent to the data subject. Some technology-focused ideas for implementing PbD are reducing retention length and amount of personal data (data minimization), working on the original data (rather than on copies) and applying anonymization or pseudonymization where possible, alongside purpose-based access controls (PBAC).

- The need persists to continuously evaluate the risks of reidentification and traceability, and include data location in the considerations for clarity on regulatory impact. Moreover, implementing PbD can lead to other positive changes such as designating a privacy officer with reach or procurement activities for new IT services, and frequently conducting privacy impact assessments.

- PbD and one of its subcomponents, privacy engineering, enable an approach to a business process and technology architecture that combines various methodologies in design, deployment and governance. Properly implemented, it yields an end result with an easily accessible functionality to fulfill the Organisation for Economic Co-operation and Development's (OECD's) privacy principles. It also helps mitigate the impact of a personal data breach by reimagining defense in depth from a privacy-centric vantage point.

- The process involves ongoing recalculation and rebalancing of the risk to the individual data owner while preserving optimum utility for personal data processing use cases. As a result, organizations can rely on the right data being available at the right time with maximized information retention and trust in a compliant operation.

- Stakeholders will also benefit from reducing the data footprint and accompanying breach exposure risk reduction. Further, PbD allows consistent delivery to subjects upon a privacy promise as well as collateral enhanced customer trust and engagement levels.

### Obstacles

- Adoption and widespread recognition of PbD has been hampered by a lack of industry-recognized principles and consistent regulatory framework support. The Information and Privacy Commissioner (IPC) of Ontario described seven key elements: proactivity, privacy by default, privacy embedded into design, full functionality, end-to-end security, visibility and transparency, and user centricity. In the U.S., a report by the Federal Trade Commission (FTC) of 2012 is the most visible early support for the PbD principle, yet worldwide standards are not yet being created
- Only over the past few years, legislative requirements start to include "data protection by design and/or by default," implying a PbD approach to all activities. Precedent-shaping rulings are slowly increasing in number and depth. Vendors have added statements like "product X was designed with PbD in mind," sometimes with little reference material to support the claim. Only when privacy is truly a more organic part of the development process, the need for and benefit rating of PbD increase.

### User Recommendations

- Tackle privacy by design in manageable steps; a wholesale shift will be too much to handle. Privacy by design is a cultural change about the processing of personal data. This pertains both to existing operations and to innovations.
- Adjust the existing operations through business process reengineering. Especially in innovative developments and new processes, the change begins by asking questions such as: Can we achieve the purpose set out by using less personal data? Can we end the personal data life cycle sooner? Can we provide the same functionality or customer experience without using the identifiable data? Can we adequately protect what we process? Do customers understand what we are processing about them and why?
- Identify use cases where privacy-enhancing computation (PEC) techniques can be adopted to support the embedding of privacy into current and future operational activities.

### Gartner Recommended Reading

Use a Privacy Impact Assessment to Ensure Baseline Privacy Criteria

16 Frequently Asked Questions on Organizations' Data Protection Programs

5 Privacy Imperatives for Executive Leaders

Quick Answer: How Can Executive Leaders Manage AI Trust, Risk and Security?

### Privacy Impact Assessments

**Analysis By:** Bart Willemsen

**Benefit Rating:** High

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Adolescent

#### Definition:

Privacy impact assessments (PIAs) enable organizations to identify and treat privacy risk. Typically conducted before implementing new processing activities and/or major changes, the PIA starts with a quick scan (looking at the process owner and description, types of data processed for specific purposes, and retention periods per purpose). A full PIA adds legal grounds, potential impact on data subjects, and mitigating measures to ensure a controlled personal data processing environment.

#### Why This Is Important

An ongoing shift in the regulatory privacy landscape mandates that organizations develop foundational insight into what personal data they process, why and how it is protected. Few organizations have the means to demonstrate insight in and control over personal data across the various repositories and silo types, let alone how they're used or intended to be used. This insight, however, is vital to proportionate and adequate deployment of privacy and security controls.

#### Business Impact

A PIA improves regulatory compliance, control over personal data throughout the data life cycle, and helps determine access management as well as data end of life following a deliberate intent toward purposefully processing people's data. Assisting in prevention of (internal) data breaches and personal data misuse, it helps security and risk management (SRM) leaders quantify risk to subjects and timely apply suitable mitigating controls. Conducting PIAs frequently and consistently provides a basis for responsible and transparent data management.

Drivers

- PIAs are one of the cornerstones of an effective privacy program. However, many organizations conduct PIAs manually, using spreadsheets and questionnaires. With increasing volumes and the need for repetition of PIAs, a manual approach becomes unmanageable.

- Overstandardization traps the skills needed to conduct PIAs with a few people rather than making them part of an organization's data-handling fabric.

- PIA automation tools allow for (API-driven) triggers to initiate the assessment process, collecting the needed information at every step and tracking it through a predefined workflow all the way until a case is closed or flagged for remediation.

- When done well, the PIA sits at the heart of connecting legal requirements and business process reengineering to practical operationalization in privacy by design and enablement of adequate security control application.

- The results of a PIA will help assess records of processing activities (RoPAs), and through intelligence of data fabric from data and analytics leaders, SRM leaders can further automate the intended personal data life cycle in terms of where it should and should not be available. In other words, the PIA outcome with purpose-based processing activities determines purpose-based access controls (PBAC). In addition, it facilitates automation of the determined data end-of-life moments.

