Presentation# 4140
Protect Your Greatest Asset - Data
by Edward Harper, Phoenix Systems Integration, LLC
731 Main, Monroe CT 06468   (203) 268-9030

As we approach the threshold of a new millennium, it becomes even more evident to our society that information will bear greater weight as a factor between organizational growth and organizational starvation. Information is the tree that bears the fruits of success. It is the engine that drives the corporate ship toward greater power of leveraging within a sea of competition, regulation, and cooperation.

The life cycle of information is quite simple; it is retrieved or acquired, stored, and utilized. Although information can be viewed as simplistic in nature, we often encounter barriers within its intended application. When we pose the question of what is more important, the raw data itself or the process relating to the data, we fail to view the importance of the interdependencies between the two. Information as an entity is more a matter of both data and process linked together than the two being independent of one another. That is, it is highly difficult to address one without the inclusion of the other when viewed as a whole.
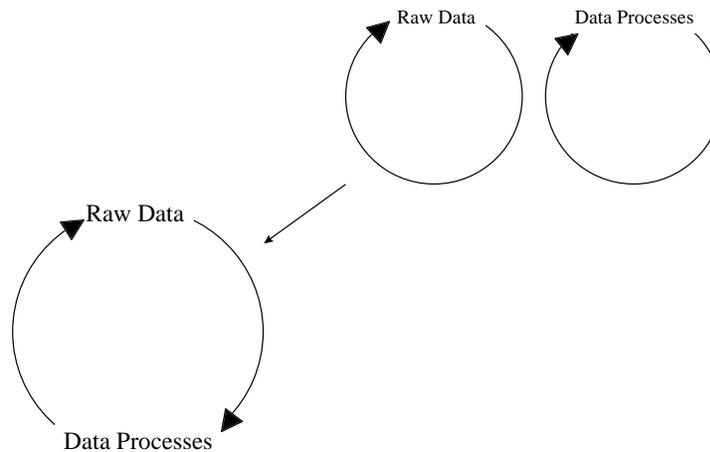
Within the information technology industry, this causal viewpoint of information as being interdependent rather than independent has brought about new innovations in hardware and software architecture, and business process strategy. The by-product of this innovative approach has unlocked a new facet of information technology, that being document management.

In the past, many business organizations viewed document management as a simple function that could easily be championed by merely implementing off-the-shelf database products that could be augmented with network hardware. We have even seen information technology professionals attempt to satisfy the document management challenge by implementing hardware capable of handling enormous amounts of data with processor speeds that would even make Albert Einstein ponder his theories. The theory was 'bigger and faster is better.'

Some specific problems encountered with the first aspect are:

1) most often, off-the-shelf software does not address an organization's specific needs, especially when confronted with the scope of user size, nature of data-specific requirements, and the safe storage environment.

2) document management, generally, was not a concern of architectural and strategic considerations during original implementation.

Most often, the main consideration of hardware professionals was to provide a system that would facilitate a powerful thoroughfare of information through the network, not to provide a system that would offer the most effective and efficient means of supporting enterprise-specific relational data.

Some specific problems encountered with the second aspect are:

1) faster speeds do not necessarily mean efficient system operation, nor does it accurately define effective system operation.

2) greater amounts of storage space do not truly address enterprise-specific processes, nor does it concentrate on curing the disease of cross-functional utilization conflicts.

3) neither speed nor size addresses today's growing concern over 'bottle-necking' and 'stove-piped' pathways or connections.
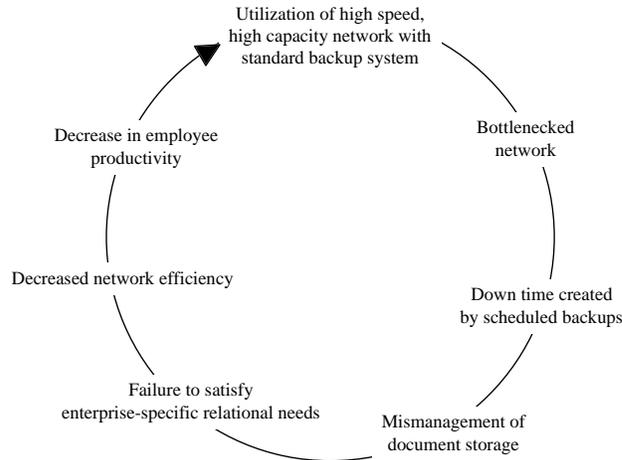
As we surge toward the future, propelled by the impetus of the new global competition paradigm, technology remains at the forefront to navigate our corporate ship through the many obstacles we face today. Along with the many milestones of technological advancement are the gravestones of obsolescence that also pave the path toward technological superiority.

