DSMM

DATA SECURITY MATURITY MODEL

Data Security Maturity Model

Version 1.0 published February 23, 2023

Table of Contents

Data Security Maturity Model Scope and Purpose **2**
Why Use the Data Security Maturity Model 2

How to Navigate the Data Security Maturity Model **3**

Defining Data Security Program Goals **5**
Identify and Classify 5
Protect 10
Detect 16
Respond 19
Recover and Improve 22

About the C3 Working Group **28**

Appendix 1 **29**
Getting Started with the Data Security Maturity Model 29
Example Scenarios 30

Data Security Maturity Model Scope and Purpose

The Data Security Maturity Model (DSMM) helps organizations protect their data and critical assets by developing a data-centric approach to security. Overall, the DSMM shares a similar structure with some of the most well-known security frameworks in the industry, most notably the NIST Cybersecurity Framework. This will hopefully allow organizations to integrate it more easily into their existing security and privacy efforts and augment their investment in other frameworks.

Why Use the Data Security Maturity Model

Many security models cover key aspects of data security and privacy, but the DSMM brings a uniquely “data-centric” approach not found in other models. While traditional data protection strategies have typically focused on narrowly defined use cases or specific threats, the DSMM aims to help organizations take a broad, more consistent approach to protect any or all of their data. Such an approach is a growing priority for security leaders as enterprise data has become incredibly dynamic both in terms of how it is used and where it resides. No longer sequestered in databases, today’s data is constantly being used, modified, and shared by users over dozens of applications. Data likewise can exist virtually anywhere, including in end-user devices, traditional and SaaS applications, and a variety of other cloud services.

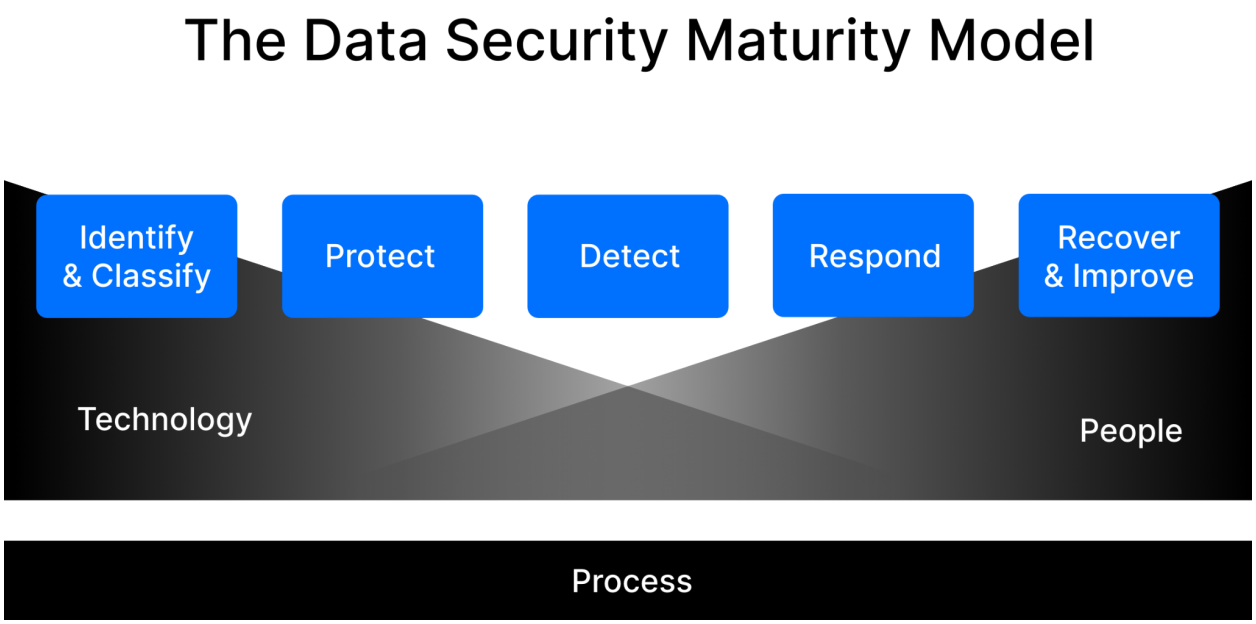
A data-centric approach to security ensures that risk context and policy enforcement can be applied to any data and can follow the data itself wherever it moves or however it is modified. Instead of relying on network-based boundaries, a data-centric approach is able to “follow the bouncing ball” of data, so to speak, without losing visibility and control. This type of approach is essential for organizations to fully leverage the power of their data while also keeping it safe from external threats, insider threats, or simple mistakes that can put the availability, confidentiality, or integrity of data at risk.

Organizations also have access to new types of data security and privacy tools that can make such a data-centric approach far more practical and reliable than ever before. In the past, the limitations of traditional tools often made data security efforts highly laborious and limited to very narrow use cases. A recent wave of innovation is now enabling organizations to extend data security principles to virtually any type of data, anywhere in the enterprise. The DSMM is designed to provide a blueprint to help organizations align available tools and practices to the unique needs and risks of their environment.

How to Navigate the Data Security Maturity Model

The DSMM is organized into five key Functions of a data security program. These Functions are:

- **Identify and Classify**
- **Protect**
- **Detect**
- **Respond**
- **Recover and Improve**



Each **Function** covers multiple underlying **Objectives**. Objectives focus on a particular aspect of security that supports the higher level Function. For example, “Data Discovery” is an Objective within the Identify and Classify Function. Each Objective is addressed at up to three **Levels** of maturity. Each Level includes Practices/Activities that are needed in order to meet the given Level of maturity and include example methods and tools that can be used to implement and fulfill those Practices. These example methods and tools are not intended to be exhaustive, but rather to provide some basic references and guidance for users of the model. Organizations are encouraged to investigate and identify the methods, tools, and processes that make the most sense for their unique needs.