- The entire PIA process eases data governance initiatives to a more controlled state, yet the current main drivers still primarily come from regulatory requirements. Additional frameworks do help, like the 2023 revamped ISO 29134.

### Obstacles

- Often considered a tedious task because of poorly conceived manual workflows and a one-size-fits-all mentality, there is a certain PIA fatigue in organizations where this activity has been mandatory for a longer period of time.
- Business partners' view of a checkbox mentality does not help the quality of the PIA.
- Others simply underestimate its relevance and position and do not complete accurate PIAs or fail to frequently keep them updated, making the initial attempt an ultimately futile one.
- PIA automation tools are hard to tailor to an organization's needs in the absence of knowledgeable and trained staff. As a result, even an automated approach fails to fulfill the purpose for subsequent automation and alignment of the personal data life cycle governance or management activities that are ideally connected to the PIA.

### User Recommendations

- Appoint and mandate business process owners with responsibility over their respective personal data processing activities, and actively involve them in optimizing the process for fluency and detail.
- Require PIAs to be conducted as a mandatory, frequently reiterated activity. Triage the necessity for PIAs in change processes and the introduction of new processing activities.
- Include the PIA's results — especially from large projects — in the corporate risk register for monitoring and follow-up. Depending on scope and focus, it may also help to integrate high-risk PIA activities to overarching business impact assessments.
- Extend the assessment's effectiveness to processing of personal data carried out by service providers by demanding that they complete and periodically revise a full PIA.
- Use a centrally provisioned tool for consistently conducting PIAs (for example, as an internal automated workflow process), or require the PIA to be conducted as a manual exercise when a less-mature procedure suffices.

### Sample Vendors

DataGrail; OneTrust; PrivacyPerfect; RESPONSUM; Securiti; Smart Global Governance; TrustArc; WireWheel

### Gartner Recommended Reading

Toolkit: Assess Your Personal Data Processing Activities

Use a Privacy Impact Assessment to Ensure Baseline Privacy Criteria

Ignition Guide to Implementing a Privacy Impact Assessment Process

## Augmented Data Catalog/Metadata Management

**Analysis By:** Guido De Simoni, Brian Lowans, Robert Thanaraj

**Benefit Rating:** High

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Early mainstream

### Definition:

Metadata management solutions are software that includes one or more of the following: metadata repositories, a business glossary, data lineage, impact analysis, rule management, semantic frameworks, and metadata ingestion and translation from different data sources. Modern AI-driven augmented data cataloging is part of the solutions automating metadata discovery, ingestion, translation, enrichment, and the creation of semantic relationships between both business and security metadata.

### Why This Is Important

Augmented data catalogs and metadata management solutions support organizations managing varied data assets. Demands for accessing, using and sharing data are not limited to IT as data-oriented citizen roles emerge. Data and analytics (D&A) and security leaders face growing security and privacy risks, necessitating new data management approaches. With the pervasive use of data across a distributed data landscape by citizen users, augmented data catalogs can support metadata discovery and inventory management automation.

### Business Impact

D&A leaders investing in augmented data cataloging (ADC) and metadata management solutions will see benefits from:

- **Collaboration:** Metadata requires the contribution of many people and ADC and metadata management solutions can provide a multiuser environment to address the requirements.
- **Automation of processes:** As data changes, metadata management solutions can streamline many recurring activities by (partial) automation.
- **Cost optimization:** Metadata management solution ensures that organizations understand the datasets, workloads, queries and tools being utilized, and highlights redundant tools and technologies.

### Drivers

- Augmented data cataloging and metadata management solutions can remediate suboptimal results in an organization's use of data due to improperly managed metadata. This saves time, effort and money, while ensuring organizations are not exposed to unnecessary risks.
- Innovation generated by active metadata, reducing human effort in inventorying and managing data assets, is accelerating augmented data cataloging and metadata management solutions. Humans are primarily validators as opposed to doers of the operational tasks associated with the metadata management solutions.
- Informal and formal teams emerge and convert to community participation with as much automation as possible when supported by augmented data cataloging and metadata management solutions. These demands are only starting to be addressed by vendors, with modern metadata management practices slowly being established within organizations.
- Enterprise data cataloging techniques are emerging from several vendors that combine business, security and privacy metadata. This enables cross-functional operations for both D&A and security products to leverage the same metadata management solutions.

### Obstacles

- The lack of maturity of strategic business conversations about metadata management solutions, as historically, enterprises have struggled to understand, define and use metadata showing business value.
- The expensive, but required, effort to integrate metadata management solutions in multivendor environments. This inhibitor has started to be addressed by vendors and community initiatives relating to openness and interoperability (see, for example, the open-source Egeria).
- The lack of identification of metadata management solutions with capabilities that meet the current and future requirements of specific use cases.

### User Recommendations

- Recognize that the augmented data cataloging and metadata management solutions market will take two to five years to reach the Plateau of Productivity as the technology continues to expand both capabilities and support for existing and emerging use cases.
- Evaluate the metadata management capabilities of your company's existing tools, including data integration, data quality and master data capabilities, before buying a modern metadata management solution.
- Pilot the use of metadata management solutions for emergent use cases, including D&A governance, security and risk, and augmented data value for analytics.
- Invest in augmented data cataloging and metadata management solutions to augment manual efforts needed to access all types of metadata and analyze the data to support your company's data fabric designs for automation.
- Drive metadata insights to mitigate risks by informing your business leaders and their teams to ask more relevant questions of the data around them.
- Cooperate with the security and privacy teams to establish if enterprise data cataloging techniques would be beneficial.

