

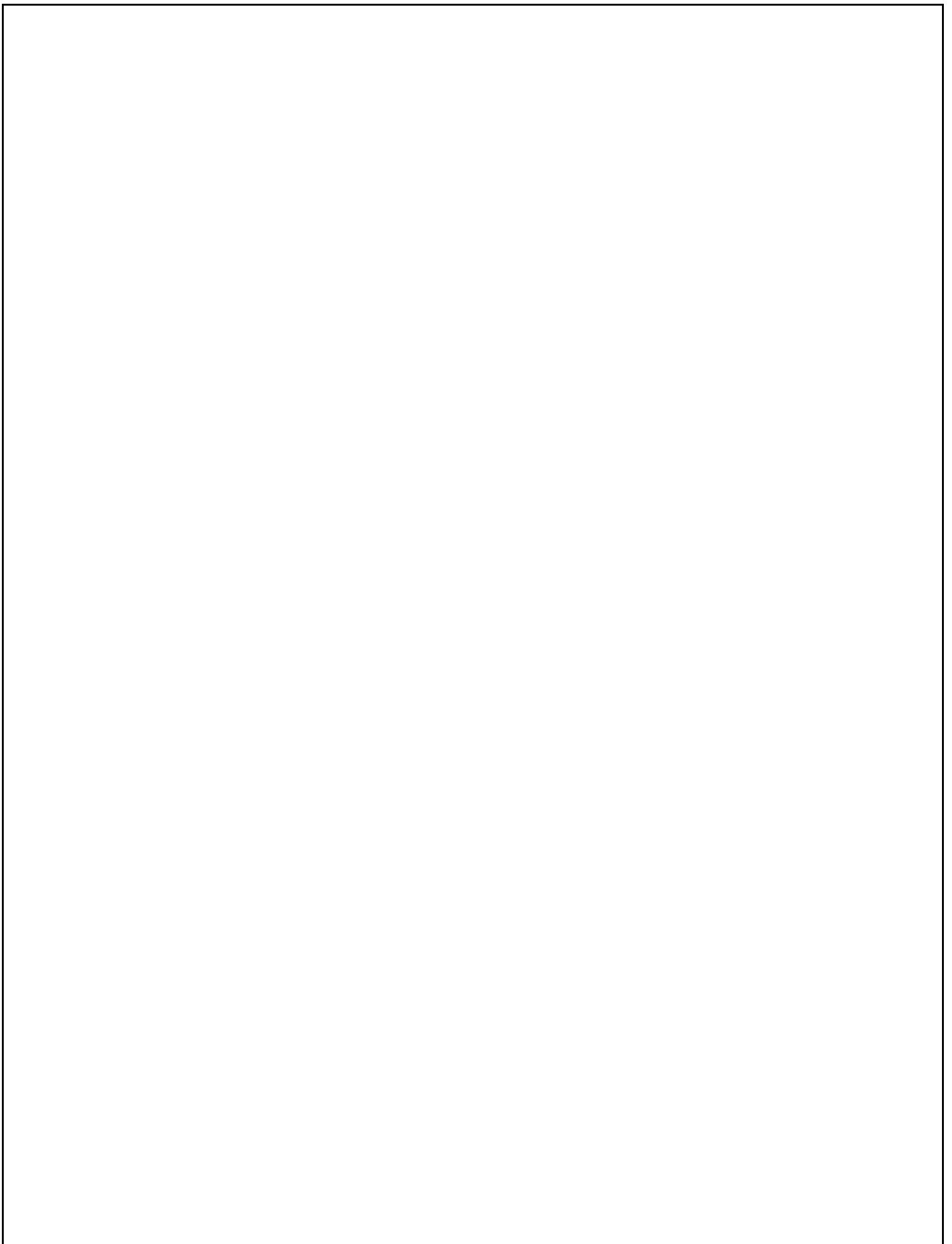
A GPR Consulting, Inc. White Paper



P.O. Box 9203  
Chesapeake, VA 23321  
888-718-1079  
[www.gprinc.net](http://www.gprinc.net)  
[info@gprinc.net](mailto:info@gprinc.net)

## **Electronic Business Records Data Storage Considerations**

*By Charles Heishman, Partner  
GPR Consulting, Inc.*





## Contents

---

<b>Introduction</b>	<b>3</b>
<b>Definitions</b>	<b>3</b>
<b>Storage Environments</b>	<b>4</b>
<b>Key Concepts</b>	<b>4</b>
<b>Summary</b>	<b>7</b>
<b>About the Author</b>	<b>8</b>
<b>About GPR Consulting, Inc.</b>	<b>8</b>

## Introduction

Retention Management is at the heart of any records and information management program. Today's electronic records management retention solutions encompass many alternatives when compared to the relatively simple options offered to their paper-based counterparts.

Before embarking on any record management initiative, you must first understand some basic definitions. You must know with what you are dealing and what your options are.

