SNIA
Advancing storage & information technology

Storage Networking Industry Association

Storage Security: Data Protection

Technical White Paper

March 2018

Abstract: *The ISO/IEC 27040:2015 (Information technology - Security techniques - Storage security) standard provides detailed technical guidance on controls and methods for securing storage systems and ecosystems. This whitepaper provides an overview of data protection and the associated guidance in the standard.*

USAGE

DISCLAIMER

The information contained in this publication is subject to change without notice. The SNIA makes no warranty of any kind with regard to this specification, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The SNIA shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this specification.

Suggestions for revisions should be directed to http://www.snia.org/feedback/.

Revision History

Revision	Date	Sections	Originator:	Comments
V0.1	*9/13/2015*	All	Eric Hibbard	Initial Draft
V0.2	*12/14/2015*	All	Eric Hibbard	Incorporation of ISO/IEC 27040 summary
V0.3	*7/17/2017*	All	Eric Hibbard	Alignment with DPCO whitepaper and initial draft of SNIA elements
V0.4	*10/10/2017*	4	Eric Hibbard	Data classification, data authenticity, due care, and retention/disposition
V0.5	*12/18/2017*	4	Eric Hibbard	Retention separated and expanded
V0.6	*1/9/2018*	4	Eric Hibbard	Retention and preservation; archives
V0.7	*1/29/2018*	All	Eric Hibbard	Review draft
V1.0	*2/20/2018*	All	Eric Hibbard	Final Approval Draft (for ballot)
V1.1	*3/6/2018*	All	Eric Hibbard	TC Approval Draft
V1.1	*3/8/2018*	All	Arnold Jones	Approved by SNIA Technical Council

Suggestion for changes or modifications to this document should be submitted at http://www.snia.org/feedback/.

Foreword

This is one of a series of whitepapers prepared by the SNIA Security Technical Working Group to provide an introduction and overview of important topics in ISO/IEC 27040:2015, *Information technology – Security techniques – Storage security*. While not intended to replace this standard, they provide additional explanations and guidance beyond that found in the actual standard.

Table of Contents

Executive Summary

Data protection is an essential element of storage security that can be nuanced, depending on industry requirements (e.g., storage, security, and privacy). This can be seen in the ISO/IEC 27040 (Storage security) standard, which while not directly addressing data protection, does identify relevant security controls. To raise awareness of data protection, this whitepaper highlights the relevant data protection guidance from ISO/IEC 27040 and then builds upon it, covering topics such as data classification, retention and preservation, data authenticity, and data disposition. As part of this expanded material, SNIA provides guidance and considerations that augment the existing storage security standard.

1 Introduction

Storage security is concerned with the application of physical, technical and administrative controls to protect storage systems and infrastructure as well as the data stored within them. The ISO/IEC 27040:2015, *Information technology – Security techniques – Storage security* standard (henceforth referred to as ISO/IEC 27040) identifies a broad range of storage security controls, including those specific to data protection.

The term "data protection" means different things to different audiences within their respective venues. This ambiguity can lead to serious misunderstandings and lead in turn to data indiscretions involving severe consequences such as data breaches (with or without exfiltrations), financial liability, and regulatory scrutiny, etc.

This whitepaper explores multiple facets of data protection, summarizes the data protection guidance from ISO/IEC 27040, and provides additional SNIA recommendations.

2 Facets of Data Protection

This section explores three different perspectives of data protection: storage, information assurance/security, and privacy. Each facet of data protection is briefly described and includes a review of the differences and similarities.

2.1 Storage

The 2017 SNIA Dictionary defines data protection as:

[Data Management] Assurance[1] that data is not corrupted, is accessible for authorized purposes only, and is in compliance[2] with applicable requirements.

The SNIA data protection taxonomy[3] elaborates on this definition by describing data protection as:

Data Protection means assurance that data is not corrupted, is accessible for authorized purposes only, and is in compliance with applicable requirements. Protected data should be usable for its intended purpose. Usability may require that steps be taken to provide data integrity, application consistency, versioning, and acceptable performance.

This definition of data protection goes beyond the notion of data availability, defined as the amount of time that data is accessible by applications during those time periods when it is expected to be available. Unacceptable performance can lower productivity levels such that access to applications and related data is effectively unavailable. Note that data security and compliance issues are also intimately involved as the ultimate goal of data protection is to reduce risks, costs, and downtime while increasing business value and agility.

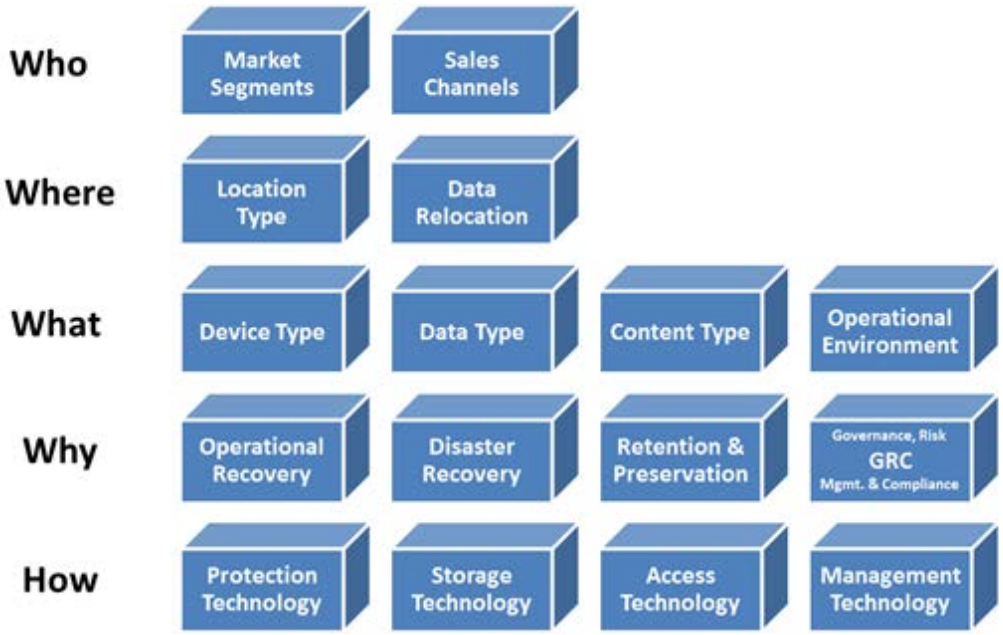


Figure 1. Data Protection Taxonomy

[1] The SNIA Dictionary also defines "assurance" as: "[Data Security] A process for demonstrating that the security goals and objectives for an IT product or system are met on a continuing basis."