### Sample Vendors

Alation; Alex Solutions; Atlan; BigID; Collibra; data.world; IBM; Informatica; Zeenea

Gartner Recommended Reading

Market Guide for Active Metadata Management

How to Succeed With Data Classification Using Modern Approaches

Quick Answer: What Is Active Metadata?

Quick Answer: What Is Data Fabric Design?

## Climbing the Slope

### Data Classification

**Analysis By:** Ravisha Chugh, Bart Willemsen, Andrew Bales

**Benefit Rating:** High

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Mature mainstream

#### Definition:

Data classification is the process of organizing information assets using an agreed-upon categorization, taxonomy or ontology. The result is typically a large repository of metadata useful for making further decisions. This can include the application of a tag or label to a data object to facilitate its use and governance, either through the application of controls during its life cycle, or the activation of metadata using data fabric.

#### Why This Is Important

Data classification facilitates effective and efficient prioritization of data within data governance and data security programs concerned with value, access, usage, privacy, storage, ethics, quality and retention. It is vital to security, privacy and data governance programs. Data classification helps organizations distinguish the sensitivity of the data that they process, promotes a risk-based approach and improves the effectiveness of data protection controls.

#### Business Impact

Data classification supports a wide range of use cases, such as:

- Implementation of data security controls
- Privacy compliance
- Enablement of purpose-based access controls
- Risk mitigation
- Master data and application data management
- Data stewardship

- Content and records management
- Data catalogs for operations and analytics
- Data discovery for analytics and application integration
- Efficiency and optimization of systems, including tools for individual DataOps

### Drivers

- Data classification approaches — which include classification by type, owner, regulation, sensitivity and retention requirement — enable organizations to focus their security, privacy and analytics efforts on important datasets.
- When properly designed and executed, data classification serves as one of the foundations supporting ethical and compliant processing of data throughout an organization.
- Data classification is also an essential component of data governance, as by classifying the data, organizations can establish data retention, data access and data protection policies that can help reduce the risk related to data exfiltration.

### Obstacles

- Data classification initiatives have often failed because they were dependent on manual efforts by users with insufficient training.
- Data classification adoption is typically a reflection of the security posture of the organization. If the purpose of data classification is not clearly defined for employees using natural language, engagement in the data classification program is minimized.
- Data classification often fails due to poor communication. Program objectives, policies and procedures should be effectively communicated to all necessary stakeholders to avoid resistance to data classification initiatives.
- Although many vendors offer automated data classification tools that can classify more data more accurately while minimizing user effort, they are not 100% accurate — especially if they use machine learning or artificial intelligence algorithms for which models require ongoing training.

### User Recommendations

- To identify, tag and store all of their organization's data, security and risk management leaders and chief data officers should collaboratively architect and use classification capabilities.
- Implement data classification with user training as part of a data governance program.
- Use a combination of user-driven and automated data classification for success in a data classification program.
- Determine organizationwide classification use cases and efforts, and, at minimum, keep all stakeholders informed.
- Combine efforts to adhere to privacy regulations with security classification initiatives. Information can be classification-based by nature (i.e., personally identifiable information, protected health information or PCI information), or by type (i.e., contract, health record or invoice. Records should also be classified by risk category, so as to indicate the need for confidentiality, integrity and availability. Additionally, records can be classified to serve specific purposes.

#### Sample Vendors

BigID; Concentric AI; Congruity360; Microsoft; Netwrix; OneTrust; SecuritiAI; Spirion; Varonis

#### Gartner Recommended Reading

Building Effective Data Classification and Handling Documents

Improving Unstructured Data Security With Classification

How to Succeed With Data Classification Using Modern Approaches

Video: What Is Data Classification, and Why Do I Need It?

### Privacy Management Tools

**Analysis By:** Bart Willemsen

**Benefit Rating:** High

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Adolescent

### Definition:

Privacy management tools help organizations facilitate compliance insights and check processing activities against regulatory requirements. They bring structure to privacy processes and workflows, enhance insight into data flows and governance maturity, and monitor and track the privacy program's maturity progression.

### Why This Is Important

The increasing maturity of data protection legislation globally forces organizations to maintain awareness and control of personal data processing operations. Roughly two-thirds of jurisdictions worldwide have requirements similar to the EU's trendsetting GDPR in place. They range from U.S. state laws like the CCPA or the CPRA to national initiatives like Brazil's LGPD and China's PIPL. The differences in detail across requirements make managing compliance overviews manually almost impossible.

### Business Impact

Privacy management tools give business leaders oversight and accountability about handling personal data, and enable transparency and control over those activities. They contain audit capabilities to demonstrate compliance especially across multiple jurisdictions. Point solutions are more and more integrated into suites to account for an increasingly automated enablement of a privacy UX, vendor risk management, records of processing activities (ROPAs), data intelligence inventories, and more.

Drivers

- In almost all cases, passed or proposed privacy laws have been heavily influenced by the GDPR. Therefore, they are introducing concepts such as subject rights, explicit consent and timely breach disclosure. Regulatory changes are likely to continue over the coming two to three years, establishing the fundamental basis for privacy at the legislative level.

- Individuals' awareness and demand for privacy continue to rise, often through confrontation with a widespread use of new technology. Amid these ongoing pressures, organizations must adapt their privacy programs to allow better scale and performance while staying within budgets that are still tight. They should do so without exposing the business to loss through fines or reputational damages.