In order to be truly successful, an organization must optimize its productivity. Whether a manufacturing or service entity, there will always remain one constant in corporate growth; the need to utilize the most productive means toward managing an organization's proprietary jewels - its information. This is usually the point where most of the confusion or, rather, misunderstanding lays. For document management to be a useful resource, one much first understand the true meaning of the term. In most cases, individuals consider document management, its terminology, and principles to mean something entirely different.

For example, in one prominent Fortune 50 company, an information technology professional was asked if his organization utilized a document management system. The veteran technician responded in a knee-jerk response, "Of course, we have a huge network that's faster than a bullet and can provide information to more users than we have." When asked about his company's backup system operation and redundant architecture as part of their disaster recovery strategy, the articulate gentleman added, "We're using the backup system that was recommended with our system. Sure it isn't perfect, however, what backup system is?"

Upon physical inspection of the organization's network, it was discovered that the network technician's perception of a document management system was solely about raw data storage capacity and the speed of the network's servers. In many circumstances, this example is much too often the case. The company had the fastest equipment and greatest amount of data storage space that money could buy, however, the system did not successfully achieve the goal of providing efficient thoroughfares, redundant processing, and enterprise-specific relational architecture.

The causal loop diagram below depicts the occurrences from that organization's document management strategy:

Utilization of high speed,
high capacity network with
standard backup system

Decrease in employee
productivity

Bottlenecked
network

Decreased network efficiency

Down time created
by scheduled backups

Failure to satisfy
enterprise-specific relational needs

Mismanagement of
document storage

So how do we protect our most valued asset - our information? And how do we provide assurance that these measures address the needs that specifically drive our organization? The critical point in understanding how to create such a solution is to first understand the primary elements of electronic document management systems.

The term *document* has been a part of our vocabulary for eons, however, we tend to assume that *document*, in the context of EDMS, refers to stored text or a metaphor of similar nature. In the EDMS paradigm, documents can contain text, however, they can also contain certain types information such as image, video, voice, or even taste and smell. As we approach heightened levels of awareness through virtual reality, tactile stimulation becomes more visible in this arena.

Documents, in the EDMS paradigm, are more than stored information files. For example, a graphic image may be stored within a system, however, the image cannot be retrieved through a database without some type of descriptive information contained within a structured index or database that is linked to the image. This is due, mainly, to the fact that graphic images do not maintain any level of granularity. That is, we cannot decompose the information any further than its current state for purposes of retrieval. The problem that is encountered with attempting to query such a document, is that we do not have the capability to retrieve the image based purely on its characteristics without the aforementioned information. Here we can clearly see how a document in the EDMS realm can be quite confusing to a user who relies upon common terminology in a new environment.

Electronic document management deals with the ability to work with varying types of document formats. One preferred method, the use of imaging, is not always the most effective or efficient one, however. What is essential in document management is the fact that the nature of the document, its interdependency and interrelationship within its function, is what dictates the content and architecture of the document and the system relational to it.
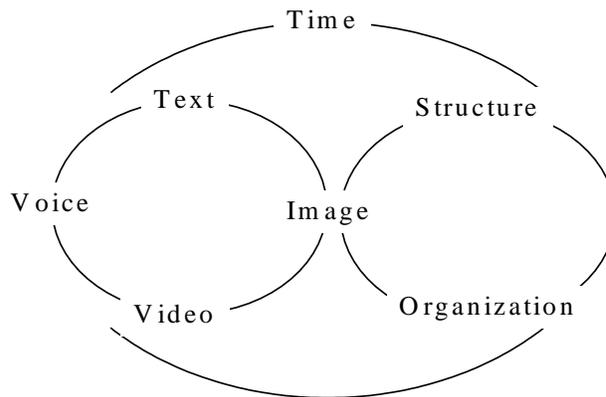
Some additional considerations regarding how we view and define a document, is based upon structure and organization. The structure of a document refers to the nature of its composition, its complexity and relationship between each item on a document. Organization, on the other hand, refers to how each document relates to one another.

It is essential that an understanding of life cycles exist when dealing with documents. Documents have a life cycle in the document management paradigm. In order to create a truly productive system to facilitate that process, it is important to understand the life cycle of a document along with its workflow.

Bayer's research and production facility in Berkeley, California was one such example of how the integration of a true document management system was able to not only provide efficiency in information retrieval and workflow, but harness greater networking power due to the optimization of the enterprise system's resources and organization-specific requirements.