[2] The SNIA Dictionary also defines "compliance" as: "1.[General] The state of being in accordance with a standard, specification, or clearly defined requirements. 2. [Legal] The state of being in accordance with legal requirements."

[3] SNIA Data Protection and Capacity Optimization Committee whitepaper, A Data Protection Taxonomy, June 2010

Figure 1 is a high-level overview of the data protection taxonomy. It is represented as boxes which represent distinct lenses through which to view a data protection solution; each lens is independent of every other lens. These lenses are in many instances interrelated, and the taxonomy encourages examination of these relationships.

Each row of boxes in Figure 1 addresses a particular question: who, where, what, why, and how. Each lens categorizes a straightforward notion and both the high-level category, and its subcategories avoid the use of unnecessary jargon.

The *SNIA Data Protection Best Practices* whitepaper[4] expands on the taxonomy by documenting SNIA's position on data protection best practices, as defined by the SNIA's Data Protection & Capacity Optimization (DPCO) Committee. These data protection best practices are organized by the following drivers:

1. Data Corruption / Data Loss
2. Accessibility / Availability
3. Compliance

The DPCO identifies the data protection technologies associated with each driver, references the appropriate existing standards (when appropriate), and finally recommends the best practices for each data protection technology. Of these drivers, as mentioned above, the compliance driver is the one that is probably the most closely aligned with the contents of this whitepaper.

2.2 Privacy

The International Association of Privacy Professionals (IAPP) Glossary[5] provided the following relevant definitions:

> **Privacy:** ***The appropriate use of personal information under the circumstances.*** What is appropriate will depend on context, law, and the individual's expectations; also, the right of an individual to control the collection, use and disclosure of information.
>
> **Data Protection:** ***The management of personal information.*** In the United States, "privacy" is the term that is used in policies, laws and regulation. However, in the European

[4] *SNIA Data Protection Best Practices* whitepaper, SNIA Data Protection and Capacity Optimization (DPCO) Committee, October 2017

[5] International Association of Privacy Professionals (IAPP). IAPP Information Privacy Certification Glossary of Common Privacy Terminology. 2011. Web PDF file listed as "CIPP Glossary of Terms," https://iapp.org/media/pdf/certification/CIPP_Glossary_0211updated.pdf

Union and other countries, the term “data protection” often identifies privacy-related laws and regulations.

It is worth noting that the current Web version of the IAPP Glossary has removed both of these terms.

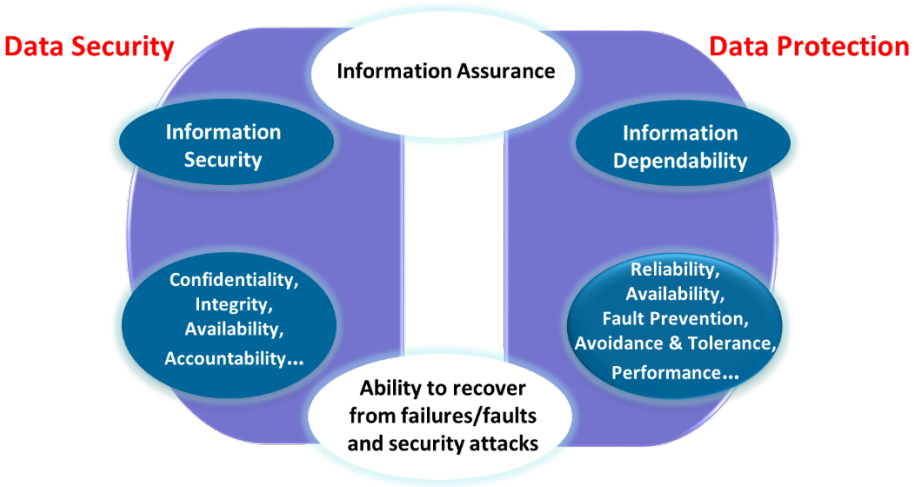
2.3 Information Assurance/Security

In contrast to the storage view of data protection, information security[6] tends to focus on the confidentiality, integrity, and availability of data. As an example, ISO/IEC 2382:2015 (*Information technology -- Vocabulary*) defines data protection as:

> *implementation of appropriate administrative, technical or physical means to guard against unauthorized intentional or accidental disclosure, modification, or destruction of data*

This definition has basically remained the same since 1993 (ISO/IEC 2382-1:1993).

Information assurance[7] expands upon the information security elements by adding information dependency elements (see Figure 2).



[6] **Information Security:** The protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability. (Source: SP 800-37; SP 800-53; SP 800-53A; SP 800-18; SP 800- 60; CNSSI-4009; FIPS 200; FIPS 199; 44 U.S.C., Sec. 3542)

[7] **Information Assurance:** Measures that protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation. These measures include providing for restoration of information systems by incorporating protection, detection, and reaction capabilities. (Source: SP 800-59; CNSSI-4009)

Figure 2. Information Assurance: Interaction Between Security & Dependability[8]

Dependability primarily focuses on how to quantitatively express the ability of a system to provide its specified services in the presence of failures, through measures of

- reliability (probability that a system provides its services throughout the specified period of time),
- availability (fraction of time that a system can be used for its intended purpose within a specified period of time),
- safety (probability that a system does not fail in such a way as to cause a major damage), and
- performability (quantitatively measures the performance level of a system in the presence of failures).

It is worth noting that the dependability and security communities remain somewhat separated, but there is recognition that interactions between them is desirable. A simple and often cited difference between the two areas is that dependability focuses primarily on faults and errors in the systems that are typically non-malicious in nature (primarily from the fault tolerance design area), while security focuses mainly on protection against malicious attempts to violate the security goals. In reality, this perceived difference is not accurate because there are significant overlaps and synergies within the two communities, but they are not always recognized.