- The privacy landscape is becoming increasingly complex. Gartner estimates that by 2025, 75% of the world's population will have its personal data covered under modern privacy regulations. Even further, we calculate that by that time over 80% of organizations worldwide will face modern privacy and data protection requirements. Fundamental capabilities carrying even an immature privacy program include those for ROPAs, privacy impact and compliance assessments, several elements of the privacy UX and incident or data breach management.

- To implement a consistent and holistic privacy program, organizations require two sets of capabilities — privacy management capabilities and data-centric control capabilities. Some organizations opt to tackle data issues first: discovery, classification, authorization and access controls, and operationalizing end of life.

- Finally, an in-control privacy-first posture as well as transparency and control over personal data processing activities are simply good for business. They enhance brand protection, customer trust and represent an ethical approach to data and the people behind that data.

### Obstacles

- After time, a sense of “good enough” can lead to delays in adoption of privacy management tools, or worse, an absence of (automated) integration to reap the maximum benefit.
- On the other hand, some organizations already have point solutions for privacy UX components, data breach response or impact assessments in place. They might overlook opportunities for more mature complementary capabilities in an integrated suite. Attempts to integrate loose components from various vendors can often be disappointing, if only remaining manual workload is considered.
- Ongoing developments, including the “shift-left” movement where certain capabilities are automated at a code level, are often technical in nature. This hinders adoption even when they could ultimately increase immediate and long-term benefits.
- As the absence of immediate sanctions or investigations makes pressure seem to subside, some organizations continue to take a wait-and-see approach and rely on a set of unmanageable practices.

### User Recommendations

- Incorporate the demands of a rapidly evolving privacy landscape into the organization’s data strategy by developing a common baseline driven by applicable regulatory guidelines and privacy frameworks available.
- Maintain a focus on overarching capabilities with relevance across the board, including privacy impact assessments (PIAs), ROPAs, consistent vendor risk management and a people-centric privacy UX.
- Accept, adapt and evolve with the new business challenges and needs to privacy by leading with a cost-optimized set of privacy capabilities.
- Assess the extent to which privacy management tools fit the organization’s criteria for standardization needs. For example, they can help in certification preparations against the EuroPriSe framework, ISO 27701, or aligning with others like the NIST Privacy Framework, cloud codes of conduct, etc.

### Sample Vendors

DataGrail; Ketch; OneTrust; RESPONSUM; Securiti; TrustArc; WireWheel

### Gartner Recommended Reading

State of Privacy — The European Union

State of Privacy — China

State of Privacy — Regional Overview Across North America

5 Privacy Imperatives for Executive Leaders

## Data Access Governance

**Analysis By:** Joerg Fritsch

**Benefit Rating:** Moderate

**Market Penetration:** 5% to 20% of target audience

**Maturity:** Early mainstream

### Definition:

Data access governance (DAG) provides data access assessments, management and real-time monitoring capabilities for the unstructured and semistructured data found in file repositories. Its primary purpose is to determine, manage and monitor who has access to which data in an organization's repositories and how that data is classified. Further, data access governance provides an audit trail of access and permission activities.

### Why This Is Important

Unstructured data repositories are frequently not well-managed, and the progressive adoption of cloud and collaboration platforms has made the situation more complex. Cyberthreats and privacy laws further complicate setting up controls for unstructured data. DAG solutions have now become critical as they provide data access assessment, access management and real-time monitoring for unstructured and semistructured data.

### Business Impact

DAG products help organizations solve data security and privacy regulation issues, especially in sectors where data access tracking is critical — including financial services, banking, legal, insurance and retail. Further, DAG products help protect intellectual property in manufacturing and energy industries. Healthcare and research data in the federal and pharmaceutical industries can also benefit from these products.

### Drivers

- DAG adoption continues to be sustained by ever-changing threat landscapes and regulatory compliance, including the growing list of data privacy laws. In addition, data breaches, ransomware and overall cybersecurity threats highlight the enhanced focus on data compromise as a significant threat from bad actors. Specifically, unstructured data sprawl (both on-premises and hybrid or multicloud) is difficult to detect and control when compared to structured data. As a result, security leaders recognize that ensuring the security of unstructured data — by using DAG products — is paramount when there are concerns about data loss. Data loss may result from either leakage, destruction or tampering caused by ransomware, for example.

- DAG solutions help to provide the right level of access for users or groups. Furthermore, DAG also helps customers migrate some of their data repositories from on-premises to the public cloud by classifying data and cleaning up permissions beforehand.

- Many organizations start using DAG technology tactically to organize and associate Active Directory (AD) groups and entitlements in file repositories. They do so by using the data discovery features of the product. Some then implement it to continuously monitor permissions and access. Others use it to identify and involve data owners in managing access to their documents. When clients review access rights or entitlements, they typically discover that several entities have excessive access to data.

### Obstacles

- The integration into all file repositories that enterprises are using — while having parity in function and feature across all file repositories — is, in practice, challenging.

- Cloud support is frequently lacking in maturity and features. Support for cloud-native file repositories will need to mature at a faster rate in order to keep up with demand and requirements.

- There is a lack of clear and actionable integration with identity and data security products and tools. A variety of data security products offer basic DAG features, but lack a full set of critical enterprise-grade functions, such as classification, access assessment, real-time monitoring and enforcement.

- Data security posture management (DSPM) is a new category that provides many of the traditional DAG features but often at higher speed. Further, the insights provided are usually more actionable.

