2022 Strategic Roadmap for Data Security Platform Convergence

Published 28 September 2021 - ID G00748558 - 31 min read

By Analyst(s): Joerg Fritsch, Brian Lowans, David Mahdi

Initiatives: Security of Applications and Data

Organizations use an increasingly complex set of security controls. Successful SRM leaders can significantly improve business utilization and data value by building a migration plan from siloed data security offerings to data security platforms enabling simpler, consistent end-to-end data security.

Additional Perspectives

- Summary Translation: 2022 Strategic Roadmap for Data Security Platform Convergence
  (27 October 2021)

Overview

Key Findings

- The relentless pace of business activity is creating a significant challenge for the data security team. Organizations are challenged by having to secure data that is constantly transposed by digital processes and business ecosystems spanning storage silos on-premises and in the cloud.
- Most organizations have outdated policies, frameworks and obsolete tooling which prevents them from leveraging their data and is leading to growing risks of data compromise.
- Patchworks of use case and silo-specific security controls are resulting in security and risk management (SRM) leaders struggling to understand the capabilities and limitations.
- This complexity is encouraging vendors to rapidly amalgamate disparate data security capabilities into data security platforms. Organizations applying these newer platforms are securing their data better and more easily.

Recommendations

To benefit from this amalgamation, as an SRM leader responsible for data security, you should:

- Inventory data security controls to implement a multiyear phaseout of siloed data security tools that are holding you back when you need to leverage your data in favor of a modern data security platform.
- Consolidate vendors and cut complexity and costs as contracts renew. For example, in database activity monitoring, data masking, data discovery, data encryption or data access governance.
- Actively engage with initiatives for data lakes and artificial intelligence (AI)/machine learning (ML) use cases in order to integrate a data security platform into the scope of project planning.Plan to implement a data security platform (DSP) that covers data located on-premises and cloud-based, next-generation data stores as part of your multiyear consolidation efforts.
- Include data security platforms in your cybersecurity mesh architecture by choosing DSP products that offer high levels of integration capability. The security of the composable enterprise requires flexible cybersecurity mechanisms with a rich set of APIs, based on interoperability standards.

Strategic Planning Assumptions

By 2024, 30% of enterprises will have adopted DSP, up from less than 5% in 2019, due to the pent up demand for higher levels of data security and the rapid increase in product capabilities.

By 2025 30% of Gartner clients will protect their data using a “A need to share” approach rather than the traditional “Need to know” approach.

Introduction

This document was revised on 30 September 2021. The document you are viewing is the corrected version. For more information, see the Corrections page on gartner.com.

Data security has become a key issue for SRM leaders. This is driven by the growth in privacy and regulations and other legislation, as well as by the transition to cloud services, on top of more traditional intellectual property protection needs. Data security has several challenges, especially to do with the variety and volume of the information organizations hold and the increased need to process and share sensitive data.

For most use cases, several data security controls and technologies will be required to address data security and privacy risks and to protect data well enough for the envisioned use. The number of technologies that teams engage in for data protection has been creeping up, as the requirements complexity grows faster than vendors are consolidating and integrating these technologies. Managing the rule set in these tools in isolation almost invariably leads to gaps or side effects in overlapping areas. Meanwhile, there are DSPs evolving that bridge gaps and carefully orchestrate the implementation of data security controls and policies.

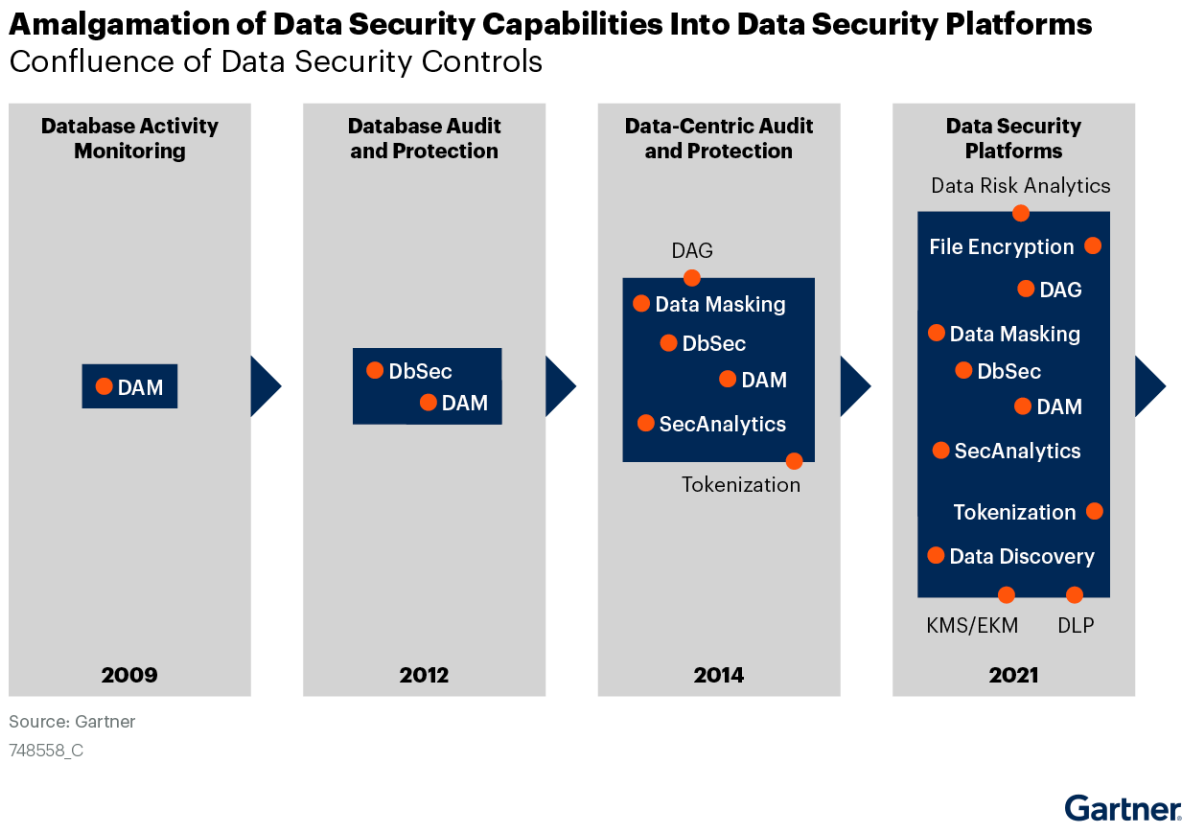
The confluence of data security into large product platforms has never been more evident than right now, but complexity grows faster than vendors are consolidating.