Most Objectives contain three Levels of maturity. Each ascending Level introduces an improvement in the degree to which the Objective is met and/or an increase in the scope of the Objective. For example, Level 1 Discovery may cover manual identification and tagging of data in select locations or databases, while Level 3 would cover the automated discovery of all

sensitive information in all locations. Level 3 Objectives have the broadest scope and level of completeness/sophistication. The use of the word “all” in these levels should not imply that an organization has to be perfect in order to achieve the Level of maturity. Naturally, perfection is rarely possible in real-world security practice. Instead, we use the word “all” to denote that the scope of the Objective is not specifically constrained, such as limiting policy to a certain type of data or a certain location.

Organizations are unique and likewise their appropriate level of maturity will vary based on their unique needs and risks. Readers can refer to the Example Scenarios section in the Appendix for additional guidance on how to pick the appropriate levels of maturity for their organization.

Additionally, Objectives and Practices are quite often interrelated and, therefore, feed into one another. For example, **Identifying** sensitive data will play a key role in **Protecting** data. Likewise, better **Detect** capabilities will often support enhanced **Respond** efforts. The model also supports feedback from later Objectives and Practices to earlier ones. For example, lessons learned from an incident review or detection should likely be incorporated into the policies that are implemented in the **Protect** phase.

As data security is in a phase of rapid evolution, we also encourage active collaboration on the DSMM. Suggestions for updates to the model or inclusion of specific controls or methods can be shared at www.datasecurity.org.

Key Terms

The nomenclature for the DSMM is similar to that found in other, mainstream security frameworks and standards, especially the NIST Cybersecurity Framework and Cybersecurity Capability Maturity Model (C2M2).

Functions provide a high-level, lifecycle-oriented view of an organization’s management of cybersecurity – in this case, focusing specifically on data security and privacy.

Objectives (alternately referred to as capabilities) are cybersecurity outcomes that are closely tied to programmatic needs. They are mid-level achievements that are accomplished by implementing the Practices that comprise them.

Practices are the most fundamental component of the DSMM. Each Practice is a brief statement describing a data security/privacy activity to be performed by an organization. The purpose of these activities is to achieve and sustain an appropriate level of security/privacy, commensurate with an organization’s tolerance for risk and its overarching business objectives. Within each Objective, Practices are organized to progress along a maturity scale.

Defining Data Security Program Goals

The first step in any data security program is to specify the business strategy and goals for the organization.

- Identify the business priorities, and the way in which data security and privacy advance the business outcomes.
- Define the scope of data to be secured: this may be a manual process to determine what data is important to the business. Include any data that, if lost, exposed, or maliciously altered/corrupted, would pose a risk to the organization:
    - Data subject to geographical and industry-aligned regulatory compliance mandates
    - Technical data protected and regulated by export control regulations
    - Personally identifiable information (PII) and similar data that, if mishandled, could adversely impact brand reputation, customer acquisition, or customer retention
    - Customer or partner data that may not be covered as part of regulatory mandates.
    - Data that is operationally needed to produce goods and services.
    - Company intellectual property (IP) and trade secrets.

Identify and Classify

This Function entails a set of core processes for continuously discovering the data that is in scope (i.e., is covered by the data security program); classifying it; and understanding the risk associated with the storage, processing or transfer of data. More specifically, risks include:

- **Data availability**: information being lost or not available
- **Data integrity**: Malicious and intended or accidental altering of data
- **Data confidentiality**: Data being exposed to unauthorized entities

The likelihood of a risk materializing, and its associated business impact depends on the type of data, its location, and whether or not the security team knows about it.

This Function can involve a variety of technical tools as well as more staff-driven processes such as meeting with internal employees and leaders to identify important data for each business unit.

Higher levels of maturity will force organizations to expand their Identify and Classify efforts beyond narrowly defined projects or regulatory requirements. This often requires leadership to take a more open-ended approach to how they assess data security and consider additional types of data or use cases that should be included in the data security program. However, the open-ended approach provides meaningful security and business outcomes by getting to and addressing the root causes of an organization's risk. In many cases, this leads to a reduction in security events and less work for security teams who can then shift from reactive efforts to a more proactive approach.

The following set of questions can help leaders take a more comprehensive approach to their data security and uncover areas of risk that might not be initially obvious.

- What data contains the organization's intellectual property or trade secrets? This may include direct forms of IP such as source code or design files, as well as more indirect sources such as emails, documents, and presentations for product plans, etc.

- What data would cause damage to the organization (financial, competitive, reputational, etc) if it were lost or exposed to the public?

- What data would be more valuable to an attacker than to the organization?

- How does the organization quantify the damage of a breach or data security event?

- How does the organization track the spread of data after it has been accessed? How are derivatives or copies of data identified?

- Is there a consistent approach for discovering/tracking sensitive data across data stores, user devices, applications, and cloud? Can the organization confirm if sensitive data was deleted on a user's device or is sitting in the recycle bin?

- How will the organization automatically classify sensitive data that may not conform to rules and signatures (e.g., non-patterned, non-textual, or other complex types of data)?

- To what extent should ongoing classification efforts be manual or automated? Will staff or end-users be required to perform classification functions?