3 ISO/IEC 27040 Data Protection Guidance

Within ISO/IEC 27040, the concept of data protection is not explicitly addressed. That said, it is possible to get some insight into the standard's guidance for data protection by consulting the ISO/IEC 27040 Index[9] that was published by SNIA.[10]

The ISO/IEC 27040 data protection controls are associated with backup/recovery systems, Continuous Data Protection (CDP), and replication technologies, which are used to ensure data reliability, availability, and resilience. The standard also emphasizes that all the data protection solutions should be viewed as data resilience mechanisms. This focus is more closely aligned with the storage view of data protection than the privacy or information assurance views.

[8] Figure is based on the "Information assurance: Interaction between security and dependability" figure in *Information Assurance – Dependability and Security in Networked Systems*, Qian, Joshi, Tipper, Krishnamurthy, 2008, New York, ISBN: 978-0-12-373566-9.

[9] Prior to publishing the ISO/IEC 27040 standard, the metadata for the index was stripped and the index was lost in the published standard.

[10] The SNIA published index for ISO/IEC 27040 can be found at: http://www.snia.org/securitytwg.

The remainder of this section will provide a summary of the ISO/IEC 27040 guidance associated with data protection as well as controls for other related areas (e.g., disaster recovery and business continuity).

3.1 Securing Backups

As part of ISO/IEC 27002:2013's "operations security," the objective for backups is "to protect against loss of data." In addition, the identified control is: "Backup copies of information, software and system images should be taken and tested regularly in accordance with an agreed backup policy." ISO/IEC 27002 also provides implementation guidance that includes:

- A backup policy should be established to define the organization's requirements for backup of information, software and systems; this policy should define the retention and protection requirements
- Adequate backup facilities should be provided to ensure that all essential information and software can be recovered following a disaster or media failure
- Backup plans should consider records of backup copies and documented restoration procedures, the extent (e.g. full or differential backup) and frequency of backups, remote location storage, the need for physical and environmental protection, regular testing of backup media, and encryption of backups (when confidentiality is important).
- Operational procedures should monitor the execution of backups and address failures of scheduled backups to ensure completeness of backups according to the backup policy
- Backup arrangements for individual systems and services should be regularly tested to ensure that they meet the requirements of business continuity plans; for critical systems and services, backup arrangements should cover all systems information, applications and data necessary to recover the complete system in the event of a disaster
- The retention period for essential business information should be determined, taking into account any requirement for archive copies to be permanently retained

ISO/IEC 27040 leverages the above recommendations by reference as well as expanding upon and emphasizing the following:

- The backup systems and storage media need to be adequately protected (e.g., encryption of media or operator authentication and authorization) against unauthorized access
- For backed up data, especially business/mission critical data, the backup approach needs to be aligned with the associated restore strategy for the data
- Storage media should always be handled by trusted individuals (including vendors); in this context, "trusted" means vetted/cleared/bonded individuals
- Not only does their need to be an audit trail showing backups are performed, but there also needs to be “proof” that restore requirements are being met

3.2 Securing Replication

ISO/IEC 27002:2013 does not address replication explicitly, but under "Information security continuity" (under "Redundancies") there is an objective "to ensure availability of information processing facilities" with a control that states, "Information processing facilities should be implemented with redundancy sufficient to meet availability requirements." The specific implementation guidance focusses on understanding the business requirements for availability and testing of failover mechanisms implemented.

As part of ISO/IEC 27040's guidance associated with data availability, the following recommendations are made:

- For replicated data, especially business/mission critical data, the replication approach needs to be aligned with its associated reliability, fault-tolerance, or performance requirements for the data
- The replication approach should provide adequate protections against unauthorized access (e.g., data in motion encryption).

In other sections of ISO/IEC 27040, the following guidance is also provided:

- Care should be exercised to ensure that compression and deduplication do not adversely affect remote replication
- Replication of sensitive or business/mission critical data that is encrypted on the primary storage should also be encrypted on replicated storage
- Replication of encrypted data (ciphertext) may necessitate additional management of data encryption keys, especially for DR/BC solutions (remote/out-of-region replication)

3.3 Securing Continuous Data Protection (CDP)

Like replication, ISO/IEC 27002:2013 does not explicitly address CDP. However, ISO/IEC 27040 does address the following:

- The CDP approach (e.g., continuous, near continuous, fixed interval, etc.) needs to be aligned with its associated restore strategy, especially when used in conjunction with business/mission critical data
- In high network bandwidth scenarios (e.g., multimedia files), throttling techniques should be employed to prioritize network traffic in order to reduce the impact of CDP on day-to-day operations
- The CDP approach should provide adequate protections against unauthorized access (e.g., data in motion and data at rest encryption).

3.4 Controls Related to Data Protection

ISO/IEC 27040 does include guidance for some other technologies that are related to data protection (storage view). These include business continuity management solutions as well as data retention (archive) and cloud technologies.

3.4.1 Business Continuity Management

ISO/IEC 27002:2013 dedicates an entire clause to "Information security aspects of business continuity management" with a key objective of "Information security continuity should be embedded in the organization's business continuity management systems." The controls associated with this objective include:

- The organization should determine its requirements for information security and the continuity of information security management in adverse situations, e.g. during a crisis or disaster.
- The organization should establish, document, implement and maintain processes, procedures and controls to ensure the required level of continuity for information security during an adverse situation.
- The organization should verify the established and implemented information security continuity controls at regular intervals in order to ensure that they are valid and effective during adverse situations.

In addition, ISO 22301[11] and ISO 22313[12] provide respectively requirements and guidance to organizations in determining their business continuity needs. ISO/IEC 27031[13] in turn provides guidance for organizations in determining their ICT resilience and recovery requirements in support of wider business continuity while the new ISO/IEC 27036[14] multi-part standard provides a broad level of guidance in terms of the acquisition of IT services from suppliers. Note also that ISO/IEC 24762:2008[15] has been withdrawn as it is no longer a relevant document and that it has been superseded by ISO/IEC 27036.

[11] ISO 22301, *Societal security -- Business continuity management systems --- Requirements*, was developed by developed by ISO/TC 223, *Societal security*. ISO 22301 is the main standard, which defines the framework for business continuity management.