Examples of market convergence areas include:

- Tokenization, encryption, database activity monitoring (DAM) and data masking. Mature tokenization products offer DAM capabilities based on database agents or network gateways and data masking capabilities are used as a post processing step after tokenizing or encryption of the data.
- Data masking, data discovery, DAM and data access governance (DAG). Traditional data masking products add DAM capabilities by reconstructing audit logs from dynamic data masking gateways. Data discovery capabilities are added to give clients greater visibility across their data stores. Frequently the information is stored in a data catalog.
- Data discovery, tokenization and data governance.
- Data discovery and data classification.

Figure 1 illustrates the confluence of data security controls since the year 2009. Controls that are on the boundaries of the blue area are about to be included, but not many DSPs have them yet.

Figure 1. Amalgamation of Data Security Capabilities Into Data Security Platforms



Gartner defines data security platforms (DSPs) as products and services characterized by data security offerings that target the integration of the unique protection requirements of data across data types, storage silos and ecosystems. The data security market is currently characterized with vendors integrating their existing capabilities into a DSP. In this market the formerly siloed capabilities will come together under a common policy instrument, considerably streamlining data security and making DSPs a key enabler of meaningful data risk analytics assessments.

Future State

Successful SRM leaders can significantly facilitate the business utilization and value of data by transitioning from siloed data security offerings to data security platforms, enabling simpler, consistent end-to-end data security. DSP offerings deliver and protect this future state (i.e., 2024 and beyond; see Table 1).

Table 1: DSP Future State

(Enlarged table in Appendix)