Data Discovery. The goal for this Objective is to find all data covered by the data security program. Achieving higher levels of maturity requires progressively expanding the scope of coverage, moving from narrowly focused or project-based data security efforts to a more holistic approach that can identify additional types of data critical to the organization. Level 1 will aim to identify data that already exists most of the time, while Level 3 will focus on ensuring data is properly identified and tagged at creation, using automation. For example, a retail organization may include well-known sources of sensitive data such as regulated customer PII as a standard

element of Level 1. Higher levels of maturity could expand the scope of the data program to include data at rest discovery across all internal assets such as user laptops and additional data types such as M&A plans, which may not be initially covered, but could have a massive business impact if compromised. The associated methods/tools also progress from ad-hoc and mostly manual to continuous and increasingly automated.

Levels	**Practices**	**Methods/Tools**
Level 1	Implement focused discovery: often project-based,this process identifies known, well-defined, high-risk data sets (e.g., database with PII).	Selective, reactive process - can include manual processes or automated tools such as tagging or other data detection and response (DDR) tools capable of automatically tracing data based on provenance.
Level 2	Employ expanded discovery based on goals at a business/organizational level. Includes discovery of other types of data important to the company, in additional locations. Encompasses less-structured data such as IP. Full understanding of what data is on which machines or apps, its age, and risk.	Interview business or functional groups to identify data important to the organization. Implement a data store review process into all projects to identify data usage for new initiatives
Level 3	Apply universal discovery, i.e., discovery by default. This continuous automated approach enables discovery of data not previously defined or recognized as sensitive. It gives visibility to drive new proactive policies and decisions.	Use tools to identify newly created databases and data repositories and automatically integrate them into a corporate data inventory. Uses tools such as automated data tracing of all data without relying on staff or user tagging. New data stores are created via

		automation and tagged appropriately.

Data Location Discovery and Context. Finding the “where” of data creation, movement, and usage, including locations such as databases, file sharing apps, endpoints, and personal cloud apps. This objective focuses on answering two critical questions - where is all sensitive data located, and how did it get there? Data can often pass through a variety of entities, such as being shared by multiple users, stored in cloud applications, modified, and shared again. The ability to track this complex journey is an increasingly essential part of seeing and controlling an organization’s true data risk. This may require the organization to develop methods to track how data is shared after it is accessed, including any copies or derivatives.

Levels	**Practices**	**Methods/Tools**
Level 1	Discover project-based data in known locations.	Manual processes (surveys, discussions).
Level 2	Incorporate data found through expanded discovery, in unexpected locations (e.g., data sprawl).	Automated discovery tools (e.g., DLP, shadow IT discovery).
Level 3	Discover all in-scope data in all locations.	Implement data tracing to maintain visibility over data stored in 3rd party systems.

Data Classification. The goal of this Objective is to label/categorize in-scope data in a manner that reflects the relative degree of importance and/or sensitivity of the data to the organization and associated parties (e.g., customers, constituents, or other organizations that are the actual “owners” of the data).

Levels	**Practices**	**Methods/Tools**
Level 1	Classify project-based data in known locations using top-down, organizational classifications (e.g., Public, Internal, Classified). Point-in-time classification.	Manual (tagging, keywords, AIP/MIP labels).

Level 2	Classify organizational in-scope data in all locations on a periodic basis.	Automated classification tools based on content scanning either at rest or in transit.
Level 3	Automatically classify all in-scope data in all locations by default. Unclassified data is treated as a security event/alert, with root-cause analysis.	Automated classification tools with provenance and enterprise context.

Data Risk Assessment. The goal of this Objective is to identify areas of risk, and assess how much damage the loss, exposure, or malicious alteration of specific data elements/sets would cause to the business. The assessment takes into account factors such as data classification, data location/exposure, and the prevalence of threats, as well as the business processes and people/customers/constituents that are potentially impacted.

Levels	**Practices**	Methods/Tools
Level 1	Assess likely impact and probability of data exposure/alteration due to data sprawl, known vulnerabilities and the organization's own experience of threats. Verify compliance of cloud and SaaS vendors.	Manual review of key data data sources and protection policies. May include interviews with internal teams and stakeholders to identify any previous incidents.
Level 2	Perform periodic assessment of risk, based on understanding of the external threat landscape.	Provide a regular review of recent industry data incidents and popular attacker tactics, techniques, and procedures (TTPs) based on news, industry bulletins, and security alerts. Evaluate existing policies and protection mechanisms in context of these risks.

Level 3	Continuous assessment to gain a complete understanding of all dimensions of data risk for all in-scope data. Address data risks posed by trusted insiders at SaaS vendors, partners, and other elements of the extended data supply chain. Audit partner data security policies and controls to verify that sensitive data is properly protected from insider and outsider risks.	Risk and threat modeling and analysis tools. Data Protection Agreements, audit controls, logs, reports, and policies related to the organization's data.

Protect

This Function entails implementing policies and practices to proactively minimize the exposure of important/sensitive data, in particular by controlling how it is accessed, used, and retained. This is a critical Function of the model as it covers the policies and procedures aimed at preventing data-related security events.

The Objectives within Protect will often build on the previous work done in Identify and Classify. While Identify and Classify is designed to help map out an organization's data risk, Protect focuses on mitigating that risk. Most organizations will have multiple types of sensitive data, and the appropriate level of protection can naturally vary based on the type of data or business use case, as guided by the Data Risk Assessment Objective. In fact, a single piece of content may contain multiple types or classifications of sensitive data (e.g PII and PHI), and the organization may need to consider which protection rules should take precedence when protecting data. Ultimately, the Protect function will require organizations to directly consider their tolerance for risk as well as how security controls could potentially impact productivity.