[12] ISO 22313, *Societal security -- Business continuity management systems -- Guidance*, was also developed by ISO/TC 223. ISO 22313 is an auxiliary standard that helps with the ISO 22301 implementation.

[13] ISO/IEC 27031:2011, *Information technology -- Security techniques -- Guidelines for information and communication technology readiness for business continuity*

[14] ISO/IEC 27036, *Information technology -- Security techniques -- Information security for supplier relationships*, is currently a four-part standard.

[15] ISO/IEC 24762:2008, *Information technology -- Security techniques -- Guidelines for information and communication technology disaster recovery services*

Recognizing that storage is typically a critical element of an organization's ICT Readiness for Business Continuity (IRBC) program or informal DR/BC activities, ISO/IEC 27040 includes the following guidance associated with DR/BC:

- ensure that the storage ecosystem is factored into the DR/BC planning and implementation;
- prepare for limited disruption events (system failures, adversarial attacks, operator errors);
- identify and document the unique staffing and facility requirements associated with the storage ecosystem;
- perform on-going planning and regular testing of assumption, which are critical to successful DR/BC; results of DR/BC testing should be fed back into on-going maintenance of the DR/BC plan.

3.4.2 Data Retention (Archive)

ISO/IEC 27040 addresses data retention from two distinct perspectives: 1) short to medium-term (less than 10 years) and 2) long-term. The short to medium-term retention drivers are often based on legal, regulatory, or statutory requirements that also include security provisions; failure to meet these requirements can result in significant liabilities for the organization.

Long-term archival storage systems introduce integrity, authentication and privacy threats that do not generally exist in non-archival storage systems. In addition, the long lifetime of data gives attackers a much larger window within which they can attempt to compromise a security system; with archival storage an assailant might have several decades of time to conduct an attack (slow attack). To address these issues, ISO/IEC 27040 recommends the following:

- Archival storage assumes a write-once, read-maybe[16] access pattern, thus the integrity of the data in the system should be actively checked at regular intervals rather than waiting to when it is read.
- When migrating archival data to newer storage technologies, introduce available security capabilities that offer enhanced security measures to better secure the data in its new location.
- Since the data in a long-term archive can out-live the data stewards, a secure, archival storage system should be able to authenticate new users and establish their relationship to resources attached to existing users.
- Secrecy mechanisms (e.g., encryption, secret-sharing, etc.) should function in the complete absence of the user that wrote the data (e.g., a new user who is given rights to read data should also be given the ability to decrypt the data).

[16] The standard uses this language, which is intended to highlight the fact that much of the data recorded in an archive is never accessed.

- Security logging should be sufficiently complete and long-lived (measured in decades) that it assists in detecting slow attacks and maintains an attack history that can be used to make decisions to adjust the data protections.
- The system should either immediately deal with any compromise or maintain a history of compromises in order to intelligently schedule corrective action.
- The use of data reduction technologies (e.g., compression and deduplication) should be used in a manner that avoids compromising data integrity (e.g., factored into copies that might not have any association with the data reduction technologies).

To assure successful retention of digital information over short to mid-term retention periods, requires utilization of data protection, Disaster Recovery, and digital preservation and curation practices commensurate with the value of the information being retained, the risk of loss from all factors, and the acceptable amount of loss over the retention period. From a storage perspective, these short and medium-term data retention scenarios usually span one or more generations of technology and require the capture and retention of associated metadata. The following should be considered for short and medium-term retention:

- Multiple physical or logical replicas of the data should be created and preserved;[17] the replicas need to be organized to be as independent as possible (e.g., geographic, administrative/management, and platform/operating system), and their number chosen according to the data's value and tolerance of risk.
- On a defined schedule, audits should be conducted to test for both obvious and latent faults (e.g., integrity checks), and the damage they cause; repair the corrupted data using the good data from other replicas before that damage spreads.
- Match the access control scheme to the legal and regulatory requirements for the information being preserved.
- Ensure that accountability and traceability measures are adequate and functional; all data accesses may require audit log entries.
- Implement mechanism to demonstrate data authenticity, provenance, and chain of custody, especially for data of an evidentiary nature.
- If encryption is used, archive/escrow the keys and keying material; rekey the data within recommended cryptoperiods or when the underlying cryptographic algorithm needs to be replaced.

3.4.3 Cloud Computing

Both proprietary and standards-based, cloud computing storage offerings are in use and they commonly provide copy capabilities (e.g., mirror some or all the storage on a system), backup and recovery capabilities, long-term retention capabilities (e.g., archives), and multi-system

[17] It is not at all about how many copies, rather about the quality and characteristics of the digital archive process.

synchronization capabilities (e.g., allows a user to synchronize data on multiple and potentially diverse types of devices). ISO/IEC 27040 provides the following guidance for cloud storage:

- Transport security, such as IPsec or Transport Layer Security (TLS), should be used for all transactions
- Data at rest encryption and appropriate key management processes should be used to prevent access by unauthorized parties (e.g., cloud service provider personnel, other tenants, adversaries, etc.) when sensitive data is stored in a third-party cloud environment
- User registrations should be handled securely, and strong password authentication should be used to protect access to data
- Access controls that guard against unauthorized access from other tenants while providing appropriate access privileges to users permitted to access the data should be used
- Sanitization capabilities should be used to clear sensitive data from the cloud computing storage

ISO/IEC 27040 provides additional, specific guidance for SNIA Cloud Data Management Interface (CDMI)[18] implementations and use.

4 SNIA Data Protection Guidance

4.1 Data Confidentiality

ISO/IEC 27000 defines confidentiality as the "property that information is not made available or disclosed to unauthorized individuals, entities, or processes." ISO/IEC 27040 points out that "within storage infrastructures, data confidentiality is typically maintained using some method of encryption. These methods are most often associated with protecting data while it is transferred (sometime referred to as in flight or in motion) within the storage infrastructure or as it is stored (or at rest) within a device or on storage media." The *SNIA Storage Security: Encryption and Key Management* whitepaper addresses many of these concepts and provides guidance for storage.