## Definitions

### *Retention Requirement*

A records retention and disposition program is that component of an organization's records management program that defines the period of time during which records are maintained and specifies procedures for the transfer and disposition of records. Recent legal and legislative decisions have placed electronic records management on par with other record formats.

For the purposes of electronic business records, there are five phases identified as being part of the lifecycle range. These are:

1. Creation and Receipt
2. Distribution
3. Use
4. Maintenance
5. Disposition

Retention focuses on Maintenance and Disposition.

### *Local Repositories*

Local repositories refer to those storage areas accessible by one individual. They are



quite often independent of any common back-up routines and as such are subject to data loss and quality inconsistencies. Local repositories typically are not involved in data sharing.

### ***Server-Based Repositories***

Server-based repositories refer to those storage areas accessible to multiple individuals, dependent upon security access privileges. Server-based repositories are the solution to data-sharing requirements.

### **Storage Environments**

Deciding the right storage solutions to meet a company's specific needs can be a daunting challenge since one solution does not fit all business cases. The reality is that data has different business values during its lifecycle. Using a combination of storage technologies (tape, disk, and storage networking) and software solutions can help companies effectively store, access, manage, share, and protect information assets.

### ***Backups***

In information technology, backup refers to making copies of data so that these additional copies may be used to restore the original after a data loss event. Backups are useful primarily for two purposes: 1) to restore a computer to an operational state following a disaster (called disaster recovery) and 2) to restore small numbers of files after they

have been accidentally deleted or corrupted.<sup>1</sup>

### ***Archives***

An archive refers to a collection of records. Archives are made up of records which have been created during the course of an individual's or organization's life. In general an archive consists of records which have been selected for permanent or long-term preservation.

### ***Off-Line Storage***

Off-line storage is computer storage that is not available for immediate, on demand use by the system without human intervention.

### **Key Concepts**

#### ***Storage Issues Have Changed***

Multitudes of storage solutions are available today to efficiently and cost-effectively leverage, manage, and access data. Gone are the days when all data is treated the same way, regardless of its age or business value.<sup>2</sup>

The December 1, 2006 changes to the Federal Rules of Civil Procedure introduced a new category of information, electronically stored information, into the legal discovery process and gave electronic data a heightened importance.

Companies should consider flexibility, scalability, performance, and costs associated with storage. The ultimate resolution blends different storage solutions for various business needs.

#### ***What is the Retention Requirement?***

Understanding the information lifecycle is critical to the management and retention of

---

<sup>1</sup> Wikipedia

<sup>2</sup> Tom Inglefield, Business Intelligence Journal



electronic records. The management of electronic data is in its infancy with regard to formal records retention programs. For the most part, we continue to manage all electronic data records the same, regardless of its position in its lifecycle.

Retention requirements define the period of time during which records have operational, legal, fiscal, or historical value. In addition, the retention requirements should define the point, and methods, at which electronic data can safely be migrated from on-line access systems to near-line or off-line storage systems.

In general, as most data ages, the probability of accessing that data substantially diminishes. When data is created, most accesses occur within the first day. During the next 30 days, the data is accessed occasionally and such data should be readily available for fast and easy access. After the first 30 days, the data is rarely accessed, and after 90 days, the data is almost never accessed.

The records retention program provides for the transfer of inactive records from primary space to less expensive secondary storage facilities. It also provides for the elimination of valueless records within the organization.

Unlike paper-based storage solutions, electronic data storage has been relatively inexpensive and easily expandable to fit ever-

increasing information repositories. The emergence of the data mart/data warehouse as a common component in the IT arsenal of data storage and retrieval is further evidence of this reality. This often leads to a “more is better” mentality when considering electronic data retention.

In the past, electronic data was often not considered the official records of an organization. The court decisions leading to and the eventual enactment of the Revised Rules of Civil Procedure have changed this. While often easier to just hold onto the data, risks associated with maintaining too much data for too long outweigh the costs and efforts involved with implementing and maintaining a routine electronic data records retention program.

Details for establishing an effective electronic records retention guideline can be found in “Electronic Records Retention: New Strategies for Data Life Cycle Management”.<sup>3</sup>

Regardless of your published retention requirement, you must also recognize and accommodate retrieval of pending litigation-related data. The average U.S. corporation is currently contending with thirty-seven lawsuits.<sup>4</sup> The retention and disposition program ensures that records are properly handled in anticipation of and during litigation or government investigation.

### ***Managing Local Repositories***

Managing local repositories is the most inefficient element in the storage management effort. The quality of data in

<sup>3</sup>

<http://www.arma.org/bookstore/productdetail.cfm?ProductID=1203>

<sup>4</sup> Coming Distractions, John Goff, CFO Magazine, 4/01/2006



local repositories can be inconsistent across an organization in that each owner determines the quality and content of the data.

Best practices would suggest you retain PC-based (local) electronic records based on official records status. Recent legislation and court decisions have explicitly declared that data stored in local repositories is subject to the same eDiscovery elements as other corporate data.