Some key considerations may include:

- How will the organization define appropriate/allowed use for each type of sensitive data? By user, group, location, context, intent, other? Which data will be accessible by which users?

- How will data be protected after an initial access? Can data be copied and shared? What channels, applications or features should staff use when sharing sensitive data?

- How will organizations monitor and control common enterprise data flows (e.g. user to removable media, user to personal cloud, social media, backend SaaS integrations, and other cloud-to-cloud)?

- How will derivatives of data be tracked and protected (e.g,. an encrypted version of a sensitive document, a presentation that embeds a table from a sensitive spreadsheet, an email that contains copied/pasted content from a sensitive internal application, etc)?

- How will the organization address the risk of a user's device or credentials being compromised?

- How will the organization address the risk of user negligence or a malicious insider? For example, are the tools provided for data access built in a way that prevents massive amounts of data from being copied to local workstations?

- Will the policy allow for users to override a blocking decision in order to limit impacts to productivity? If so, how will this be administered and tracked?

- How will data be protected in cases where the data itself is not visible (such as due to encryption or archival)?

- What capabilities will be available to detect and prevent attempts to evade protection policies (e.g., changing file type/extensions, zipping files, etc)?

Controlling Access. The goal of this Objective is to prevent unwanted/unnecessary access to in-scope data. Pursuing the principle of least privilege in this regard is an effective means of reducing the potential for downstream exposure/incidents.

Levels	**Practices**	**Methods/Tools**
Level 1	Enforce ad-hoc policies using native application controls.	Static access control lists based on user identities. Native application access and authorization.
Level 2	Centralized access controls to establish consistent rules around what data can be accessed. Ensure ability to enforce fine-grained entitlements (e.g., all execs are allowed read-only access to company metrics spreadsheet but only the CFO can edit the sheet and see certain tabs.)	Centralized user or role-based access. IAM and user provisioning tools.

Level 3	Dynamic access controls that adapt based on context or risk. Ensure ability to enforce risk-based authorization at user level, based on context of access (e.g., challenge for token code and image recognition when user accesses highest-value data from unknown device.) Provide access to data via applications built to prevent high-volume copying of data.	Dynamic, contextual access controls providing access based on attributes such as: endpoint configuration, patch status and connection origin. Fine-grained entitlements tools. Step-up authentication tools. Data owner approval workflows

Identifying and Preventing Misuse. The goal of this Objective is to go beyond the principle of least privilege to further ensure that in-scope data is not used improperly – that is, in ways that violate policies or regulations, or that otherwise put the data at risk of being unnecessarily exposed. Measures should be taken not only to prevent accidental or intentional misuse or undesired movement of data by those users with a legitimate need to access/use it, but also to ensure that data cannot be used by attackers even if stolen.

Levels	**Practices**	**Methods/Tools**
Level 1	Define clear data handling policies. Ensure violations of the policies are monitored for known sensitive data in known locations and any detected misuse is manually remediated by the security team. Note: data handling policies may restrict locations for storing or channels for transmitting sensitive data, place restrictions on sharing the data inside or outside of the organization, or impose data retention requirements.	Enterprise data loss prevention (DLP), Platform DLP, cloud access security broker (CASB), Insider threat protection/detection — in monitor-only mode.

Level 2	Ensure data handling policies are enforced (i.e., block user upload to personal file sharing app) for all known sensitive data in both known and unknown locations. To further reduce the potential for misuse, ensure that data is tokenized, masked, or encrypted.	Level 1 tools run in enforcement mode, plus tokenization, masking, and encryption.
Level 3	Enforce data handling policies for all data and all of its derivatives (e.g., data converted into other formats or cut/pasted or embedded in other documents) OR prevent the creation of derivatives. Enforce controls for data even after it leaves approved locations or goes outside the organization.	Level 2 plus digital rights management (DRM) and data detection and response (DDR) tools that provide continuous tracing of all data and derivatives.

User Education and Feedback. This Objective focuses on end-user training, coaching, and interaction to reduce risks and increase compliance with established policies. Organizations will need to balance policy enforcement and user productivity, which may require providing users with the ability to override a blocking decision or acknowledge a violation without blocking. Increasing levels of maturity are characterized in part by progressively broader options for enabling user inputs (overrides, policy/config change requests) based on business need and adjusting policies/settings accordingly.

Levels	**Practices**	**Methods/tools**
Level 1	Train users on the established policies for protecting data including what types of data are particularly sensitive and what applications and features are approved for handling sensitive data. Ensure that users acknowledge established policies and are tested to verify that they have adequately understood the training.	Periodic education and training. Can be delivered in a classroom setting or self-paced remote training.

	Enable users with self-approved one-time override of data blocking policies/controls to facilitate and align with necessary business functions/practices.	
Level 2	Implement incident-based training and reinforcement based on policy violations. Document violations and identify corrective actions to avoid future violations. Enable users with workflow-governed override of data blocking policies/controls to facilitate and align with necessary business functions/practices.	Detection of policy violations. Manual follow-up by an end-user's supervisor. DLP tools.
Level 3	Incorporate automated real-time coaching and user training. This includes contextual coaching within the end user's regular workflow without the need for manager intervention. Enable users to provide feedback and justification where policies/settings may need to be adjusted to better align with business goals/practices and/or fine-tune how the organization protects sensitive data.	Real-time training platforms or data security platforms, ideally featuring adaptive response capabilities.