While cryptographic mechanisms are one of the strongest ways to provide confidentiality, additional mechanisms may also be required to assure data confidentiality:

- Authentication processes
- Authorization and access controls
- Data classifications and policy

[18] SNIA Technical Position: Cloud Data Management Interface (CDMI) v1.1.1, SNIA, March 2015; also known as ISO/IEC 17826:2016 (SNIA). Information technology -- Cloud Data Management Interface (CDMI)

- Proof of controls and audit logging

From a data protection perspective, maintaining data confidentiality is one of the most important aspects of ensuring protection of personal data.

4.2 Data Classification

As part of the *SNIA Data Protection Best Practices Whitepaper's* description of confidentiality, a simple data classification scheme (Figure 3) is proposed that uses production versus non-production and sensitive versus non-sensitive aspects (see table). Production systems are often treated as being in a different security domain than development systems, which may have less stringent security requirements and controls. This situation is recognized in ISO/IEC 27040 with guidance that encourages separation of the two environments.

Importance/Priority	Production	Non-production
Sensitive	High	Medium
Non-sensitive	Medium	Low

Figure 3. DPCO Simple Data Classification Scheme

Using this simple scheme, it is possible to establish a basic set of priorities or a relative importance of certain data. However, this may not be adequate for many organizations. Consider that "sensitive" may require sub-categories to address regulatory requirements associated with PII (e.g., GDPR), health care may have special versions of PII (e.g., HIPAA/HiTECH), and national security imposes another dimension of categories. It is also worth noting that ISO/IEC 27040 does not explicitly address the subject of data classifications, but it does suggest that focusing on data sensitivity or criticality can help an organization begin their analysis for identifying relevant storage security controls (e.g., sanitization, access control, authentication, encryption, and key management) necessary for their environment.

As a general recommendation, SNIA encourages the use of the smallest number of sensitivity categories as possible, but this should be driven by a clear understanding of organizational risk.

4.3 Due Diligence/Due Care

In many instances, the regulations associated with data protection of personal data or PII (privacy) do not include details on the specific security controls that must be used. Instead, organizations are required to implement appropriate technical and organizational measures that meet their obligations to mitigate risks based on the context of their operations. Put another way, organizations must exercise sufficient due care and due diligence to avoid running afoul of the regulations.

There is no "safe harbor," so you must do things right. Even then, you may still have a data breach.

To help understand these concepts, consider the following:

- ***Due Diligence*** – Measure of prudence, responsibility, and diligence that is expected from, and ordinarily exercised by, a reasonable and prudent person under the circumstances. [BusinessDictionary.com] The diligence reasonably expected from, and ordinarily exercised by, a person who seeks to satisfy a legal requirement or to discharge an obligation. [Black's Law Dictionary (10th ed. 2014)]
- ***Negligence*** – Failure to exercise the *care* toward others which a reasonable or prudent person would do in the circumstances, or taking action which such a reasonable person would not. [law.com]
- ***Care*** – Level of active concern, or lack of *negligence*, towards avoidance of possible dangers, mistakes, pitfalls, and risks, demanded of a party as a duty or legal obligation. See also *due care* and *duty of care*. [BusinessDictionary.com] Under *negligence* law, the conduct demanded of a person (or entity) in a given situation. [Black's Law Dictionary (10th ed. 2014)]
- ***Reasonable Care*** – As a test of liability for negligence, the degree of care that a prudent and competent person engaged in the same line of business or endeavor would exercise under similar circumstances. Generally, reasonable care is the application of whatever intelligence and attention one possesses for the satisfaction of one's needs. The term is always relative, depending on the particular circumstances. What is reasonable care in one case (for example, involving an adult) might be gross negligence in another (for example, involving an infant). [Black's Law Dictionary (10th ed. 2014)]
- ***Due care (Duty of care)*** – Degree of *care* that an ordinary and reasonable person would normally exercise, over his or her own property or under circumstances like those at issue. The concept of *due care* is used as a test of liability for *negligence*. [BusinessDictionary.com]
- ***Standard of Care*** – Degree of prudence and caution required of an individual who is under a *duty of care*. [Merriam-Webster's Dictionary of Law. Merriam-Webster. 1996] In the law of negligence, the degree of care that a reasonable person should exercise. [Black's Law Dictionary (10th ed. 2014)]

Failure to take basic steps to understand risk exposures as well as addressing any identified risks is a quick way of demonstrating a lack of due care or due diligence, which can have significant negative consequences. This situation can be further complicated when required breach notifications[19] are not done or are mishandled, especially when the data breach can be attributed to the lack of due care or due diligence.

Storage systems and ecosystems are such integral parts of ICT infrastructure that these concepts frequently apply, but this situation may not be understood by storage managers and administrators who are responsible and accountable. It is important for these individuals to recognize that:

- Protections are often necessary in the storage infrastructure to guard against unauthorized, accidental or intentional corruption, modification, or destruction of data
- The risks associated with data breaches can be significant for some organizations, so prudence dictates the use of reasonable measures such as SNIA best practices and guidance in ISO/IEC 27040 to guard against these breaches
- Proper data preservation and disposal activities (see 4.6) are necessary for an organization to meet its legal obligations
- Policies are important administrative controls to facilitate proper data handling

4.4 Retention and Preservation

There are many instances in which the terms "retention" and "preservation" are used interchangeably and incorrectly. This can result in different and conflicting requirements that govern how the same information is maintained, how long it must be kept, and whether and how it is protected and secured.

ISO TR 18492:2005 notes that electronic document-based information constitutes the “business memory” of daily business actions or events and enables entities to later review, analyze or document these actions and events. As such, this electronic document-based information is evidence of business transactions that enable entities to support current and future management decisions, satisfy customers, achieve regulatory compliance and protect against adverse litigation. To achieve this goal, this electronic document-based information should be retained and appropriately preserved (e.g., addressing evidentiary requirements, which includes authenticity).

Preservation requirements often take on legal (e.g., holds) and/or usability focuses. Usability preservation addresses processes and operations involved in ensuring the ability to read, interpret,

[19] Breach notifications have become a mandatory element for much of the data protection regulations around the world.

authenticate, secure and protect against the loss of data or information throughout its lifecycle.[20] Usability preservation can also involve transformation of data (e.g., either conversion of files written by obsolete word processors or preservation of the associated ecosystem).