By understanding the nature of what was truly required of the system at a fundamental level, the members of the Phoenix Systems Integration team were able to focus on greater EDMS functionality at a much higher level. Hewlett-Packard's network hardware played an integral part in this process. During the preliminary design phase of the Bayer DMS project, the identification of critical data and data formats was essential. To provide the most productive structure for managing the large workflows at Bayer's Berkeley facility, a data center solution was utilized vs. a stand-alone architecture. This strategy provided all components of the document management process a high availability of technology and a high degree of redundancy.

The final product resulted in not only the successful development of a custom document management system, but the creation and implementation of several key industry innovations. Listed below is an overview of the system:

- Documentum Enterprise Document Management System
- Client/Server Architecture
- Robust Client Applications
- Integrated E-Mail
- Customized Menus and Screens
- Electronic Signature and Verification

- Document Security and Integrity

- High Availability
- Redundant Server as Fail-Over
- Windows 95 and Mac Clients

Listed below is an overview of the system's software components:

- Documentum Enterprise Server 3.0
- Documentum Client
    - Workspace 3.1
    - MS Office Suite
    - Adobe Acrobat Exchange 2.1
    - Lotus Notes 4.5
    - Oracle 7
    - PenOp - Capture/View/Verify

The objective in the implementation of hardware was to provide a solid base architecture which will enable DMS to expand throughout the future. This was a critical aspect in attempting to provide Bayer with a system flexible enough to endure the challenge of technology vs. time.

Listed below are some benefits of the DMS hardware strategy:

- Remote Diagnostics (HP 9000 K420 Server - Production System)
- Open Architecture, Reduced Footprint, "Best of Breed" Technology (EMC Storage)
- Fail-Over, Development, and Test Configuration (HP 9000 K220 Server - Development System)
- HP MC/ServiceGuard Strategy - Enterprise Clustering
- HP Extended Warranty Program
- Cost Effective Long Term Strategy - Overall

Bayer was provided a system, comprised of overlapping components, that bundled high technology with an architecture complimentary to each of its component's strengths. For example, EMC's state of the art proactive monitoring, continuous(24x7) support, full scaleability with multi-processor support, and its open system approach was an immeasurable compliment to the leading-edge HP system.

Using HP system components was a critical decision in Bayer's document management solution. To develop a system that is capable of facilitating the challenging functions and processes of the Berkeley facility, Phoenix System Integration had to provide a upper tier system capable of moving in unison with the demanding requirements of technology and time.

Along with the standardized features of Bayer's document management system, certain customizations were made to exploit mission-critical needs. Listed below are some of those customizations:

- Additional Documentum Workspace Functionality
    - New Interface (Custom Menus and Screens)
- Additional Documentum Server Processes (Workflow)
    - Overdue Routers, Object ACLs (Access Control Lists)
- Adobe Acrobat Plug-Ins
    - Print Function, Electronic Signature Modules
- Server Scripts
    - On-Line Backup, System Admin., Batch Import

Addressing the value-added benefits for future implementation are:

- Multi-CPU Upgradable to 4 CPUs
- Multi-Storage Growth Path
- Next Generation Board Upgrades Built Into the Backplane of the Systems
- Fully Scaleable to the Growth Needs of the EDMS Production System

Highlighting Bayer's powerful and productive DMS, is the implementation of two innovative items:

1) The implementation and integration of PenOp's state-of-the-art electronic signature software. The PenOp software implements biometrics for a person's signatures. PenOp provides the capability to capture, view, and verify a person's signature. PenOp signatures are bound to the document and cannot be copied, moved or deleted. In the Bayer DMS, the attributes of the signatures applied to the document are extracted and stored in the attributes of the document. This functionality allows DMS users to view who signed and when by simply displaying the Documentum attribute screen for that document.

2) The creation of a true HOT backup system that allows the network to remain on-line while performing backup tasks. In most cases, when encountered with the network environment, system down-time is not only a requirement, but rather a necessity. The custom software architecture avoids the many pitfalls encountered, technically and administratively, when performing such a task. The HOT backup software synchronizes the backup of the Oracle database and the document content files while allowing users to view documents. This design helps meet the high availability requirement.

As our world becomes more global in nature, decentralization and reengineering will create an even greater demand upon our tasking requirements. To support this challenge, the retrieval, storage, and cross-functional utilization of electronic documentation will grow exponentially in order to remain in alignment with an organization's most basic survival needs.

On February 8, 1984, U.S. astronaut Bruce McChandless walked in space … totally free of any connection to earth. No ropes or cords to the spacecraft, no planet beneath his feet. For the first time ever, a human being was completely on his own, an individual satellite drifting in space. He was quite literally set free by technology. It is here, within the scope of electronic document management, where social and business communities will turn to, to free themselves from the domination of their current paradigm.