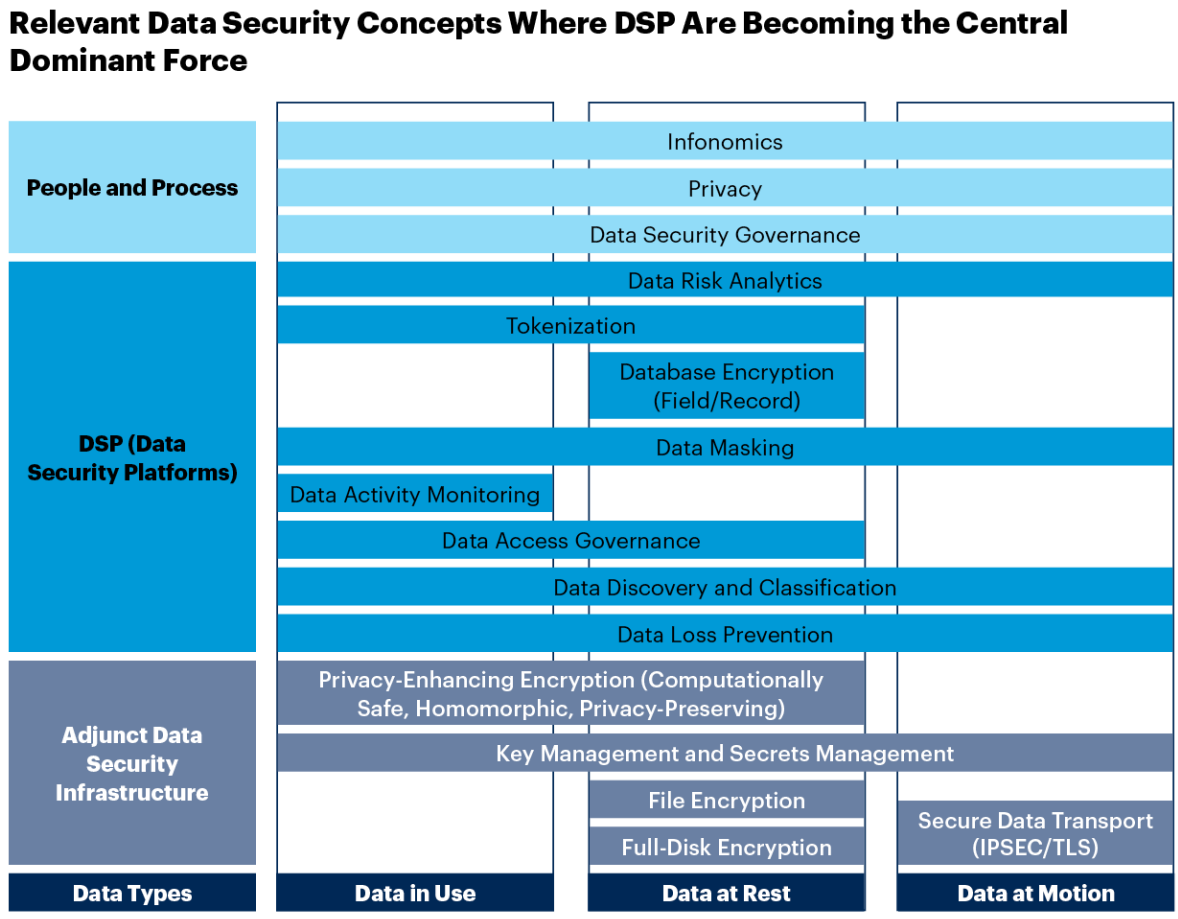
Future State	Description
Consistent visibility on (sensitive) data, data stores, policies and applicable regulations.	Sensitive-data visibility and control is a critical capability of DSP. This is enabled using a consistent overview of data silos, data security policies, controls and the applicable regulations. The best DSPs will have semantic capabilities for data classification — judging what something really is, rather than relying on preconfigured identifiers.
High levels of integration capability and simplified deployment models.	The use of API integration and cloud-delivered services will increase in importance. DSP provides a number of integration options with popular data stores, such as API integration, agent software or network gateways. Customers can choose the DSP implementation architecture that is least/noninvasive in their environment. Customers can choose to position enforcement points as close to data assets as possible. This will maximize protection, while minimizing user impact.
Ease of administration via a consolidated policy control plane.	The DSP management control plane is decoupled from data types and control objects, allowing centralized administration. The administrative interface will allow data security policy to be managed from a single console and applied regardless of the data silo or the required control objective. AI and ML will be integral to automate policy creation. Full API enablement allows automation and integration with existing processes and tools.
Democratized access to the technologies required to secure data thoroughly.	DSPs are available as stand-alone tools and cloud-based service offerings. Cloud-based offerings will make most of the DSP security objects available via low threshold API integration, making best in class data security controls achievable and affordable for many.
Roles moving from being focused on a single product or technology to being multidimensional.	DSP engineers and administrators live at the intersection of securing data and sharing data. The primary roles of DSP engineers are to secure data well enough so that it can be used as required (without compromising security), instead of locking data away by maintaining a single product-driven data security control.
Data stores for AI and ML, for example data lakes, will increase in importance.	The use of identifiable and regulated data for AI model training raises security and compliance concerns around personal data usage and the potential negative effects in production deployments. Gartner expects DSPs to make an important contribution to the mitigation of those. For example, providing data security for next-generation use cases and enabling migration to public clouds and providing scalable and reliable controls that can support immense data volumes.

Source: Gartner (September 2021)

Technology

A more detailed view of the future state of DSP is shown in Figure 2. DSP controls are displayed in the center where the color indicates a high-level relative placement on the control family. The Section “Data Security Platforms” provides an overview of the scope of these technologies, the role they perform for data security and the best practices for using them.

Figure 2. Relevant Data Security Concepts Where DSPs Are Becoming the Central Dominant Force



Source: Gartner
748558_C

Gartner

People and Process

Data Security Governance (DSG)

DSG refers to a subset of information governance that deals specifically with protecting corporate data (in both structured databases and unstructured file-based forms) through defined data policies and processes, and implemented via technologies that are by and large drawn from DSPs.

Privacy

Privacy compliance, and its accompanying architecture requirements, has become a predominant driver for many elements of data security in Gartner inquiry. Privacy and related regulations require a data protection impact assessment and usually avoid specifying specific data security controls. But may imply that data protection solutions — such as privacy-enhancing computation techniques, encryption, tokenization and data loss prevention (DLP) — must be used to assist in compliance. Generally, these controls are part of DSPs. Data security controls play a significant role in the support of privacy programs.

Infonomics

Enabling the secure and lawful monetization of information within the business ecosystem has become both a competitive differentiator for leading organizations and a data and analytics strategy priority. For example, an emergent product category are data exchange vendors who have a data sharing platform and services that companies can license — allowing them to operate their own secure closed-ecosystem sharing of data, containing only approved participants.

Data Security Platforms (DSPs)

Data Loss Prevention (DLP)

DLP (also known as “data leakage prevention”) technology is designed to stop data being moved, used or located where it shouldn’t. The technology has been available for a long time but used with varying degrees of success. DLP vendors use the concepts of data at rest, data in motion and data in use to structure DLP architectures. DLP can be found in a wide variety of products, so it can help to think of DLP as a capability rather than a market. DLP capabilities of increasing competence can be found not only as “enterprise DLP” suites, but also commonly within DSPs, SaaS and infrastructure as a service (IaaS) environments as native controls. They can also be found in secure web and email gateways, endpoint protection platforms, cloud access security brokers (CASBs) and firewalls.

Data Discovery and Classification

Data discovery and classification tools search and assign data to a category. Most security tools in this area do so based on pattern matching and the relative level of sensitivity, then record that outcome in a repository or as a tag or label on the document. A range of tools use alternate, less granular categorizations, such as personally identifiable information (PII), patient health information (PHI) or credit card data, to achieve their goals.