An organization should have a records management policy that defines what is a record[21] and how records will be managed.[22] In addition, the organization should have a retention schedule that classifies its records into record series, with associated retention periods and metadata. It is important to note that not all documented information in the possession, custody, or control of the organization should have record status. Instead, only documented information regarding the operation of the organization's business that it is legally required to keep, or which has legal compliance or business value should be a record.

At any given moment, the same information can exist in multiple "states," meaning the purpose for which the information is kept, rather than its physical location or medium. Recognizing these various states as well as using consistent language when describing these states is critical to ensuring applicable retention and preservation requirements are identified.

4.4.1 General Data Retention

Record-quality information should be retained in the ordinary course of business pursuant to the retention schedule, regardless of the medium of the record (such as paper, digital data, or micrographics). Retention periods should be codified in policy and determined based upon legal requirements and legal considerations, and also by considering the business value and business need for the information. Legal compliance and business considerations may also dictate the manner in which the records are retained, including how they are protected and secured. And once, in the ordinary course of business, a record has been retained for the length of time that the retention schedule indicates, it should properly be disposed of because its compliance and

At least one law firm in the U.S. has noted that there are more than 56,000 legal requirements and considerations in the statutes and published regulations of the United States federal system and the fifty states.

[20] SNIA definition for "preservation."

[21] An example of a definition for a record is: "A record is broadly defined as documentary material, in any media, that is created or received in the normal course of business, is worth preserving, either temporarily or permanently, because it provides evidence of the organization's policies, procedures, activities, and decisions and has technical, administrative, historical, and/or legal value."

[22] The SNIA Cloud Data Management Interface (CDMI) specification states that retention management includes implementing a retention policy, defining a hold policy to enable objects to be held for specific purposes (e.g., litigation), and defining how the rules for deleting objects are affected by placing either a retention policy and/or a hold on an object.

business value has expired. While it may seem contradictory that retention schedules should include a data disposition policy, organizations that keep everything are exposing themselves to considerable risk.

With all the requirements on record retention, it can be difficult to design a storage infrastructure that adequately protects an organization's records. When sampling some of the U.S. records retention requirements (see Table 1), it becomes clear that these requirements fall into either "permanent" or "temporary" retentions, with temporary being 10 years or less.

Document Type/Contents	Retention
Articles of Incorporation, charter, bylaws, minutes, and other incorporation records	Permanently
Copyright, trademark, patent registrations	Permanently
Deeds, mortgages, bills of sale	Permanently
Depreciation schedules	Permanently
Mission Statements, Strategic plans	Permanently
Workers compensation documentation	10 years after 1st closure
Contracts, mortgages, notes and leases (expired)	7 years
Stock and bond certificates (cancelled)	7 years
Personnel files, terminated employees	7 years after termination
Insurance policies	3 years after expiration
Correspondence with customers and vendors	2 years
Grants, un-funded	1 year

Table 1. Sample documents types with minimum retentions.[23]

ISO/IEC 27040 approaches data retention (see 3.4.2) from the perspective of long-term versus short/medium-term retention with the latter being driven by legal, regulatory, or statutory requirements that are shorter than traditional archives (less than 10 years). The evidentiary nature of the short/medium-term retention is thought to have noteworthy differences that could impact security.

4.4.2 Archive

ISO 14721 points out that the term "archive" has come to be used to refer to a wide variety of storage and preservation functions and systems, and further, that traditional archives are understood as facilities or organizations which preserve records, originally generated by or for a government organization, institution, or corporation, for access by public or private communities. The archive accomplishes this task by taking ownership of the records, ensuring that they are understandable to the accessing community, and managing them so as to preserve their information content and authenticity.

The *SNIA Data Protection Best Practices Whitepaper* characterizes an archive as a collection of data objects that represent an official working copy of the data, but is managed separately from more active production data, for such purposes as long-term preservation and better cost economics. Further, archives are often used for storing data sets that need to meet specific regulations and/or legal/contractual obligations, and they are normally used for auditing or

[23] *Records Retention and Disposition Guidelines*, Prepared by the Collaborative Electronic Records Project, Rockefeller Archive Center, Revised November 2008

analysis rather than for application recovery. In addition, the whitepaper notes that retention requirements can vary (e.g., short, medium, and long-term), but the archive must ensure proper integrity, immutability, authenticity, confidentiality and provenance.

ISO TR 18492:2005 defines "long-term preservation" as the "period of time that electronic document-based information is maintained as accessible and authentic evidence" and further notes:

> *This period of time can range between a few years to hundreds of years, depending upon the needs and requirements of the organization. For some organizations, this period of time would be determined by regulatory compliance, legal requirements and business needs. For other organizations, such as archival repositories holding public records, the period of time required to retain electronic document-based information is usually thought to be hundreds of years.*

ISO TR 18492:2005 also identifies six key issues that storage repositories should consider when they are developing a long-term preservation strategy:

- *Readable electronic document-based information* – the bit stream comprising electronic document-based information should be accessible on the computer system or device that initially created it, currently stores it, currently accesses it, or will be used to store it in the future; media obsolescence and data formatting are also considerations.
- *Intelligible electronic document-based information* – intelligibility of electronic document-based information is a function of information about what the bit stream in fact represents and the processing software's capacity to take appropriate action based on this information.
- *Identifiable electronic document-based information* – document-based information should be organized, classified and described in such a way that it is possible for users and information systems to distinguish between information objects based upon a unique attribute such as name or ID number; facilitating search and retrieval is a consideration.
- *Retrievable document-based information* – discrete information objects (or parts of them) can be retrieved and displayed. Retrievability is typically software-dependent in that it requires keys or pointers that link the logical structure of information objects (e.g., data fields or text strings) to their physical storage location.
- *Understandable document-based information* – conveying information to both computers and humans beyond the document contents, including context of creation and use (i.e., metadata) as well as relationships among other documents
- *Authentic electronic document-based information* – ensure the information is what it purports to be (i.e., information that over time has not been altered, changed or otherwise corrupted); focusses on a) transfer and custody, b) the storage environment, and c) access and protection.