User education: traditional user education is broad-based and generally takes place only periodically – during onboarding, or a few times a year – and covers corporate security policies, anti-phishing, anti-malware, and other diverse categories. However, data exposure through accidental user events are a common cause of security incidents. Studies have proven that one of the most effective ways to change unwanted user behavior is via just-in-time, contextual training, which alerts users in real time when they are engaging in potentially risky behavior. This "adaptive response" allows users to interact, give feedback and learn about the proper data handling requirements. In turn, this can bring down the risk of unwanted data exposure by a factor of twenty.

Data Retention: The goal of this Objective is to reduce the risk of unwanted exposure, by proactively eliminating in-scope data that is no longer needed or out of date, or which data owners have requested be erased/removed from the organization's records. On the other hand, many regulations require records to be retained for defined periods of time (e.g., PCI-DSS, HIPAA, Fair Labor laws, etc). Data retention must balance the need to ensure unnecessary data is not retained while preventing other types of data from being inadvertently deleted.

Levels	**Practices**	**Methods/Tools**
Level 1	Note: Driven by core regulatory requirements and basic best practices. **Delete:** Ensure that legal requirements for data retention and deletion have been identified. Enforce policies to ensure only the minimal amount of personal data is collected and stored in organizationally controlled locations. **Retain:** Build policies to properly archive and protect any regulated and sensitive data to avoid loss or inadvertent deletion.	Data minimization and erasure, consent management tools. Data protection policies, and manual data governance procedures.
Level 2	Apply an automated and enterprise-wide approach to lifecycle management. Manage all copies of the data, not just in the central database. Automatically enforce policies related to retention.	Use of an enterprise document management solution to store and identify potentially outdated data and all copies..
Level 3	Implement fully automated retention orchestration. Automatically enforce retention policies for all locations and for all in-scope data.	Level 2 plus retention orchestration and automated erasure to standards. Data has a lifecycle policy attached to it at creation, ensuring it CAN'T violate retention policies.

Detect

The Detect Function covers the collection and analysis of data to identify data-related security events or policy violations. While the preceding Protect Function focuses on preventative enforcement measures, the Detect Function aims to uncover risks or violations that were not stopped by Protect measures. However, Detect efforts should not be considered purely retrospective. It is critical for organizations to detect risks and threats as early as possible, in order to minimize any impacts.

Detect efforts can include the discovery of risky or anomalous behaviors (e.g., unusual data downloads, risky application usage) as well as direct, first-order incidents (e.g., data leakage, unauthorized copies of sensitive data) related to usage and movement of in-scope data.

Organizations will want to consider a variety of factors with regard to their Detect maturity level including:

- What are the goals in terms of time-to-detection (Near real-time? Hours? Days) and how do they vary based on the defined dataset?

- What data or signals are available for analysis in addition to logs and events captured by Protect controls? User access logs? Security events? Packet/flow capture? Endpoint and host logs?

- How will the organization identify and track behaviors or anomalies at the network or endpoint level? What systems or methods are or will be used to support this effort?

- What additional data or signal is required in order to reliably identify risks and how can these gaps be addressed?

- What methods will the organization use for data analysis (e.g., correlation, data analytics, flow analysis, machine learning, graph analysis, etc.)? What systems or tools will be used for these efforts (e.g., SIEM, analytics platforms, homegrown analysis tools)?

- How much time and effort will be required of staff in order to perform deeper analysis and triage? What is the team's tolerance for false positives and false negatives? How much effort is required to detect false positives?

Signal Collection. The purpose of this Function is to collect data/telemetry that can be used to reveal data security events, policy violations, or anomalies. Telemetry sources should focus on

the access, movement, or modification of in-scope data. They should provide additional context, and/or enable staff to verify events or hunt for additional risks or violations.

Levels	**Practices**	**Methods/tools**
Level 1	Collect logs and alerts, including: ● Content attribute matches (content) ● User access logs (user) ● Violation events	Manual triage of logs and incident invocation (where applicable)
Level 2	Record data movement events and behaviors for priority enterprise applications (application and user actions). This could include recording user actions when using data in a cloud application or on a host device such as editing or renaming a file, copying/pasting data, uploading/downloading, etc. Record data lineage across user actions and behaviors to maintain context. Incorporate risk and threat data from external sources to identify threat families and techniques targeting data. Collect additional data that can be valuable for analytics and anomaly detection. This may include data not limited to security events (e.g. network logs, network flows, data access and application logs, etc).	DRM, data lineage tools, DDR tools. DLP tools. External threat feeds, risk sources. Other logging tools.
Level 3	Collect all user actions across all applications (including both unmanaged/personal and managed/corporate applications). Record data lineage across multiple applications and locations (e.g., relate data copied/pasted from database to a presentation.)	DDR tools.

Analysis Methods. This Objective is based on a range of increasingly sophisticated and automated techniques and technologies to process/analyze collected signals for the purpose of uncovering data movement or modification that represents a threat to data security or privacy.

Levels	**Practices**	Methods/tools
Level 1	Implement signature- and rule-based detection models based on pattern matching (regular expression) or defined metadata or tags. Use these techniques to detect the movement of sensitive data and to find previously undiscovered data in unsanctioned locations.	IDS, DLP, File system auditing, CASB, SIEM/automated log review.
Level 2	Implement data analytics and log correlation of multiple data sources including data from applications, users, and systems. Leverage multiple data sources to baseline activity and identify event-based anomalous behavior (e.g., unusual download sizes). Note that these practices provide limited coverage of evasion techniques (encryption, file format conversion, etc.) used. Use of “honey” data, which are fake data entries that are never to be used, to detect malicious or accidentally policy violating data movement.	NDR, UBA, SIEM, SOAR, data tracing, honey data.
Level 3	Perform advanced analysis based on full enterprise lineage of data including data provenance, all application and user behaviors. Apply AI models to identify potentially malicious actions based on multiple dimensions of user behaviors.	AI/ML/deep learning, graph analysis

Respond

This Function focuses on immediate, short-term actions to be taken upon detection of a potential incident. Key objectives include validating and establishing the scope of the incident, taking steps to minimize/halt the impact, and maintaining communication with business stakeholders and other affected parties.

