Information Lifecycle Management

Future Standards, Policies and Technology Outlook

V1.2

Created: November 27th, 2006

Authored By: Ryan Walters

INTRODUCTION

Information Lifecycle Management (ILM) is not a unique product or technology, but rather an amalgamation of disciplines garnered together to create an organization specific solution. This solution is a well defined combination of systems, policies, and processes. Technology does however play a significant role. Thus all organizations must eventually consider technology as an integral part of this process. An organization's ILM strategy must take into account structured, semi-structured, and unstructured information. This information may be kept on a myriad of platforms in multiple formats.

ILM is currently a topic only being discussed by a particular group of people within the IT community, and yet it is fundamental to all who use information. IT must convince Business managers and senior executives to put ILM on their agendas, albeit within a context that ensures understanding, and IT management should avoid all talk of hardware and software. Instead, the focus should be on business information value management rather than terabytes and tiered storage.

Several issues arise when looking at ILM. First there is the issue of long-term information storage which can no longer be ignored as government mandates now are specifying the time and level of information that needs to be retained by an organization. An example of this is the growing number of litigation cases that require organizations to present electronic mail messages alongside other disclosures of information. Simply referring to a back-up tape or another data source does not constitute Compliance.

Next, the growing complexity and size of storage environments is causing system back-up problems which are leading to extended data recovery times, increased security concerns, information assurance issues, and administration costs spiraling out of control. Compound this with the degradation of application performance due to the striation of data and the issues become unmanageable.

Today, many IT departments are not able to keep up with the data and information placed within their care. Philotek believes that this predicament will only get worse. The time has come to give due consideration to the management of information through out its entire lifecycle and to treat information as a valuable business asset, rather than a by-product of a business process.

ILM is something all organizations know they should do, and indeed many believe they do it already; but as data volumes and the use of IT accelerates, Philotek believes that few organizations have a real handle on their long-term information storage requirements and obligations.

ILM automation is clearly still at an early lifecycle stage, as is ILM itself; but an increasing number of areas are being catered for by products, solutions, and technologies. Vendors across the board are creating proprietary solutions that capture relevant metadata and transform an enduser's environment into a striated environment. Systems from the major hardware vendors currently dominate the ILM world. However, Philotek believes that in the near future the focus in the ILM market will shift from hardware to software, as organizations move from their initial storage strategies that were hardware centric at an array level to a broader focus of storage integration and virtualization across platforms. This change of core focus will implore that more intelligence be built into the switching layers with less reliance on the arrays.

THE BUSINESS NEED FOR INFORMATION LYFECYCLE MANAGEMENT

The amount of electronic data needing to be held by the enterprise is increasing exponentially. Research from Storage Networking Industry Association (SNIA) has estimated that most large organizations have more than doubled their storage capacity since 2000 with, many far surpassing this growth rate. This growth in the amount of data, both in terms of hardware capacity and the drain on resources to manage it, has forced the topic of ILM onto the IT investment agenda.

Businesses need the development of a storage strategy and policies, along with a phased plan of action to evolve the data and information storage infrastructure to a shared resource and thus eliminate separate data stores tied to particular platforms, applications, and technologies. Managing a common pool of storage through ILM policies will not only improve utilization and scalability, but it will lead to a lower total cost of storage and better archiving/recovery methodologies. Information clearly has a lifecycle: It is created, it changes or stales, and then it is either destroyed or kept in perpetuity. ILM manages this lifecycle to optimize the use of resources, meet regulatory requirements, and to ensure the integrity of the information that is used to run a business or an organization.

Business Issues

When one looks at addressing the business drivers from an ILM perspective the discipline can be broken into relevant business drivers as follows:

• **Assurance**. A great deal of importance is placed on the chain of custody of information. Where did the information come from? When was it created? ILM can help organizations to answer these difficult questions.

• **Efficacy**. IT resources are finite. Organizations must implement an ILM strategy that ensures resources assigned to the management of information are relative to its business value and usefulness. Please note, this should not be a one-off decision, as information and the value of that information changes.