### User Recommendations

- Deploy DAG products to secure the access to data, and to address concerns about regulatory compliance and the security of public cloud storage data.
- Check if the DAG product has sufficient support for the cloud-native repositories that are in use in your organization.
- Implement DAG as part of an enterprisewide data security governance strategy. Organizations should work actively with business units to define data classification and sensitive data handling policies.
- Use DAG tools to discover and classify the data, and implement rules for monitoring and protecting the data per established policies.
- Implement DAG to complement database activity monitoring (DAM) and data loss prevention (DLP). Unlike DLP, DAG addresses entitlements and analyzes user activity when accessing the data in file repositories.

### Sample Vendors

AvePoint; Brainwave GRC; Concentric AI; Cyral; Lepide; Netwrix; SailPoint; SolarWinds; Varonis

### Gartner Recommended Reading

4 Critical Steps to Accelerate the Adoption of Data Security Governance

Innovation Insight: Data Security Posture Management

Use a Data Security Steering Committee to Realize Data Security Governance Objectives

## Entering the Plateau

### Enterprise Key Management

**Analysis By:** Brian Lowans, Joerg Fritsch

**Benefit Rating:** Moderate

**Market Penetration:** 20% to 50% of target audience

**Maturity:** Early mainstream

#### Definition:

Enterprise key management (EKM) provides a single, centralized software or hardware appliance for multiple symmetric encryption solutions. Encryption is used to enforce consistent data access policies across different structured and unstructured storage platforms, on-premises and in public cloud services. It facilitates key distribution and secure key storage, and maintains consistent key life cycle management.

#### Why This Is Important

Cryptographic products that implement encryption or tokenization are critical components of a data security strategy to mitigate growing data residency and privacy requirements. These products control access to data and prevent data breaches/theft due to hacking, malicious insiders or accidental disclosure. Cryptographic products require EKM to provide consistent key life cycle management to help mitigate these risks and reduce the possibility of accidental shredding of data in case keys are lost.

#### Business Impact

EKM supports an enterprisewide data security and privacy strategy, which limits risk, and reduces operational and capital costs. Enterprises can purchase a variety of specialized cryptographic products that protect data across specific on-premises and multicloud data repositories and processes, or native platform cryptographic products. Without EKM, this results in many independent key management (KM) consoles that do not integrate and will increase the security and compliance risks.

#### Drivers

- EKM policies are needed to define a consistent granularity of cryptography that is required to enforce appropriate user access controls to data.

- EKM is a means to ensure encryption is applied consistently across a variety of storage platforms. This is done by enabling cryptography to protect the data in storage, protect individual files, or protect fields stored within or accessed from SQL and NoSQL platforms. It also supports privacy-enhanced computation technologies such as secure multiparty computation, confidential computing and homomorphic encryption.

- EKM can also be deployed outside trusted environments or with hardware security module (HSM) technology, for example, to support key management as a service (KMaaS) solutions.

- EKM is increasingly required to support enterprisewide data security governance (DSG) policies that complement a broader set of product controls, such as database activity monitoring (DAM), data access governance (DAG), data loss prevention (DLP) and data access privileges, and data security platforms.

- EKM is increasingly evolving to provide an enterprisewide plan for disaster recovery situations throughout the key life cycle, including key backup, recovery, escrow processes or changes to algorithms due to advancing postquantum threats.

- The use of EKM products will greatly simplify the ability to provide consistent KM policies and data access policies across disparate data silos and multicloud architectures.

- EKM can reduce the number of vendors and native KM products in use, if storage and self-encrypting-drive vendors (that do not offer EKM products) are complying with standards. Example standards include the OASIS KMIP, OASIS PKCS#11 and NIST FIPS 140-2 standard, ranging from Level 1 (software lowest security) to Level 3 or 4 (pure implementation or integration to a hardware security module).

Obstacles

- Although EKM products typically comply with KM standards, vendor cryptography products generally only integrate with their own KM using proprietary interfaces. This means that cryptography products often cannot be managed by a different vendor's EKM.

- A number of on-premises and cloud databases, key vaults, and SaaS provide their own native encryption and KM offering that may require custom integration by each vendor's EKM solution. This makes implementation of EKM a huge challenge in the wake of many incompatible products and disparate native platform encryption offerings managed by a variety of administrators.

### User Recommendations

- Reduce the number of KM and cryptographic products deployed by different vendors.
- Ensure the EKM can manage third-party cryptography or native products compliant with KM standards.
- Prepare for deployment of postquantum cryptography in the next few years.
- Verify that EKM offers KMaaS to integrate with cloud-native KM solutions using bring your own key (BYOK) or hold your own key for each cloud service.
- Assess whether EKM should be deployed as a software or a hardware appliance, and can achieve the appropriate security-accredited standard under NIST FIPS 140-2.
- Review carefully each vendor's EKM ability to operate across various structured and unstructured encryption solutions that require protection.
- Be careful when selecting EKM and cryptographic products from multiple vendors. EKM vendors rely on a protectionist strategy, and their products do not typically integrate with other vendors' cryptographic products.