Recently, the market for data discovery and classification tools has undergone two important changes:

1. *A specific and growing market for privacy management and data tracking that includes classification capabilities that specifically discover, for example, personally identifiable information (PII) as a category.*
2. *Vendors strive to move away from algorithmically pattern matching and are adding AI/ML-based capabilities that are enabling semantic capabilities to find out what something means. For example, the number 72 could be a house number, a temperature — literally anything. You can hardly find out what something is when a product is limited to pattern matching for discovery and classification.*

Data Access Governance (DAG)

These products are focused on the implementation of data security access policies for unstructured data. DSPs that include DAG generally provide:

- Data discovery
- Classification
- Data owner identification
- Activity monitoring and auditing for file shares
- Network attached storage (NAS)
- Document repositories, such as Microsoft OneDrive
- SaaS content collaboration platforms such as Box
- Directory services, such as Azure Active Directory

DAG products that are extending into the DSP space are also applying their governance controls to data in relational data stores. For example, popular products have the ability to collect user and role assignments (and analyze permissions), discover sensitive data, and monitor database user activity and configurations in selected relational data stores.

Data Access Control

Security problems for structured and unstructured data can be addressed in many ways, but at the root of most issues is access to the data. Poor control of access to data can lead, for example, to data exposure or loss of data through ransomware, and also complicate other controls such as DLP. This concept applies to structured and unstructured data, whether located in the cloud or on-premises. The principle of least privilege should be implemented as much as possible, and tools to do this now cross the boundary between identity and access management (IAM) and DSPs. Gartner observes that the enforcement is frequently done using IAM technologies while data access governance is occasionally part of DSPs.

Data(base) Activity Monitoring (DAM)

The ability to detect changes and create alerts for privilege escalation, or for changes to data, is important in order to detect potential malicious insider or external hacking activities and to meet certain compliance mandates. Over the past 10 years, DAM products have constantly added features and many have evolved into DSPs or have become part of them.

Currently, DSP vendors are redefining DAM from an agent-based capability on traditional database servers into products that use database APIs and cloud APIs to get the required information for finding data misuse. This has been started by the fact that the traditional DAM agents cannot be installed on cloud databases. However, the new DAM generations may not offer the ability to assess highly privileged users such as database administrators, system administrators or cloud service provider (CSP) engineers. Therefore, it is also important that DSPs have additional features, such as DLP or tokenization, enabling CISOs to close the new gaps.

Data Masking

Data masking transforms data so that it is either unreadable or at least deidentified, allowing processing to happen in a compliant fashion. DSPs include data masking either as a dedicated capability or as part of the tokenization capabilities where masks can be applied as a post-processing step after the actual tokenization.

Static data masking (SDM) is used to create a deidentified copy of a dataset, and this copy is then used in one or more applications or processes. The deidentification does not happen in real time, and is often initiated manually, as a scheduled batch task or automatically, as part of a workflow. The SDM can be used to cross otherwise-separated application and data environments, such as production and training systems, without connecting them in real time.

Referential integrity, that is masking the same unique identifier in the same way across many databases, is the de facto standard and does not need to be a concern — although clients frequently ask for it in inquiries with Gartner.

Dynamic data masking (DDM) is essentially a late binding access control. DDM applies masking operations in real time when an application or person accesses data. The original data resides in the data repository and, therefore, DDM provides protection for data in use only. The original data is accessible to a consuming entity when authorized by policy. Entities that are not authorized to access the sensitive information are provided with masked data instead.Unstructured/semistructured redaction (USR) can protect sensitive unstructured (PDF, Excel files, text files, log files, etc.) and semistructured (XML, JSON, etc.) content with data redaction technology. Demand for USR and semistructured data redaction from DM vendors is not as strong as demand for SDM and DDM for relational and data warehousing platforms. However, some DSP vendors are adding USR capabilities as it is also required for data lakes (that are used for AI/ML) where unstructured and semistructured data uis stored.

Database Encryption (Field/Record)

Field-level encryption (FLE) can protect individual fields and documents with all key management, encryption and decryption operations occurring exclusively outside the database server. With FLE enabled, a compromised administrator or user obtaining access to the database, the underlying filesystem or the contents of server memory, (e.g., via scraping or process inspection) will only see unreadable encrypted data.

For the protection of data in relational databases, FLE has reached little relevance. Organizations that require FLE for relational databases are generally using tokenization. Column-level encryption, that is part of Microsoft SQL Server 2018 and later, is more widespread. Clients frequently abandon FLE or column-level security because the encrypted fields and columns cannot easily be processed (for example sorted).

Tokenization

Tokenization and format-preserving encryption (FPE) protect a data field by replacing its value with a substitute when it is loaded into an application or data store. This protects data at rest, as well as data in subsequent use. If an application or user needs the real data value, the substitute can be transformed back because the algorithms are designed to be reversible. The algorithms are designed to provide unique and consistent mapping between the real value and the token, where referential integrity is maintained.

Tokenization is gaining popularity for many use cases besides Payment Card Industry Data Security Standard (PCI DSS). Vendors are responding by amalgamating adjunct features such as discovery and monitoring into comprehensive DSPs rather than specialist products that are tied to single use cases.