The potential evidentiary nature of archives and the need to address data authenticity, provenance, and chain of custody are noteworthy because the archive may need to retain, protect, and maintain significant amounts of metadata. This means that the following security services identified by ISO 14721 apply to both the information and metadata:

- *Identification/authentication service* confirms the identities of requesters for use of information system resources. In addition, authentication can apply to providers of data. The authentication service may occur at the initiation of a session or during a session.
- *Access control service* prevents the unauthorized use of information system resources. This service also prevents the use of a resource in an unauthorized way. This service may be applied to various aspects of access to a resource (e.g., access to communications to the resource, the reading, writing, or deletion of an information/data resource, the execution of a processing resource) or to all accesses to a resource.
- *Data integrity service* ensures that data is not altered or destroyed in an unauthorized manner. This service applies to data in permanent data stores and to data in communications messages.
- *Data confidentiality service* ensures that data is not made available or disclosed to unauthorized individuals or computer processes. This service will be applied to devices that permit human interaction with the information system. In addition, this service will ensure that observation of usage patterns of communications resources will not be possible.
- *Non-repudiation service* ensures that entities engaging in an information exchange cannot deny being involved in it. This service may take one or both of two forms. First, the recipient of data is provided with proof of the origin of the data. This protects against any attempt by the sender to falsely deny sending the data or its contents. Second, the sender of data is provided with proof of delivery of data. This protects against any subsequent attempt by the recipient to falsely deny receiving the data or its contents.

These security services may need to be applied during the storage and transfer of the data and metadata to and from the archive. Equally important, care must be exercised when the security services/controls are being adjusted/replaced to avoid exposing the archived data to attack and/or disclosure (i.e., risk).

In many of the standards and publications, privacy is often not directly addressed in the context of archives; however, with the increase in privacy (protection of PII) regulations around the world, this is something that SNIA believes should be addressed.

Provenance and authenticity are essential elements of most archives, which means that proper metadata handling is required. SNIA also notes that chain of custody measures may be

necessary as well to address evidentiary requirements and this could complicate the nature of the archive solutions used (e.g., cloud storage may not be able to provide the needed details).

Many archives are concerned with "proving" data has not been changed (authenticity), but an alternate strategy is to employ immutability measures (e.g., WORM storage) instead of integrity verification approaches.

4.5 Data authenticity and integrity

The terms *data authenticity* and *data integrity* are often used together and sometime interchangeably, but frequently without a clear understanding of what they mean. ISO/IEC 27000 defines *integrity* as the "property of accuracy and completeness" and ISO 7498-2 further defines *data integrity* as the "property that data has not been altered or destroyed in an unauthorized manner." ISO/IEC 27000 also defines *authenticity* as the "property that an entity is what it claims to be;" there are few definitions for *data authenticity*, but those that exist are frequently aligned with "genuineness of data", "guarantee of the data provenance", and/or "assurance about the source of data." The relationship of the two concepts can be described as data authenticity is achieved when data integrity and authentication are joined together, or alternatively, data authenticity is authentication applied to a piece of data through integrity.

Data integrity is a core concern for storage systems and ecosystems with significant resources expended on ensuring data integrity as part of replication, data migrations, etc. Both data integrity and data authenticity take on additional significance within digital archives. For example, an archive solution may be designed to have multiple independent copies on different technologies wherein the integrity on each technology must be maintained as well as the integrity between the various copies. For data authenticity, the archive may serve as the mechanism for authenticating data. For this reason, SNIA recommends that data integrity and data authenticity elements of digital archives be implemented such that the digital archive becomes responsible for maintaining the evidentiary nature of the materials after the records have left the control of the organization that created them.

4.6 Monitoring, Auditing, and Reporting

ISO/IEC 27040 provides relevant audit logging guidance for storage, which is in line with common security guidance. When privacy is factored into audit logging, there are additional issues and complications that need to be consider. To explore some of these issues, the European Union (EU) General Data Protection Regulation (GDPR) is used as an example of what an organization may encounter.

There are at least three distinct aspects of GDPR that may impact a logging strategy: retention for a purpose, retention for a time, and anonymized retention. A general logging strategy is likely to record all accesses and/or all updates of data. If one has not designed the system with "privacy by design" in mind, personal information may be swept up in the data stream that is being logged. GDPR requires that if any EU citizen's personal information is retained, then that data must be anonymized such that the individual cannot be directly identified, that data must be tagged with the purpose for which it was collected, and that data must be tagged with a lifespan after which said data will be expunged (completely removed) and can be expunged on demand by that individual. It is highly unlikely that a general logging strategy would have been implemented with those needs in mind, so it is highly likely that current logging attempts that involve EU PII are rendering organizations culpable under GDPR for penalties that range up to 4% of annual worldwide revenue. The GDPR does allow retention of information for appropriate authorities, but does not define what those are, so it's unclear which authorities might be allowed to access what. Nor does the GDPR clearly delineate whether purposes can be aggregated (e.g., "healthcare" might reasonably include "dental", "medical" and "vision"), so caution suggests fine granularity until case law more clearly defines the boundaries.

4.7 Data Disposition/Sanitization

Within common records and information management (RIM) frameworks[24], disposition is the last stage of a record's life cycle. Within these frameworks disposition may not mean destruction, but rather, transfer to archives. In the latter case, this may simply delay when destruction occurs for most records (few records outside of government must be retained indefinitely). When records (data) are no longer needed, the destruction of the data becomes a critical, and often required, component of an effective data governance program. Data destruction is the process of removing information in a way that renders it unreadable (for paper records) or irretrievable[25] (for digital records).

A record isn't ready for final disposition until confirmation can be given that the information it contains will no longer be required for operational, legal, governmental or professional association compliance reasons. In addition, it is the organization's responsibility to ensure compliance with all electronic records disposal regulations governing operations and the organization's records retention policies.

In today's world, it might not be enough to remove all traces of data from digital and electronic records. Increasing concerns about privacy and security means electronic data disposal must be

[24] ISO 15489-1:2001, *Information and Documentation – Records Management – Part 1: General* is one of many frameworks for planning and implementing a records management program.

[25] In the digital world, making data irretrievable is caveated to a specified level of effort to retrieve it.

carefully and systematically handled to minimize the risk of illegal and/or unauthorized access to information. Proper sanitization of media[26] as well as maintaining proof of sanitization records may be required to meet legal obligations.