### Sample Vendors

Fornetix; Fortanix; IBM; OpenText; PKWARE; Protegrity; StorMagic; Thales; Townsend Security

### Gartner Recommended Reading

Use the Data Security Governance Framework to Balance Business Needs and Risks

Use Enterprise Key Management to Provide Stronger Data Security and Privacy

Select the Right Key Management as a Service to Mitigate Data Security and Privacy Risks in the Cloud

Preparing for the Quantum World With Crypto-Agility

Getting the Most Out of Your Investment in Encryption

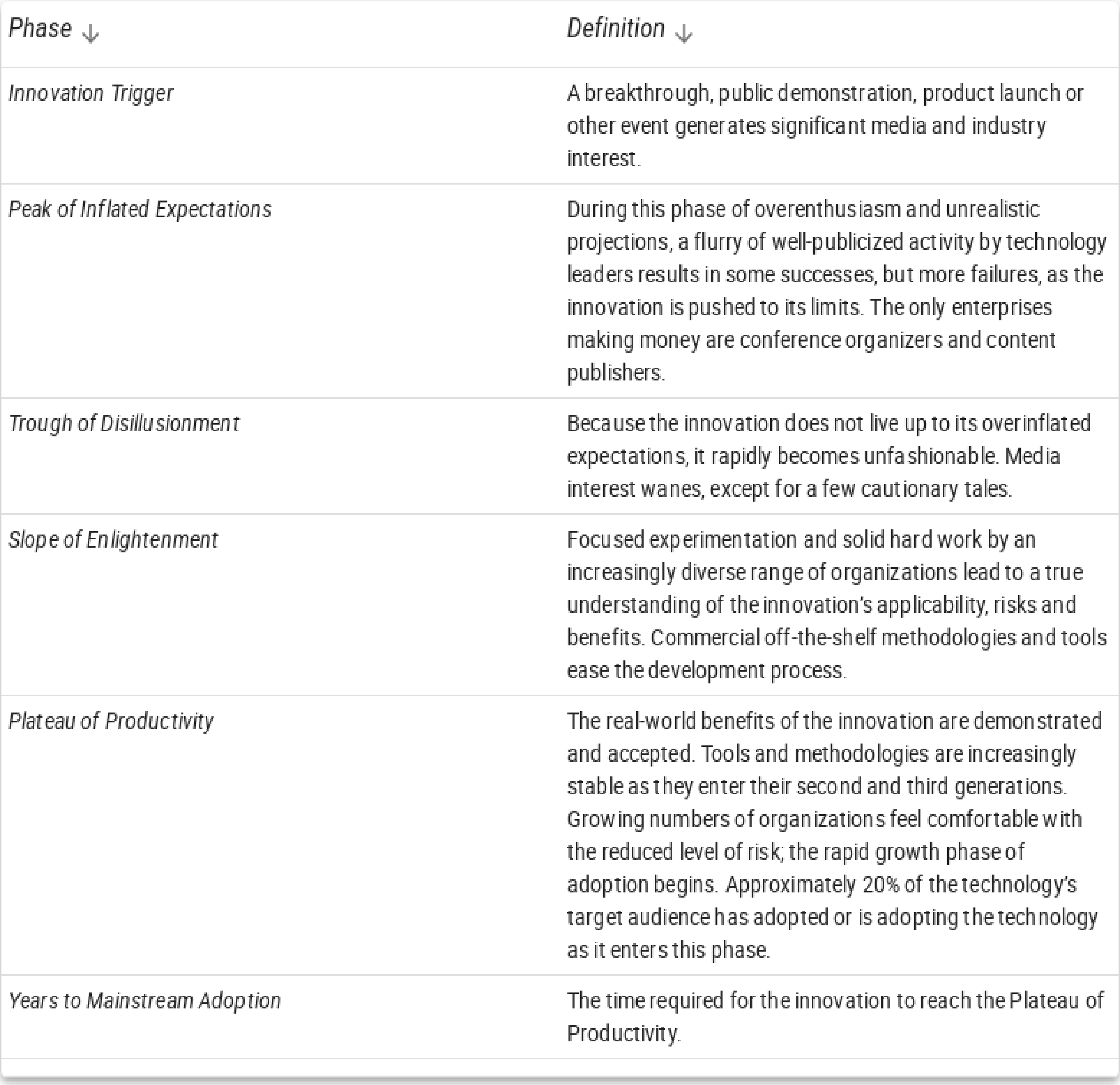
## Appendixes

See the previous Hype Cycle: Hype Cycle for Data Security, 2022

## Hype Cycle Phases, Benefit Ratings and Maturity Levels

### Table 2: Hype Cycle Phases

(Enlarged table in Appendix)



| *Phase* ↓ | *Definition* ↓ |
|---|---|
| *Innovation Trigger* | A breakthrough, public demonstration, product launch or other event generates significant media and industry interest. |
| *Peak of Inflated Expectations* | During this phase of overenthusiasm and unrealistic projections, a flurry of well-publicized activity by technology leaders results in some successes, but more failures, as the innovation is pushed to its limits. The only enterprises making money are conference organizers and content publishers. |
| *Trough of Disillusionment* | Because the innovation does not live up to its overinflated expectations, it rapidly becomes unfashionable. Media interest wanes, except for a few cautionary tales. |
| *Slope of Enlightenment* | Focused experimentation and solid hard work by an increasingly diverse range of organizations lead to a true understanding of the innovation's applicability, risks and benefits. Commercial off-the-shelf methodologies and tools ease the development process. |
| *Plateau of Productivity* | The real-world benefits of the innovation are demonstrated and accepted. Tools and methodologies are increasingly stable as they enter their second and third generations. Growing numbers of organizations feel comfortable with the reduced level of risk; the rapid growth phase of adoption begins. Approximately 20% of the technology's target audience has adopted or is adopting the technology as it enters this phase. |
| *Years to Mainstream Adoption* | The time required for the innovation to reach the Plateau of Productivity. |

Source: Gartner (July 2023)