While a highly desirable goal, enforcement and execution of this standard is easier said than done. The records management professional often is unaware of all the layers of data stored in local repositories. Even when aware, his effort is complicated by inaccessibility to the data. The very nature of a local repository makes access to and management of locally stored data a daunting challenge.

Where possible, a more efficient and safe method for managing locally created and maintained data is to force it into a server-based environment.

### ***Managing Server-Based Repositories***

Storage area network (SAN) solutions deliver an entirely new level of service to support the growing storage needs of today's digital enterprise. Companies can achieve economies of scale through shared storage resources and simplify storage management through consolidation. In a SAN infrastructure, servers are optimized

for business applications and rely on high-bandwidth connections with other devices to capture data. Despite the complexity of SAN deployments, they provide a scalable and flexible option to consolidate and support both instant and serial access to data.

### ***Storage Environments***

Companies must ensure the data and information format for a planned application is compatible with the storage solution.

There are several storage options companies may consider to meet a multitude of business needs and budget requirements, including disk, tape, storage networking and virtual solutions.

### ***Backups and Archives***

Backups and Archives are two very different activities and as such require very different storage considerations. Differentiating between backups and archives can be critical and difficult. Under amended rule 34(a) of the revised Civil Procedures, a reviewing party has the right to test or sample electronic data, even if it resides only on backup media. This data, if so ordered, must be reproduced such that it can coherently be sampled.

To properly address the two distinct but equally important topics, we will address each one individually.

### **Backups**

Backups are used for data protection. They are organized for emergencies and urgent restoration processes to restore operations. If the operation works well, data in backups are *not* used again. Backups are not organized for business access. Producing data in its native format from backups, if even possible, can be expensive.



While there is language in amended rule 26(b)(2)(B) that provides some relief to the producer (i.e. owner of the backups) if the source data is not reasonably accessible on account of either undue burden or undue cost, the exact definition of “undue” is still in question.

### Archives

In an archive system, it is expected that the data *will be* used again and with some frequency. Archives must be organized for access to data and information with appropriate indexes and access points based on business usage.

A key consideration in selecting an archive solution is data usability. Under subsections 34(b)(ii) and 34(b)(iii) of the revised Civil Procedures, the default form for producing electronically stored information is that “in which it is ordinarily maintained [or] reasonably usable.”<sup>5</sup>

Commercial Off the Shelf (COTS) software solutions undergo regular, routine upgrades and obsolescence. Data files routinely accessed today may be impossible to read with the in-use software version tomorrow. Therefore, “software-neutral” storage solutions for your data archives are advocated wherever possible.

---

<sup>5</sup> Supreme Court Approves e-discovery Amendments to Federal Rules of Civil Procedure, Foley & Lardner, LLP, April 2006

### Off-Line Storage

Tape is an automated, cost-effective solution to ensure business continuity and maximized IT productivity. Tape will continue to be many times more cost effective than other storage solutions as it provides large inexpensive storage capacities. However, accessing data stored on tape seems slow because access is sequential rather than random. Human intervention is required to find and load tapes for eDiscovery. The amount of data stored on tape ranges from four to 15 times the amount of data stored on disks. Therefore, tape is ideal for serial processing applications, long-term storage and archiving options. Tape is also very good for back-up data that is further into the data lifecycle and can afford the time to respond to serial requests.

### Summary

In spite of what some marketers will tell you, no one has a one-size-fits-all solution.

We have Certified Business Intelligence Professionals who can help you sort out your data storage options. Using ARMA standards, we coach our clients toward a best-practices solution, and then help them implement that solution.

Using these same ARMA standards, we have developed a proven methodology that identifies dormant versus active data, resolves intricate related data dependencies, separates true dormant data from active production data, and gives you access when and where you need it.

We have demonstrated success in reducing by as much as sixty-percent the data volume in production OLTP systems with no negative impact on production operations. Backup times have dropped by as much as twenty-percent.



### **About the Author**

Charles Heishman is a founding partner of GPR Consulting, Inc. He is a Certified Business Intelligence Professional, has over twenty years of business and information technology experience, and has spent the past fourteen years working on ERP implementations and upgrades.

### **About GPR Consulting, Inc.**

GPR Consulting has experience and expertise in developing Information Technology solutions for a variety of businesses. Our clients come from diverse industries like Global Banking, Medical Research, Education, Utilities and Healthcare Providers.

We provide Professional Resources to compliment existing staff or create new teams to develop and implement your IT solution. GPR consultants are trained to work in small teams to successfully maintain and upgrade your IT systems without a lot of costly supervision on your part. We can provide on-site services or off-site consultants, matching our services to your needs.

GPR Consulting, Inc.

P.O. Box 9203

Chesapeake, VA 23321

Email: [info@gprcinc.net](mailto:info@gprcinc.net)

Phone: 888-718-1079

Fax: 800-863-6850