Within the context of data protection, data disposition, specifically data destruction, can be a major source of risk for an organization. Destroying data that must be retained as well as failing to properly destroy data using sanitization techniques or failing to destroy data that must be eliminated can result in significant exposures.

5 Summary

With the publication of ISO/IEC 27040, the storage industry has been presented with a broad range of guidance to help secure storage systems and ecosystems. While the standard doesn't call out data protection as an explicit topic, it does provide relevant controls as highlighted by this paper. That said, data retention and preservation, data authenticity, archive security, and data disposition are elements of data protection that are not addressed very well by the standard. SNIA recognizes the importance of these elements and has addressed them within this whitepaper, leveraging guidance from other standards as well as offering some of its own guidance.

[26] The *SNIA Storage Security – Sanitization* whitepaper addresses many of these concepts and provides guidance for storage.

6 Bibliography

ISO 7498-2:1989, *Information processing systems -- Open Systems Interconnection -- Basic Reference Model -- Part 2: Security Architecture*

ISO 14721:2012, *Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model*

ISO 15489-1:2001, *Information and Documentation – Records Management – Part 1: General*

ISO TR 18492:2005, *Long-term preservation of electronic document-based information*

ISO 22301:2012, *Societal security -- Business continuity management systems --- Requirements*, was developed by developed by ISO/TC 223, *Societal security*

ISO 22313:2012, *Societal security -- Business continuity management systems – Guidance*

ISO/IEC 2382:2015, *Information technology -- Vocabulary*

ISO/IEC 2382-1:1993, *Information technology -- Vocabulary -- Part 1: Fundamental terms*

ISO/IEC 24762:2008, *Information technology -- Security techniques -- Guidelines for information and communication technology disaster recovery services*

ISO/IEC 27000, *Information technology -- Security techniques -- Information security management systems -- Overview and vocabulary*

ISO/IEC 27002:2013, *Information technology -- Security techniques -- Code of practice for information security controls*

ISO/IEC 27040:2015, *Information technology – Security techniques – Storage security*

ISO/IEC 27031:2011, *Information technology -- Security techniques -- Guidelines for information and communication technology readiness for business continuity*

ISO/IEC 27036 (multiple parts), *Information technology -- Security techniques -- Information security for supplier relationships*

A Data Protection Taxonomy, SNIA Data Protection and Capacity Optimization (DPCO) Committee, June 2010, https://www.snia.org/sites/default/files/A_Data_Protection_Taxonomy_V51.pdf

SNIA Data Protection Best Practices whitepaper, SNIA Data Protection and Capacity Optimization (DPCO) Committee, October 2017,

https://www.snia.org/sites/default/files/DPCO/Data%20Protection%20BP%20White%20Paper%20Final%20v1_0.pdf

SNIA Index for ISO/IEC 27040, SNIA, February 2015, https://www.snia.org/sites/default/files/SNIA-WD_ISO-IEC-27040-Index.pdf

SNIA Storage Security: Encryption and Key Management whitepaper, SNIA, August 2015, https://www.snia.org/sites/default/files/technical_work/SecurityTWG/SNIA-Encryption-KM-TechWhitepaper.R1.pdf

SNIA Storage Security: Sanitization whitepaper, SNIA, August 2015, https://www.snia.org/sites/default/files/technical_work/SecurityTWG/SNIA-Sanitization-TechWhitepaper.R2.pdf

SNIA Technical Position: Cloud Data Management Interface (CDMI) v1.1.1, SNIA, March 2015, https://www.snia.org/sites/default/files/CDMI_Spec_v1.1.1.pdf

SNIA Dictionary, https://www.snia.org/education/dictionary

Records Retention and Disposition Guidelines, Prepared by the Collaborative Electronic Records Project, Rockefeller Archive Center, Revised November 2008

European Union (EU) General Data Protection Regulation (GDPR), L 119/1 REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016, https://publications.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en

Archival Authenticity in a Digital Age, https://www.clir.org/pubs/reports/pub92/hirtle.html

7 Acknowledgments

7.1 About the Authors

Eric Hibbard is Hitachi Vantara's CTO Security & Privacy, leveraging over 30 years of experience in ICT infrastructure with a specialty in data/storage security. He is the Chair of the SNIA Security TWG and holds leadership positions in the ABA, IEEE, CSA, and INCITS. He is and has served as the editor of multiple ISO/IEC and IEEE standards, including ISO/IEC 27040 (Storage security), ISO/IEC 20648 (TLS Specification for Storage Systems), and ISO/IEC 27050 (Electronic discovery). Mr. Hibbard currently holds the (ISC)2 CISSP and CCSP certifications as well as the ISSAP, ISSMP, and ISSEP concentrations credentials along with the ISACA CISA certification. See also www.linkedin.com/in/ericahibbard/.

Gary Sutphin has been a member of the SNIA Security TWG since 2007. Gary also served as a SME for several SNIA Certification exams, Storage Networking World Conference volunteer, SNIA Hands-on-Lab program volunteer and instructor, and an active member of the former SNIA End User Council. He started in IT with Sperry Univac then went on to work for Entrex/Nixdorf Computer, Prime Computervision, Sequent, and IBM. He recently completed the Cisco training program at the St. Petersburg College and resides in the Tampa Bay area. See also www.linkedin.com/in/garysutphin/.

7.2 Reviewers and Contributors

The Security TWG wishes to thank the following for their contributions to this whitepaper:

Thomas Rivera, CISSP, CISA	Co-Chair, SNIA DPCO
Gene Nagle	Co-Chair, SNIA DPCO
Richard Austin	Retired
Tim Hudson	Cryptsoft Pty Ltd
Bruce Rich	Cryptsoft Pty Ltd
Glenn Jaquette	IBM
Tim Chevalier	NetApp
Srinivasan Narayanamurthy	NetApp
Mark Carlson	Toshiba Memory America
Mike Wellman	SNIA/Colorado Technical University
Steven Teppler, Esq.	Abbott Law Group, P.A.

8 For More Information

Additional information on SNIA security activities can be found at https://www.snia.org/security. Additional SNIA storage security whitepapers related to ISO/IEC 27040 can be found at: https://www.snia.org/securitytwg.

Suggestion for revision should be directed to http://www.snia.org/feedback/.

The ISO/IEC 27040 standard can be purchased at http://www.iso.org.