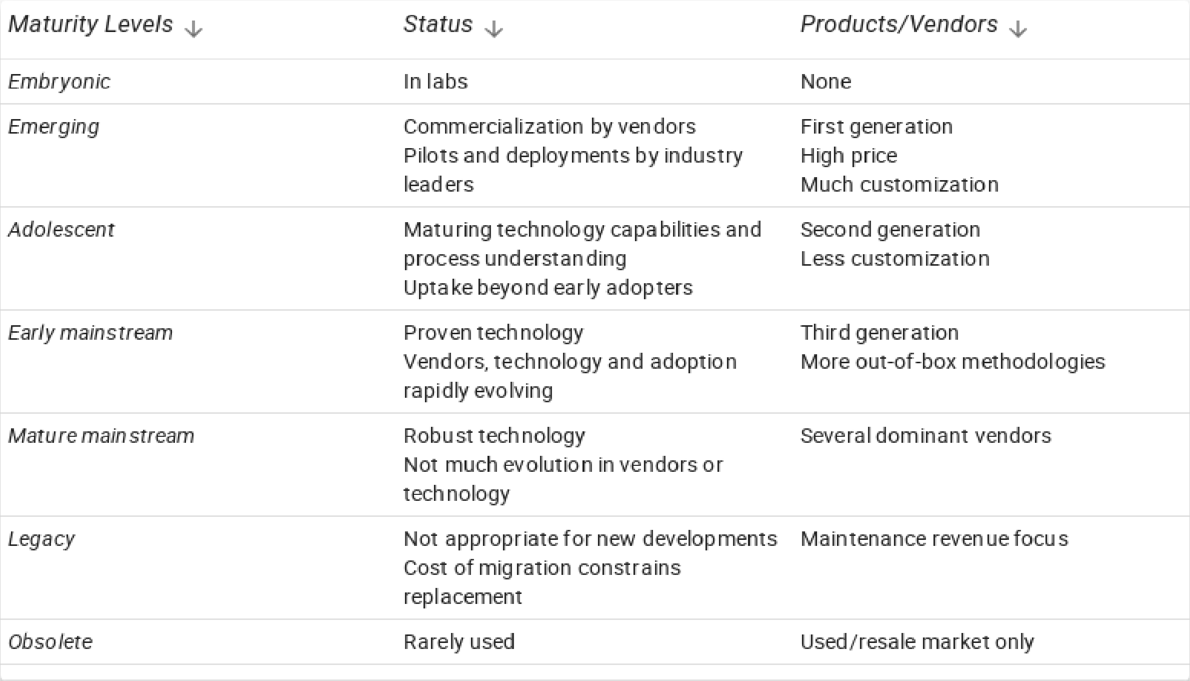
Table 3: Benefit Ratings

| *Benefit Rating* ↓ | *Definition* ↓ |
|---|---|
| *Transformational* | Enables new ways of doing business across industries that will result in major shifts in industry dynamics |
| *High* | Enables new ways of performing horizontal or vertical processes that will result in significantly increased revenue or cost savings for an enterprise |
| *Moderate* | Provides incremental improvements to established processes that will result in increased revenue or cost savings for an enterprise |
| *Low* | Slightly improves processes (for example, improved user experience) that will be difficult to translate into increased revenue or cost savings |

Source: Gartner (July 2023)

### Table 4: Maturity Levels

(Enlarged table in Appendix)



| Maturity Levels | Status | Products/Vendors |
|---|---|---|
| *Embryonic* | In labs | None |
| *Emerging* | Commercialization by vendors<br>Pilots and deployments by industry leaders | First generation<br>High price<br>Much customization |
| *Adolescent* | Maturing technology capabilities and process understanding<br>Uptake beyond early adopters | Second generation<br>Less customization |
| *Early mainstream* | Proven technology<br>Vendors, technology and adoption rapidly evolving | Third generation<br>More out-of-box methodologies |
| *Mature mainstream* | Robust technology<br>Not much evolution in vendors or technology | Several dominant vendors |
| *Legacy* | Not appropriate for new developments<br>Cost of migration constrains replacement | Maintenance revenue focus |
| *Obsolete* | Rarely used | Used/resale market only |

Source: Gartner (July 2023)

## Evidence

[1] 2022 Data Breach Report, Identity Theft Resource Center (ITRC).

[2] Cost of a Data Breach 2022 Report, IBM.

[3] **2022 Gartner Shifting Cybersecurity Operating Model Survey**: This survey was conducted to determine the impact of the changing technology governance environment on the security operating model at the macrolevel. It was conducted online from October through November 2022, with over 450 respondents from North America (n = 148), Europe (n = 216), Latin America (n = 33) and Asia/Pacific (n = 61). Respondents were required to be cybersecurity or information security leaders.

*Disclaimer: The results of this survey do not represent global findings or the market as a whole. They reflect the sentiments of the respondents and companies surveyed.*

## Document Revision History

Hype Cycle for Data Security, 2022 - 4 August 2022

Hype Cycle for Data Security, 2021 - 27 July 2021

Hype Cycle for Data Security, 2020 - 24 July 2020

Hype Cycle for Data Security, 2019 - 30 July 2019

Hype Cycle for Data Security, 2018 - 24 July 2018

Hype Cycle for Data Security, 2017 - 28 July 2017

Hype Cycle for Data Security, 2016 - 13 July 2016

Hype Cycle for Data Security, 2015 - 21 July 2015

Hype Cycle for Data and Collaboration Security, 2014 - 21 July 2014

Hype Cycle for Data and Collaboration Security, 2013 - 25 July 2013

Hype Cycle for Data and Collaboration Security, 2012 - 27 July 2012

Hype Cycle for Data and Collaboration Security, 2011 - 26 July 2011

Hype Cycle for Data and Application Security, 2010 - 26 July 2010

## Recommended by the Author

Some documents may not be available as part of your current Gartner subscription.