Data Risk Analytics

Data risk analytics is an approach to data security that focuses on data to better configure proactive security measures. For example, data classification labels, data access privileges, behavioral data and the configuration or vulnerabilities of data stores can be used to calculate data risk scores and other data risk metrics that can be used to remediate security gaps before data breaches occur.

DSPs can have data risk analytics capabilities to varying degrees. For example, some DSPs have a dashboard where the data risk is calculated with a vendor proprietary methodology and displayed with a color code or using a percent score (methodologies are not standardized and still maturing). Gartner expects that data risk analytics capabilities in DSPs will be one of the benefits of using a DSP rather than products that are not integrated with each other or from multiple vendors.

Adjunct Data Security Infrastructure

Full Disk Encryption

Full-disk encryption (FDE) software encrypts the complete contents of an endpoint's storage device, which is typically the full hard disk, including the operating system. The only files excluded from encryption are the ones needed for the actual booting of the system. FDE is the most important protection against loss of sensitive data through endpoint/device loss.

The use of encryption requires key management. Not all encryption options — for example transparent data encryption (TDE) — come with integrated enterprise key management. Frequently a separate product is needed to avoid losing the keys, which could result in losing the data forever.

Secure Data Transport

The encryption of data in transit has become the norm. Transport layer security (TLS) at the application layer has largely displaced IPSEC (a network level protocol) for the majority of use cases. For example, many APIs, API edge gateways or reverse proxies include TLS encryption. Moreover, modern zero trust network access (ZTNA) products are based on TLS encryption rather than IPSEC. IPSEC is still relevant for data secrecy over wide-area networks (WAN).

File Encryption

In some cases, individual files or folders on an endpoint need to be protected from unauthorized user access. Just because multiple users have authorized access to an endpoint does not mean they are all authorized to decrypt the files stored there. For example, administrators need remote and local access to employee devices for troubleshooting and maintenance purposes. However, administrators are not authorized to decrypt sensitive employee files, and they should not be able to do so.

Several DSPs have file encryption capabilities that can replace the TDE capabilities of the database vendors. However, with the inclusion of TDE into the base license of Microsoft SQL Server 2019 and later, this became less relevant.

Enterprise Key Management (EKM) and Secret Management

Managing keys and secrets, such as credentials that are needed to programmatically authenticate to an API, is vital to maintaining security and availability. It is especially critical for data-at-rest encryption, such as TDE, database field level encryption or tokenization, where loss of keys may result in permanently inaccessible data. With encryption being ever more pervasive, the number of keys and secrets is rapidly growing, and the number of systems and applications that use encryption become more diverse. Hardware security modules (HSMs) are at the core of every EKM strategy.

DSPs frequently have mature integrated EKM based on the vendor's HSM product.

If you do not know whether your organization requires HSM or not, then most likely you do not need it. If you need to go down the HSM route, then be prepared to buy three to six HSMs for availability and resilience. With a single HSM you risk losing the data. For more information on EKM read the Gartner research note Select the Right Key Management as a Service to Mitigate Data Security and Privacy Risks in the Cloud.

Privacy-Enhancing Computation (PEC) Techniques

PEC techniques comprises several technologies that protect data while it is being used in potentially hostile environments, such as computing clouds or in data ecosystems (also see section “Infonomics”), to enable secure data processing and data analytics.

Hardware techniques, such as trusted execution environments, provide secured environments in which data can be processed (this is also referred to as “confidential computing”). Software-based techniques rely on transformation of the data and/or algorithms such that the original data cannot be determined, but computations and analytics are still possible. This includes homomorphic encryption (HE), secure multiparty computation (SMPC) and zero knowledge proofs (ZKP), for example private information retrieval (PIR) and private set intersection (PSI).

Key use cases for PEC techniques are:

- *AI model training and sharing models with third parties.*
- *Usage of public cloud platforms amid data residency restrictions.*
- *Internal and external analytics and business intelligence activities.*

Current State

As of the beginning of 2021, a mix of data security controls and products make their own decisions, according to policies that are configured in the tools themselves — security tools running in a silo. The use of different vendors, for example, for data discovery, data masking, tokenization, DAM, multicloud DAM functions, and separate teams and roles for each data security control have created a complex and unmanageable collection of vendors, product architectures and consoles. In most cases locking away your data rather than securing it well enough to make it an asset class ready to multiply its wealth through commercialized digital markets (see Table 2).

Table 2: Data Security Current State

(Enlarged table in Appendix)