For organizations that already have solid processes, procedures, and toolsets in place to respond to cybersecurity incidents, those capabilities will remain applicable and, ideally, should be re-used extensively. Adjustments and augmentation will almost certainly be needed, however, to increase the focus on in-scope data – effectively putting it at the center of many of the response activities and decisions being made. Higher levels of maturity will result primarily from increasing levels of automation (i.e., efficiency, accuracy, and speed) for all aspects of the Function: incident validation, scope and impact assessment, and mitigation.

The following questions can help initiate the shift to a more data-centric approach to incident response.

- Is there a clear definition/plan identifying all stakeholders (both internal and external) and other parties (e.g., legal counsel) that should be involved when responding to an incident involving in-scope data? Does it include those who should be notified (and when) regarding pertinent details?
- Is there a clear understanding of the requirements (and possibly restrictions) stemming from applicable data protection laws and regulations that need to be accounted for in your response plans and activities?
- Do the tools already in place to support incident response provide sufficient visibility of and control over the data itself?
- Are the steps typically being taken to mitigate an incident effective down to the data level?
- Has sufficient consideration been given to the need to find a balance between maintaining essential business operations/functions and the impact of potential options for mitigating an incident (e.g., blocking a user or locking down a data repository)?
- Is the type of data that is being impacted a consideration when it comes to the potential participants in the response process, as well as those who will be informed of the relevant issues and outcomes?

Triage and Mitigation. The purpose of this Objective is to establish the magnitude of the data security incident and take appropriate steps to minimize impact. Related practices apply not only to incidents involving internal/insider misuse of data, but also those involving external/malicious exfiltration of data.

Levels	Practices	Methods/tools
Level 1	Incidents are reported manually by customers, employees, service providers, or others. Incident response plans include specific steps for identifying the data impacted and tailoring response and reporting requirements. Incident responders have access to a mapping of data source to data owner and owners are trained on their responsibilities during an incident. Mitigate by blocking access or otherwise stopping unwanted activity; primarily manually initiated.	Log management tools, SIEM. Manual efforts for validating scope and assessing impact. Ability to manually suspend/revoke access to network and individual systems.
Level 2	Incidents are reported manually, but may also be detected automatically by rules-, signature-, or behavior-based tools. Automated correlation tools enrich data to help analysts quickly identify impacted systems, users, and data. Automate user/device-centric mitigation and containment, such as: ● Blocking access to data ● Blocking access to external ● Fully isolating/disabling offending user/device Ensure collection and preservation of forensic data.	UEBA, XDR, event correlation. Ability to effectively isolate systems.

Level 3	Execute fully automated IR narrative (i.e., push-button ability to view all details related to scope and impact, along with responses already taken and recommended). Implement orchestrated/automated mitigation responses including user suspension, system quarantine, and leaked data sanitization..	SOAR, Open-XDR, automated incident response tools, DDR.

Communications: This Objective entails engaging and coordinating response activities with applicable business stakeholders, notifying relevant external parties (e.g., partners, customers) of pertinent details, and taking any other actions required for regulatory/legal compliance.

Levels	**Practices**	**Methods/tools**
Level 1	Implement a single, reactive communication plan that is activated during incident response and details participants/members, roles, legal/compliance requirements, communication mechanisms, and timeframes Applies to internal stakeholders (including legal team), affected external parties, law enforcement agencies, and regulatory authorities. Coordinate and communicate response and recovery activities among key stakeholders and affected/essential business partners. Perform incident/breach notifications and updates within prescribed time frames for applicable jurisdictions (i.e., regulatory driven) and contractual agreements (i.e., business driven). Engage and cooperate with relevant authorities on investigative, forensic preservation, and disclosure processes. Manage related public relations activities.	Mostly manual.

Level 2	Implement multiple communication plans (or modules) by type/class of incident. Includes proactive, voluntary information-sharing with key business partners, extended supply chain, unaffected customers/constituents (when deemed appropriate), and infosec community at large to enable greater, collective situational awareness. Reputation is proactively repaired after an incident.	Extensive automation of notification plans and processes. Intelligence sharing portal/platform (i.e., reverse feed).
Level 3	**Reserved for future development.**	

Recover and Improve

This Function encompasses those actions taken not only to restore normal operations (as they pertain specifically to data), but also to build back stronger. An overarching goal for any organization should be to evolve its overall data security program regularly and steadily. Core objectives include data backup and recovery, identifying and incorporating lessons learned, and adaptive user education.

Key questions that help to further frame this Function and serve as a jumping off point for building it out are as follows.

- Does the organization have a policy for data backup and recovery? Does the policy cover all in-scope data?
- How fast is the organization able to restore data in a “worst-case scenario” situation?
- Does the organization have a way (procedural or technical) to control the locations where backups reside and how they are able to be used and/or transferred?

- Does the organization know the retention period for the backups and logs captured by their security/DLP solution?
- To what extent does the organization examine the cause of incidents affecting in-scope data, and subsequently take action to prevent them from happening again? Does the organization look for proximate cause only, or perform root-cause analysis? Does it implement proximate remedy only, or systematic fixes that extend coverage to prevent an entire class of incidents across all systems that could be affected?
- Does the organization have a policy/process to periodically reevaluate its data security program from top to bottom?
- Does the organization have cyber insurance coverage and a policy or process to periodically reevaluate it?
- How are users made aware of data security and privacy policies? How are they trained on those policies? To what extent are they able to override those policies, or request/recommend changes to them in order to align with necessary business activities/functions?
- What mechanisms are in place to take advantage of the data security practices/knowledge of peer organizations or the security industry/community at large?

