



BUILDING A TERMINOLOGY BRIDGE:

Guidelines for Digital Information Retention and Preservation in the Datacenter

Retention
Preservation

Deletion
Long-term

Deduplication
Classification

Authenticity

ILM

Archive

Metadata

WHAT'S IN IT?

Terminology used in the datacenter can be very confusing. IT professionals, records administrators, security, legal, and compliance officers, along with the business groups, all have their own vernacular and they all hold a portion of the responsibility for maintaining corporate information assets. Communicating effectively with one another, in order to establish appropriate policies and procedures related to the retention and preservation of digital information, is the challenge. If you can communicate effectively, you can make the right decisions for your organization.

This report gives you the framework you need.

If the objective of your organization is to empower and improve collaborative communications across departmental boundaries, then a common terminology is essential. Agreeing on terminology before establishing new datacenter practices or methodologies, will go a long way in improving the maintenance and accessibility of your organization's information assets for years to come.

WHO NEEDS IT?

- Any organization operating an information governance-style committee or developing service management practices that needs to develop business requirements and policies for information assets. This report provides the common terminology and practice understanding needed to communicate retention and preservation practices among all stakeholders.
- An organization needing to better understand retention and preservation principles and have a common terminology that spans internal departments and business units, external partners, customers, and vendors.
- An organization dealing with eDiscovery, litigation holds, reducing risk and exposure, regulatory compliance, and/or long-term preservation of digital information, needs a tool to guide the development of key information management practices and preservation services. This report will help develop a common understanding of practices and requirements within the datacenter.

HOW DO I APPLY THIS REPORT TO MY ORGANIZATION?

This report is designed to aid individuals from disparate departments agree on common language and practices. It will help organization's get started with implementing their collaborative governance committees, service management methodologies, and Information Lifecycle Management (ILM)-based practices in the datacenter.

The objectives of this report are:

- **Stimulate adoption of ILM:** by reducing communication barriers, thereby building bridges between departments and encouraging implementation of ILM-based practices across the organization.
- **Improve communications:** by creating a comparative terminology between an ILM-based context and other key information management, archival, and preservation oriented industry glossaries to act as a bridge to better communications within the datacenter.
- **Explain terminology and practices:** by improving the understanding of what each retention and preservation oriented service attempts to achieve in the context of ILM-based practices.

BUILDING A TERMINOLOGY BRIDGE:

Guidelines for Digital Information Retention and Preservation in the Datacenter

HOW DOES THIS REPORT HELP IMPROVE WHAT I AM DOING TODAY?

Here are three examples of new approaches to current IT practices that are covered in detail in this report:

- **Archive:** the report advocates that IT practices adopt a more consistent usage of the term 'archive' with other departments within the organization. To the archival, preservation, and records management communities, an "archive" is a specialized repository with preservation services and attributes. Typical IT use of the verb "archiving" actually refers to a practice based on ILM called "tiering," or the migration of inactive, reference, or expired information to a lower tier of storage to reduce cost and improve storage efficiencies. A lower tier of storage is not an "archive" with preservation-class services. Another misuse happens when "archive" is confused with backup. According to this report, backup media saved offline or offsite does not constitute an 'archive' (a preservation store with preservation services) nor should backup media be confused with an 'archive' or with "tiering."
- **Preservation:** managing information in today's datacenter with requirements to safeguard information assets for eDiscovery, litigation evidence, security, and regulatory compliance requires that many classes of information be preserved from time of creation. Preservation is a set of services that protect, provide availability, integrity and authenticity controls, include security and confidentiality safeguards, and include an audit log, control of metadata, and other practices for each preservation object. The old IT practice of placing information into an archive when it becomes inactive or expired no longer works for compliance or litigation support, and only adds cost.
- **Authenticity:** is defined in a digital retention and preservation context as a practice of verifying a digital object has not changed. Authenticity attempts to identify that an object is currently the same genuine object that it was "originally" and verify that it has not changed over time unless that change is known and authorized. Authenticity verification requires the use of metadata. The critical change for IT practices is that metadata is now very important and must be safeguarded with the same priorities the data is. IT practices that damage, merge, ignore, or scramble metadata are no longer appropriate.

"Achieving business alignment has been the holy grail of IT for 20 to 30 years, but the dialogue is broken between business and IT. The reason? They don't have a language that works. The business wants to talk in business terms, and IT talks in technical terms" says Rudy Puryear, global IT practice head at Bain & Co.

Source: CIOInsight: "Why IT and Business Can't Get In Sync"
Tony Kontzer, 2009-06-04

HOW DO I GET A COPY OF THIS REPORT?

You can obtain a copy of Building a Terminology Bridge from the DMF's website at http://www.snia.org/forums/dmf/knowledge/term_bridge/ and you can participate in active discussion about it and other Data Management topics at the DMF Community site, <http://community.snia-dmf.org>.

Your feedback is welcomed and encouraged.

ABOUT THE DATA MANAGEMENT FORUM:

The SNIA Data Management Forum is a cooperative initiative of IT professionals, vendors, integrators, and service providers working together to conduct market education, develop best practices and promote standardization activities that help organizations become Information-Centric Enterprises. Areas of focus include the technologies and services that support information lifecycle management, data protection, and information retention, and preservation. For more information, visit www.snia.org/forums/dmf or participate in our open online community <http://community.snia-dmf.org>.

ABOUT THE STORAGE NETWORKING INDUSTRY ASSOCIATION:

The Storage Networking Industry Association (SNIA) is a not-for-profit global organization, made up of some 400 member companies spanning virtually the entire storage industry. SNIA's mission is to lead the storage industry worldwide in developing and promoting standards, technologies, and educational services to empower organizations in the management of information. To this end, the SNIA is uniquely committed to delivering standards, education, and services that will propel open storage networking solutions into the broader market. For additional information, visit the SNIA web site at www.snia.org.