Mobility. As the workforce becomes increasingly mobile, information one needs to stay productive is taken well beyond the confines of a corporate network, and as a result it is no longer under the organization's direct control. Mobile devices, such as laptops, Personal Digital Assistants (PDAs), and cell phones, are now traditional business tools. While these devices provide increased productivity and greater freedom they bring into question the security of information as 1 in every 10 mobile devices is likely to be lost or stolen. Theft or carelessness can expose an organization to the loss of confidential or commercially sensitive information. An organization's ILM strategy must therefore take into account the security of information held on all devices and platforms.

• **Privacy**. ILM policies can significantly reduce the risks associated with breaches of privacy. While more important in some industries than others, the ability to track a piece of information throughout its lifecycle, from creation to destruction or throughout time, is something all organizations should strive for.

• **Protection**. ILM forces organizations to evaluate their informational assets in a manner similar to that of traditional physical assets. The level of protection afforded to a particular piece of information will vary throughout its lifetime, and it will be the ILM rules and policies set by the business that will determine this.

• **Productivity**. The amount of data stored is increasing everyday, and so is the amount of information that an organization must handle and process. Being able to consolidate and destroy information which is of no business value is just as important as being able to find and retrieve information which is valuable.

• **Security**. Knowing that one has a particular piece of information is different from knowing where it is located. Thinking that only one copy of a particular document or business record exists, when in fact there may be several, can cause greater problems than if it where to not exist at all. A well defined set of ILM policies should provide an organization with an 'information inventory' together with a corresponding audit trail.

• **Regulation**. Regulatory Compliance is driving 90% of the ILM decisions organizations are facing today. Government, corporate, and customer driven regulations that stipulate requirements for information security, retention and distribution, are all issues that can be addressed by ILM policies and practices. Organizations will have to accommodate greater degrees of regulation in the future and so information management automation, through the use of ILM policies, practices, and technologies, will be the only cost-effective way to meet compliance obligations of the future.

Technology Issues

For a variety of reasons, many organizations now find themselves with a complex heterogeneous IT infrastructure containing disparate hardware and software from numerous vendors. In Philotek's opinion this cannot be allowed to continue. When it comes to managing the lifecycle of that most valuable of assets – i.e. information – simplification and optimization must become the watchwords of the IT department. Approaches, such as using industry standards, consolidation, and virtualization, must be adopted in the information storage and management arena. Philotek is not recommending tying your organization to a single product or technology, but rather to a set of protocols and industry standards that will allow efficient management of your limited resources. Adopting this Gestalt approach to ILM can benefit organizations in many ways, least of which will be an improvement in storage utilization costs, which in many cases, can be reduced by as much 50% – if not more.

In order to meet the challenges facing the IT department, there is a need for a step-change in the way in which information storage solutions are architected, including broader focus on Enterprise Architecture and overall IT strategy. Organizations should also consider adopting a virtualized approach to data and information storage technology. Doing so breaks the link between the application and the underlying storage technology, and as a result avoids any form of vendor lock-in. Virtualization can also make for a far more flexible information storage infrastructure, and as a result an IT department might find that it can offer a better service to its business customers.

Technologies that protect data and information have always been a prerequisite for efficient operations, as every organization is vulnerable to data loss and corruption. A data protection strategy must be developed that employs effective backup, replication, synchronization, and mirroring tools. Data integrity has become more relevant because of increasing regulatory demands. Not only does this include being able to prevent unauthorized changes to information and data, but organizations must be capable of substantiating this through audit trails and tamper resistant system features.

The tiered storage concept is becoming increasingly important for enabling information migration and for exploiting the optimal storage media for the age, value, and usefulness of information. Where policies that are driven by business rules are used to allow information to be automatically migrated to lower levels of more cost effective storage media over time, then a full range of media options should be considered.

Market Issues

The ILM market is still developing, and as such Philotek does not believe that any one vendor has all of the pieces in place to offer a 'complete' ILM solution. While ILM solutions do exist today for specific ILM issues organizations wanting to deploy an enterprise-wide solution will have to be patient. Enterprise solutions are still three to four years off at the very least.

Today the market for ILM solutions can be broken down into the following sub-categories: Capture, Creation, Classification, Policy Management, Distribution, Storage and Media Migration, Discovery, Retrieval, Archiving, and Disposal. Each of these subcategories represents an entire field of storage related discussions and will not be covered for the sake of brevity. It must be duly noted that these sub categories are all relevant when attempting to create an enterprise wide ILM solution.