Current State	Description
Complex administration using disparate products that frequently have no API integration.	Data security products' APIs expose some functionality, but are used mostly for integration or user experience connectivity, and act as veneers on otherwise still siloed controls.
Rudimentary or nonexistent sensitive-data visibility and control.	Some offer data discovery capabilities, others partner, while others offer only basic data discovery based on, for example, the column metadata of relational databases. Very few offer sensitive data discovery across data silos and locations, for example on-premises data stores and in the cloud.
Lack of hybrid cloud capabilities. DSPs that are partially still not ready for cloud or solutions that lack support for on-premises data stores.	Many products lack consistent support for hybrid cloud architectures. Products' line support is handled either in the public cloud or via on-premises data centers. Products that can consistently support data security in public clouds, private clouds and on-premises; data centers are rare.
Very long planning and implementation cycles. The implementation of a new data security product takes one year or longer. Eventually it delivers only part of what was envisioned.	Clients need to bring in place complex (product) architectures and processes before they can configure the first policies or reports. Meaningful proof of concepts have become rare. Vendors frequently shore up prices through feature licenses or complex licensing constructs that no mere mortal can understand.
Immature or nonexistent capabilities for privacy enhanced computation technologies (PECT).	Established vendors frequently do not have the adoption of PECT on the roadmaps of their DSP, whereas newer products from startup companies frequently have PECT capabilities included. On either side, the benefits, scope and limitations of PECT are not communicated well making it difficult for clients to adopt them.
Not all vendors currently address the full set of required and recommended DSP capabilities illustrated in Figure 1.	Some DSP offerings only focus on adapting traditional product platforms rather than replatforming their data security capabilities entirely. DSPs sometimes provide only a partial set of capabilities or have cobbled them together from different acquisitions; multiple control centers, appliances, gateways or agents may be required.
Pattern-driven data discovery capabilities that find patterns or well-known identifiers without considering the context.	Data discovery rarely makes use of AI/ML support to do semantic analysis to find out what something really is. For example, depending on the context, a date can be a date of birth, a transaction date or the dateline of a newspaper article. Each data type will need the appropriate level of protection.
Frustrating amount of false positives making it close to impossible to understand what is really happening to your data.	Since DAM was conceived as stand-alone technology, the created audit trails remain to be mostly false positives. No one can really reconstruct what happens to your data by looking at these logs. This contributes to the poor breach detection rate in organizations.
Separate and siloed teams responsible for each data security capability.	Product specialists with vendor training and certifications for any one data security control are working without much interfacing. Gartner finds that clients frequently entertain redundant data security capabilities, for example, multiple data catalogs or multiple data masking capabilities. While many data security and privacy decisions have enterprisewide impact, they are rarely based on the insights of cross-functional teams.

Source: Gartner (September 2021)

Gap Analysis and Interdependencies

Supporting DSPs and the ever evolving data security, compliance and data sharing needs requires that SRM leaders shed the familiar 10% incrementalism. The most significant gaps that will inhibit DSP migration include:

- **Organizational silos and existing investments** — The status quo of siloed data security is simply not scalable and does not support the pace of digital business. A DSP implementation requires a coordinated and cohesive approach across data security teams, compliance staff and data scientists. For midsize enterprises, this is an easier problem to address, as separate roles may not have been flashed out. Within large organizations, these organizational structures, budgeting processes and responsibilities are quite rigid.

- **Semantic sensitive-data visibility and control** — This is a high-priority capability. Most of the vendors converging on the DSP opportunity have some sort of data discovery capability included. However, large gaps remain because the existing capabilities are examining data stores either for metadata only or for well-known identifiers, and are not finding out what something really is. For example, if a data classification tool finds a date, then it does not know whether it is a date of birth, a transaction date or the dateline of a newspaper article. Effectively making data discovery useless in complex environments such as large organizations. Semantic data discovery must be delivered natively by the DSP offering, and provide options for where the sensitive data is protected, such as data masking or tokenization.

- **Composable architecture** — Monolithic appliance and virtual appliance architectures need to be migrated into composable DSPs. Individual controls are deployed where they are most needed, in a manner that is composable, scalable, flexible and resilient — rather than every security tool running in a silo. The use of cloud-delivered DSPs and data security as a service (DSaaS) provided components may impact adoption for some regions and verticals. While it is possible to use a DSP entirely on-premises, Gartner observes that cloud-based deployment models can significantly reduce complexity, Capex and make more data security controls available. Every enterprise has different requirements for data residency and compliance.

- **Paradigm shift from need-to-know to need-to-share** — Traditionally organizations have had a limited view of how they can help business leaders commercialize data — a view focused on the perceived data security and compliance obstacles. All too often best practices, security policies or regulations seem to mandate that you lock your data away, limiting access to employees that have a need to know. However, data is valuable only if it can be shared with audiences who need it, but, to the extent that it is more widely shared, innovative approaches to data security, such as DSP and data ecosystem governance, are required. Although awareness of the value of data is growing, these CDOs and CIOs rarely have the data security knowledge or experience to have their data secured for virtually any use case.

- **Limited number of comprehensive DSP offerings** — At the start of 2021, less than 10 DSP offerings provide all of the core capabilities outlined in this research note. Data security vendors are frequently looking to extend the end of life (EoL) of a portfolio of siloed data security products rather than replatforming their portfolio into a comprehensive DSP. Over the next five years, acquisitions and further market consolidation will address these gaps. As an interim step, even converged DSP vendors that avoid replatforming their legacy product portfolio are being pressured by customers who are more frequently considering cloud-based DSP in RFP.

Migration Plan

Based on the gap analysis above, we propose the following roadmap and action items over the next several years to be used as a template for DSP adoption and migration planning suitable for most enterprises. While a single-vendor approach for providing all DSP capabilities in Figure 2 may be possible, every data security planner must determine the degree of data security convergence that makes sense for their organization, and mapping that across a practical time frame.

The pent up demand for higher levels of data security, and the rapid increase in product capabilities means that by 2024, 30% of enterprises will have adopted DSP, up from less than 5% in 2019.

We have divided the recommendations into high-, medium- and lower-priority sections based on the expected timeline for typical enterprise DSP adoption.