Understanding Gartner's Hype Cycles

Tool: Create Your Own Hype Cycle With Gartner's Hype Cycle Builder

A Data Risk Assessment Is the Foundation of Data Security Governance

Use a Data Security Steering Committee to Realize Data Security Governance Objectives

4 Critical Steps to Accelerate the Adoption of Data Security Governance

Innovation Insight: Data Security Posture Management

2023 Strategic Roadmap for Data Security Platform Adoption

Top 5 Data Security Use Cases You Must Address With Unstructured Data

Preparing for the Quantum World With Crypto-Agility

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner's Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity." Gartner research may not be used as input into or for the training or development of generative artificial intelligence, machine learning, algorithms, software, or related technologies.

Table 1: Priority Matrix for Data Security, 2023

| *Benefit* | *Years to Mainstream Adoption* | | | |
|---|---|---|---|---|
| ↓ | ***Less Than 2 Years*** ↓ | ***2 - 5 Years*** ↓ | ***5 - 10 Years*** ↓ | ***More Than 10 Years*** ↓ |
| Transformational | | Security Service Edge | Data Risk Assessment<br>Data Security Governance<br>Data Security Posture Management<br>FinDRA<br>Homomorphic Encryption | |
| High | Data Breach Response<br>Privacy Management Tools | Augmented Data Catalog/Metadata Management<br>Data Classification<br>Digital Communications Governance<br>Machine Identity Management<br>Postquantum Cryptography<br>Privacy Impact Assessments<br>Synthetic Data | Cloud-Native DLP<br>Crypto-Agility<br>Data Security as a Service<br>Data Security Platforms<br>Multicloud DAM<br>Sovereign Data Strategies | |

| *Benefit* | *Years to Mainstream Adoption* | | | |
|---|---|---|---|---|
| ↓ | ***Less Than 2 Years*** ↓ | ***2 - 5 Years*** ↓ | ***5 - 10 Years*** ↓ | ***More Than 10 Years*** ↓ |
| Moderate | Data Access Governance<br>Enterprise Key Management | DevOps Test Data Management<br>Multicloud KMaaS<br>Privacy by Design | Confidential Computing<br>Data Discovery<br>Differential Privacy<br>Format-Preserving Encryption<br>Zero-Knowledge Proofs | |
| Low | | Quantum Key Distribution | | |

Source: Gartner (July 2023)

### Table 2: Hype Cycle Phases

| *Phase* ↓ | *Definition* ↓ |
|---|---|
| *Innovation Trigger* | A breakthrough, public demonstration, product launch or other event generates significant media and industry interest. |
| *Peak of Inflated Expectations* | During this phase of overenthusiasm and unrealistic projections, a flurry of well-publicized activity by technology leaders results in some successes, but more failures, as the innovation is pushed to its limits. The only enterprises making money are conference organizers and content publishers. |
| *Trough of Disillusionment* | Because the innovation does not live up to its overinflated expectations, it rapidly becomes unfashionable. Media interest wanes, except for a few cautionary tales. |
| *Slope of Enlightenment* | Focused experimentation and solid hard work by an increasingly diverse range of organizations lead to a true understanding of the innovation's applicability, risks and benefits. Commercial off-the-shelf methodologies and tools ease the development process. |
| *Plateau of Productivity* | The real-world benefits of the innovation are demonstrated and accepted. Tools and methodologies are increasingly stable as they enter their second and third generations. Growing numbers of organizations feel comfortable with the reduced level of risk; the rapid growth phase of adoption begins. Approximately 20% of the technology's target audience has adopted or is adopting the technology as it enters this phase. |
| *Years to Mainstream Adoption* | The time required for the innovation to reach the Plateau of Productivity. |

| *Phase* ↓ | *Definition* ↓ |
|---|---|

Source: Gartner (July 2023)

**Table 3: Benefit Ratings**

| *Benefit Rating* ↓ | *Definition* ↓ |
|---|---|
| *Transformational* | Enables new ways of doing business across industries that will result in major shifts in industry dynamics |
| *High* | Enables new ways of performing horizontal or vertical processes that will result in significantly increased revenue or cost savings for an enterprise |
| *Moderate* | Provides incremental improvements to established processes that will result in increased revenue or cost savings for an enterprise |
| *Low* | Slightly improves processes (for example, improved user experience) that will be difficult to translate into increased revenue or cost savings |

Source: Gartner (July 2023)

Table 4: Maturity Levels

| *Maturity Levels* ↓ | *Status* ↓ | *Products/Vendors* ↓ |
|---|---|---|
| *Embryonic* | In labs | None |
| *Emerging* | Commercialization by vendors<br>Pilots and deployments by industry leaders | First generation<br>High price<br>Much customization |
| *Adolescent* | Maturing technology capabilities and process understanding<br>Uptake beyond early adopters | Second generation<br>Less customization |
| *Early mainstream* | Proven technology<br>Vendors, technology and adoption rapidly evolving | Third generation<br>More out-of-box methodologies |
| *Mature mainstream* | Robust technology<br>Not much evolution in vendors or technology | Several dominant vendors |
| *Legacy* | Not appropriate for new developments<br>Cost of migration constrains replacement | Maintenance revenue focus |
| *Obsolete* | Rarely used | Used/resale market only |

Source: Gartner (July 2023)