Data Backup and Recovery. With this Objective the goal is to ensure that in-scope data is regularly backed up and capable of being restored when needed. In this case, what is being protected – or more accurately preserved – is not the data itself, but rather the business processes that rely on it.

Levels	**Practices**	**Methods/tools**
Level 1	Define sanctioned backup tools and locations. Ensure backups of in-scope data are conducted, maintained, and tested. Ensure retention of multiple versions of data. Restore from backup as needed (i.e., in response to an incident that impacts data integrity or availability). Ensure confidentiality of backups.	Backup and recovery tools.

	Ensure isolation at off-site, offline storage locations.	
Level 2	Prevent backups to non-approved locations or applications (e.g., employee's personal backup). Control backups on a per-file and per-account basis (e.g., prevent users from backing up data to personal Dropbox instead of corporate location). Ensure backups are encrypted, and that encryption keys are stored in a redundant, non-deleteable way, with strict access control. Ensure and verify retention of data in cloud/SaaS applications. Ensure compliance and ability to delete the data in cloud and SaaS backups based on user request or regulatory requirement. Establish back-up policies based on specific use cases (e.g. data relevant to a legal case may require raw data to be preserved for an extended period of time). Back-ups of IaaS environments is done towards separate IaaS accounts, using system accounts that have write-only privileges, to prevent even a compromised administrator account from deleting the data and backups at once.	Backup and recovery tools.

Level 3	**Reserved for future development**	

Incident Review / Lessons Learned. For this Objective, the intent is to assess both the cause and handling of data security incidents, for the purpose of identifying areas for improvement within each of the other Functions. Findings and corrective measures are then incorporated in relevant Functions to avoid repeat episodes of similar incidents and strengthen data security/privacy posture overall.

Levels	**Practices**	**Methods/tools**
Level 1	Perform incident-specific review & adjustments for high- and medium-severity incidents. Adjust/fix policies, practices, and configuration settings of technical countermeasures to prevent occurrence of similar incidents going forward.	Primarily a manual exercise. SIEM.
Level 2	Extend incident-specific review and adjustments to cover all severity levels. Perform root-cause analysis for high/medium severity incidents to uncover and fix upstream issues that caused the associated data to be at risk in the first place (e.g., role sprawl; insufficiently granular roles or policies; etc.)	Technology-based facilitation of appropriate adjustments/fixes (e.g., system proposes changes, but operator reviews and manually implements). May include SIEM, SOAR, DDR, and NDR tools, or other security or network management tools.

Level 3	Extend scope for root cause analysis to cover all severity levels. Ensure adjustments/fixes are applicable enterprise-wide and are inclusive of all incidents of similar type/class, other root causes, and across other apps/systems. Example: after analyzing a data violation from a user sending data over personal GMail, the company may enforce policies to distinguish personal vs corporate GMail and also apply those lessons to other applications such as Dropbox. Reassess implementation of security model and adjust controls and target levels as needed.	SIEM, SOAR. Orchestration / automation of appropriate adjustments/fixes.

Collaboration and Research. This Objective involves taking additional steps, both within and beyond the organization, to continuously improve its data security practices. The purpose is to account for macro-level changes not only to the organization and how it operates, but also those affecting the broader threat, security technology, and business landscapes.

Levels	**Practices**	**Methods/tools**
Level 1	Maintain constant communication with all key functions of the organization and adapt data security objectives to evolving business goals. Monitor and account for ongoing changes to the threat landscape, e.g., by leveraging advanced threat intelligence sources providing insights into the evolution of threats, threat actors, and their tactics, techniques, and procedures (TTPs). Achieve/maintain relevant certifications for internal security team/practitioners, e.g., ISC2 CISSP, ISACA CISM/CISA, CompTIA CySA+, EC-council Certified Ethical Hacker (CEHv11).	Quarterly data security/privacy reviews with line-of-business leaders. Customized threat intelligence feeds.

Level 2	Monitor and account for the emergence of new security technologies and innovations. Share (/obtain) information on threats, practices, and lessons learned with (/from) industry peers and associations.	Primarily manual. Threat intelligence platform; ISAC participation; Automated Indicator Sharing.
Level 3	Investigate, adapt, and adopt concepts and technologies from other markets (e.g., graph analytics).	Primarily manual.

About the C3 Working Group

The Comprehensive Cyber Capabilities Working Group (C3WG) is working to define for the cybersecurity community a comprehensive list of capabilities needed to secure and defend the full range of cyber assets within an organization. Comprised of security leaders from across industries, the group has deep expertise in the people, process, and technology used to solve security challenges.

Chair

Sounil Yu, CISO, JupiterOne

Members

Aaron Stanley, VP of Security, dbt Labs

Arkadiy Goykhberg, CISO, Branch

Brian Markham, CISO, EAB

Chris Hodson, CSO, Cyberhaven

Dan Walsh, CISO, VillageMD

Guillaume Ross, Deputy CISO, JupiterOne

John Sullivan, CSO, Boston Scientific

Kevin Paige, CISO, Flexport

Louis Holt, CEO, ESPROFILER

Merike Kaeo, former CISO, Uniphore

Richard Rushing, CISO, Motorola Mobility

Ross Young, CISO, Caterpillar Financial

Appendix 1

Getting Started with the Data Security Maturity Model

The Data Security Maturity Model helps organizations protect their data and critical assets by developing a data-centric approach to security, ensuring that risk context and policy enforcement follow the data no matter how it moves or is modified. This ability to protect any type of data across devices, applications, and cloud assets is essential if organizations are to take advantage of the power of modern collaboration and digital transformations without exposing their data to external threats, insider threats, or simple mistakes.