The first group of products that debuted on the market were the storage-oriented technologies that had content addressable storage functionality. These systems (not solutions) were merely revisions of existing storage products and thus came from traditional vendors in the Storage area. Adopting a hardware-led approach to ILM may appear to be an 'easy option' when considering the many information management challenges faced by organizations today, but Philotek believes this is only likely to provide a short-term solution. Philotek sees software technologies, such as audit and policy products, search and classification tools, and information security solutions as the emerging technologies in the ILM space.

INFORMATION LYFECYCLE: IT'S A DISCIPLINE NOT A PRODUCT

An organization's ILM strategy is reflected in its hardware, software and procedures and thus becomes a core business issue. ILM should be considered when discussing topics such as Disaster Recover and Back Up. Infrastructure planning must also be a critical piece of the ILM discussion, as it spans many areas of storage including storage allocation, media management, and high availability either through redundancy, backup or recovery.

An organization's LM Policy should be clearly defined as to avoid confusion, even at the expense of functionality. Philotek believes that complexity an increased leads to increased Total Cost of Ownership.

Many believe that ILM is a lot like Data Lifecycle Management (DLM), which was made popular by storage vendors a couple of years ago. DLM is also a policy-based process, in that it has rules and takes into account a lifecycle. The difference here, however, is that ILM focuses on information, not data. When referring to data, remember that it is unstructured bits and bytes; whereas information has context that differentiates it. Data is dependent on applications for meaning, whereas information (in theory) is independent of applications.

INFORMATION LIFECYCLE MANAGEMENT IMPLEMENTATION STRATEGIES

In October 2004, the Storage Networking Industry Association (SNIA) published its ILM roadmap subtitled, *Vision for Information Lifecycle Management*, it defines ILM as a new management practice for the datacenter. The problem is ILM is as much about business processes as it is IT processes.

SNIA's definition for ILM is as follows:
Information Lifecycle Management is comprised of the policies, processes, practices, and tools used to align the business value of information with the most appropriate and cost-effective IT infrastructure from the time information is conceived through its final disposition. Information is aligned with business requirements through management policies and service levels associated with applications, metadata, and data.

Philotek is comfortable with this as a vision for ILM, but we recognize that implementation strategies are all too often being led by IT departments rather than business units. Ask yourself : *Who is responsible for your information?* Odds are the response is: *The IT department*. Then ask the business units of your organization: *Does the IT department understand your information requirements?* Most business units organizations reply: *No they do not.* In order for any ILM implementation strategies to be successful the answers to the above questions must change.

SNIA's ILM roadmap describes five primary phases:

1) CONSOLIDATION AND THE IMPLEMENTATION OF NETWORKED STORAGE- This Phase focuses on the consolidation of data services and the implementation of network storage.
2) STANDARDIZATION OF DATA AND DATA SERVICES- The objective of this phase is to present a reliable, consistent, and efficient platform for the processes and policies to build upon. The difference in this phase is that SNIA espouts Data Classification, whereas Philotek believes that information-centric classification policies are required. Today, there exist very few systems that can place value on data above and beyond the scope of aging the data. Granted, the identification of an e-mail, a word document, or a CAD file can be determined, but systems cannot interpret the value of the information contained within. This dilemma forces organizations to promote an age-based approach to information classification and valuation. Older items generally are considered to be of lesser value than those more recently created. Philotek believes that XML will enable value to be associated with information, but for now organizations must use a broad approach to information classification in order to make ILM work effectively.
3) SOLUTION STACKS - In Phase 3 SNIA discusses Solution stacks which create a complete ILM environment based around a given application. 'Solution stacks' will enable organizations to automate ILM processes, and will be developed to focus on applications, such as Electronic Document and Records Management, e-mail management, enterprise applications (e.g. PeopleSoft, Saleforce.com, and SQL). As the market develops Philotek anticipates that a greater 'understanding' of the information contained within, and created by, these applications will be the core impetus of the ILM push.
4) ILM AUTOMATION – The Objective of this phase is to automate the management of information based on a given attribute or business trigger.
5) APPLICATION CENTRIC ILM SOLUTION STACKS – Lastly, true Enterprise ILM comes into existence. This represents the application-centric ILM 'solution stacks' beginning to interoperate and co-exist. The ultimate result will be information that will behave as though it were a single solution. In actuality it will be many different ILM 'solution stacks' interacting though standard protocols and interfaces.