Higher Priority

In the next 18 to 24 months:

- Use a two-pronged approach and include the benefits of DSP to:
  - Streamline your current data-centric security architecture and overcome known pain points and obstacles. Form a joint D&A, compliance and security team to develop a three to five year roadmap for DSP transformation of your entire data supply chain.
  - Demonstrate greater return value by enabling data processing and data sharing use cases that have not been possible (yet).
- Engage with D&A leaders to identify where traditional data security hinders the organization to improve data's value. Develop a pragmatic and common vision that will enable "good enough" data security and data utility for all silos and use cases.
- Start with advanced controls. DSPs that focus on advanced controls generally have broader coverage than DSPs focusing on a traditional control. For example, DSPs that offer tokenization or (field-level) encryption capabilities frequently include data masking (as a post processing step for the token), cell level authorization and database activity monitoring, whereas DSPs that focus on DAM rarely include advanced controls.

- Cut costs and reduce complexity by consolidating vendors when renewing data discovery, tokenization and data masking. All three are commonly offered now by DSP and are good starting points. Evaluate single vendor offerings, ideally including DAM and PECT.
- Prioritize the consolidation of DSP architectures that consolidate several components into, for example, API-based approaches, cloud services or at a minimum, single agents. Set a three- to five-year goal to replace 90% of stand-alone data security controls.
- Capitalize on external expectations. External expectations, for example compliance requirements or legislation, are opportunities to accelerate the deployment of DSP for the data (stores) in scope of the compliance requirement or legislation at hand.
- Identify data security capabilities that are not fit for purpose in your environment. Examples are data security controls that:
    - Form control silos, which are stand-alone products that do not integrate well with your remaining controls or products.
    - Are data security products that do not deliver the anticipated value and have evolved into checkbox controls rather than actively contributing to data security goals. Selected examples are:
        - DAM implementations that do not give you insight into what really happens on your databases because they deliver several thousands of non-relevant log entries per minute.
        - Automated data classification tools that constantly misclassify or mislabel your data to the extent that no one takes the data classifications seriously any more.
        - Data classification tools that do not classify data because the implementation required so many components and dependencies that it has never been finished.

- Update your data security policies and DSG frameworks to ensure that they are not stuck in the stone age. Reassess the efficacy and benefits of the policies, processes and standards in place. For example, outdated policies are generally written around the “need to know” principle that is locking away your data by default. However, only data that can be processed and shared can unfold its value. When adopting DSP you will want to maximize collaboration and monetization of data through secure data sharing, and use a “need to share” principle instead.

Medium Priority

Over the next 24 to 36 months (note that the recommendations in this section may be accelerated to coincide with, for example, privacy goals, cloud migration projects or data lake projects for AI), enterprises should:

- Retire and replace the data security controls that have been identified as being not fit for purpose by selecting adequate DSP capabilities. Limiting your selection to only equivalent controls and capabilities could unnecessarily hold you back.
- Define the processes and technical requirements used to support a policy, once it has been established. Standards or guidelines should detail the business, technical and security requirements to be considered and describe how this is to be reflected in the DSP. Processes and procedures should define how the DSP will be implemented, maintained and monitored.
- Benefit from capabilities that were formerly only available in one silo but can now be extended across several silos. Examples are data classification and governance, data access governance and DLP.
- Enhance their data risk assessment and use DSP integration points to bring metrics together. Many DSP will already have a centralized data risk dashboard based on vendor-defined metrics. Although certainly not perfect, these metrics are a great input into both regular data risk assessments and data risk analytics.

Lower Priority

- Look for opportunities to use DSaaS as a “data bank.” Gartner observes that DSaaS offerings are evolving to both take and secure your data and share it for you. Very similar to the model of traditional banks that take your money and secure and pay it for you based on your requirements. This can very well be the next transformational step.

Evidence

1. As more cybersecurity point solutions come to market, SRM leaders have reached the critical point for vendor integration and management. They must rationalize their information security portfolio to determine if a consolidation strategy or best-of-breed is the right approach. The 2020 Gartner Security & IAM Solution Adoption Trends Survey finds that:

- Eighty-five percent of organizations currently pursuing a vendor consolidation strategy show a flat or increased number of vendors in the past year.
- Vendor consolidation strategies require at least two years to yield meaningful results.
- About one-quarter of organizations are pursuing a vendor consolidation strategy now, but an additional one-half of the organizations surveyed plan to do so in the next two to three years.

Gartner conducted the study online during March and April 2020. The study's aim was to learn which security solutions were benefiting organizations and which factors were affecting organizations' choice of such solutions. The 405 respondents came from North America, Western Europe and the Asia/Pacific region. Their companies were in different industries and had annual revenue below $500 million. Respondents were at manager level or above (excluding the C-suite) and had been primary involvement in, and responsible for, risk management.

Gartner analysts developed the study collaboratively with the Research Data and Analytics team that follows security and risk management.

2. In 2020 multiple acquisitions and announcements demonstrated vendor interest in building out DSP products:

- IBM signed a reseller agreement with 1touch.io to offer discovery and classification features within the IBM Security Guardium portfolio.
- Imperva acquired jSonar to extend the reach and integration possibilities of their DSP.
- Informatica acquires GreenBay Technologies gaining AI capabilities for example for data governance.

- Netwrix acquired Stealthbits adding DAG and SQL protection capabilities to their DSP.
- PKware acquired Dataguise gaining data discovery and masking capabilities for their DSP.
- Varonis acquired Polyrize to expand their data security capabilities to SaaS applications.