The DSMM is organized into five key Functions – **Identify and Classify**, **Protect**, **Detect**, **Respond**, and **Recover and Improve**. Each Function covers multiple underlying Objectives, which focus on a particular aspect of security that supports the higher level Function.

While the DSMM provides the details of each Function, this Appendix provides a guide to the overall process so that organizations can get started easily and improve their data security posture quickly.

1 - Assess Organizational Data Needs
Every organization is unique both in terms of data real estate and tolerance for risk. The Data Security Maturity Model is designed to adapt to the needs of each organization. The appropriate maturity level for each objective will vary based on each organization's unique risk profile.

However, many data security programs have long been constrained by the limitations of old-school security tools – often limiting data security to narrowly defined DLP and regulatory use cases. Recent innovations in data security technologies enable organizations to apply data security controls far more universally than ever before, including the ability to protect any data, in any location, and at any time. As such, it is important - from both a security and a business point of view - to take a fresh look at the organization through the lens of its data. This means evaluating all of the organization's data assets in terms of the impact of that data being lost, exposed, or otherwise misused or rendered unavailable. As a result, organizations may want to take the following steps:

- Collaborate with all business units to identify the data critical to each group or operational function.
- Identify all intellectual property and trade secrets regardless of data or file type.
- For each type of data, assess the competitive, financial, and reputational impact if the data were exposed or unavailable.

2 - Identify Target Data Security Levels and Identify Gaps
The assessment done in Step 1 is the foundation for using the model, since it answers the question of precisely what data is in scope for the model. Based on this assessment, organizations will next address each of the key Functions of the maturity model. For each Function, several Objectives are presented, containing (generally) three Levels of maturity. Each ascending Level introduces an improvement in the accuracy and reliability of meeting the Objective and/or an increase in the scope of the Objective. For example, Level 1 Discovery covers manual identification and tagging of data in select locations or databases, while Level 3 calls for the automated discovery of all sensitive information in all locations.

It is important to note that organizations will likely have different target levels of maturity for various types of data and use cases. For example, source code or product designs may warrant Level 3 for the “Preventing Misuse” objective, while internal HR data may only require Level 2. Refer to the Example Scenarios section below for more detailed examples of how to align target levels of data security maturity to the needs and risks of an organization.

Once the appropriate targets are defined, the team will need to evaluate their existing security processes and tools to identify potential gaps that will be addressed in the following step.

3 - Develop and Implement the Data Security Plan
Based on the previous analysis, the organization can build a coordinated plan to reach the needed Level of maturity for each data type or use case. For each Objective and maturity Level, the DSMM identifies methods and tools that can be used to achieve the target goal. However, these are provided as examples only, and should not be considered an exhaustive list or the only ways that an organization can achieve the appropriate level of maturity.

Teams should evaluate any newly implemented controls to ensure they are functioning properly and delivering the desired effect.

Once deployed, organizations should continue to monitor the efficacy of the program and regularly re-evaluate their data security program to adapt to changing workflows, data usage, and business needs.

Example Scenarios

Naturally, not every organization will need to be at Level 3 or even Level 2 for all Objectives. The appropriate Objectives should be defined based on each organization’s business needs and tolerance for risk. An organization can reasonably be content at Level 3 in one area and at Level 1 in others. The following scenarios provide some examples of how an organization’s need might align to various target levels of maturity.

Scenario #1: Online health service provides health and genetic testing by mail, and provides a centralized portal for customers to review their results and recommendations. Due to the sensitive nature of patient and regulatory requirements (e.g., HIPAA, PCI, GDPR), protecting customers is the #1 priority for the firm.

In this case, the primary dataset that the firm will focus on is quite well-defined. Since the data is relatively well-known, data discovery and classification may not be a top priority. The organization may start with a target of Level 1 for Identify and Classify or may need to target Level 2 to address any cloud-to-cloud sharing or to identify potential misuse of sensitive data by the firm's internal administrators. However, the ability to protect data and respond to incidents will be a clear priority and the organization will target Level 3 for these functions.

- **Identify and Classify -** Level 1 or 2
- **Protect -** Level 3
- **Detect -** Level 2
- **Respond -** Level 3
- **Recover and Improve -** Level 1 or 2

Scenario #2: Technology Integrator and Services firm develops and implements a variety of proprietary software solutions for its clients. Each project is highly customized and confidential and tied to specific dedicated teams. Projects can also be relatively dynamic as team members regularly rotate on and off a project when needs change. The firm deals with a wide range of sensitive data including source code, product designs, and development plans, as well as sensitive customer communications.

In this case the firm must protect a wide range of datatypes in which many users will need to interact with the data. This will require the organization to target Level 3 of **Identify and Classify** in order to identify many types of data that will naturally be distributed across many endpoints, applications, and cloud services. **Protect** and **Detect** efforts will also be at Level 3 in order to prevent the loss of sensitive data either due to external threats, malicious insiders, or internal errors. However, achieving Level 2 for **Respond** and **Recover** may be acceptable based on the organization's risk tolerances.

- **Identify and Classify -** Level 3
- **Protect -** Level 3

- **Detect -** Level 3
- **Respond -** Level 2
- **Recover and Improve -** Level 2

www.datasecurity.org