3. In 2019 Thales finalized its acquisition of Gemalto and, with it, the former Vormetric — now building out the CipherTrust Data Security Platform.

Recommended by the Authors

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

Three Critical Use Cases for Privacy-Enhancing Computation Techniques

Hype Cycle for Data Security, 2021

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

Security of Applications and Data Primer for 2021

Top Strategic Technology Trends for 2021: Cybersecurity Mesh

Security Vendor Consolidation Trends — Should You Pursue a Consolidation Strategy?

Table 1: DSP Future State

Future State	Description
Consistent visibility on (sensitive) data, data stores, policies and applicable regulations.	Sensitive-data visibility and control is a critical capability of DSP. This is enabled using a consistent overview of data silos, data security policies, controls and the applicable regulations. The best DSPs will have semantic capabilities for data classification — judging what something really is, rather than relying on preconfigured identifiers.
High levels of integration capability and simplified deployment models.	The use of API integration and cloud-delivered services will increase in importance. DSP provides a number of integration options with popular data stores, such as API integration, agent software or network gateways. Customers can choose the DSP implementation architecture that is least/noninvasive in their environment. Customers can choose to position enforcement points as close to data assets as possible. This will maximize protection, while minimizing user impact.
Ease of administration via a consolidated policy control plane.	The DSP management control plane is decoupled from data types and control objects, allowing centralized administration. The administrative interface will allow data security policy to be managed from a single console and applied regardless of the data silo or the required control objective. AI and ML will be integral to automate policy creation. Full API enablement allows automation and integration with existing processes and tools.
Democratized access to the technologies required to secure data thoroughly.	DSPs are available as stand-alone tools and cloud-based service offerings. Cloud-based offerings will make most of the DSP security objects available via low threshold API integration, making best in class data security controls achievable and affordable for many.

Roles moving from being focused on a single product or technology to being multidimensional.	DSP engineers and administrators live at the intersection of securing data and sharing data. The primary roles of DSP engineers are to secure data well enough so that it can be used as required (without compromising security), instead of locking data away by maintaining a single product-driven data security control.
Data stores for AI and ML, for example data lakes, will increase in importance.	The use of identifiable and regulated data for AI model training raises security and compliance concerns around personal data usage and the potential negative effects in production deployments. Gartner expects DSPs to make an important contribution to the mitigation of those. For example, providing data security for next-generation use cases and enabling migration to public clouds and providing scalable and reliable controls that can support immense data volumes.

Source: Gartner (September 2021)

Table 2: Data Security Current State

Current State	Description
Complex administration using disparate products that frequently have no API integration.	Data security products' APIs expose some functionality, but are used mostly for integration or user experience connectivity, and act as veneers on otherwise still siloed controls.
Rudimentary or nonexistent sensitive-data visibility and control.	Some offer data discovery capabilities, others partner, while others offer only basic data discovery based on, for example, the column metadata of relational databases. Very few offer sensitive data discovery across data silos and locations, for example on-premises data stores and in the cloud.
Lack of hybrid cloud capabilities. DSPs that are partially still not ready for cloud or solutions that lack support for on-premises data stores.	Many products lack consistent support for hybrid cloud architectures. Products' line support is handled either in the public cloud or via on-premises data centers. Products that can consistently support data security in public clouds, private clouds and on-premises; data centers are rare.
Very long planning and implementation cycles. The implementation of a new data security product takes one year or longer. Eventually it delivers only part of what was envisioned.	Clients need to bring in place complex (product) architectures and processes before they can configure the first policies or reports. Meaningful proof of concepts have become rare. Vendors frequently shore up prices through feature licenses or complex licensing constructs that no mere mortal can understand.
Immature or nonexistent capabilities for privacy enhanced computation technologies (PECT).	Established vendors frequently do not have the adoption of PECT on the roadmaps of their DSP, whereas newer products from startup companies frequently have PECT capabilities included. On either side, the benefits, scope and limitations of PECT are not communicated well making it difficult for clients to adopt them.

Not all vendors currently address the full set of required and recommended DSP capabilities illustrated in Figure 1.	Some DSP offerings only focus on adapting traditional product platforms rather than replatforming their data security capabilities entirely. DSPs sometimes provide only a partial set of capabilities or have cobbled them together from different acquisitions; multiple control centers, appliances, gateways or agents may be required.
Pattern-driven data discovery capabilities that find patterns or well-known identifiers without considering the context.	Data discovery rarely makes use of AI/ML support to do semantic analysis to find out what something really is. For example, depending on the context, a date can be a date of birth, a transaction date or the dateline of a newspaper article. Each data type will need the appropriate level of protection.
Frustrating amount of false positives making it close to impossible to understand what is really happening to your data.	Since DAM was conceived as stand-alone technology, the created audit trails remain to be mostly false positives. No one can really reconstruct what happens to your data by looking at these logs. This contributes to the poor breach detection rate in organizations.
Separate and siloed teams responsible for each data security capability.	Product specialists with vendor training and certifications for any one data security control are working without much interfacing. Gartner finds that clients frequently entertain redundant data security capabilities, for example, multiple data catalogs or multiple data masking capabilities. While many data security and privacy decisions have enterprisewide impact, they are rarely based on the insights of cross-functional teams.

Source: Gartner (September 2021)