A SEMINAR

ON

DATA ARCHIVAL MANAGEMENT SYSTEM

BY

OLOPADE OLUWASEGUN MOSES

MATRIC NO: 16/69/0009

A SEMINAR SUBMITTED TO THE

DEPARTMENT OF COMPUTER SCIENCE

MOSHOOD ABIOLA POLYTECHNIC, ABEOKUTA.

SUPERVISOR:

DR MRS ALARAN

DECEMBER, 2022

TABLE OF CONTENT

Title page i

Table of content i

Chapter one 1

Introduction 1

1.0 Introduction 1

1.1 Importance of data archiving 1

1.2 Data archiving strategy 2

1.3 Common data archiving challenges 3

1.4 Rational approach to data archiving 3

Chapter Two 5

Literature Review 5

2.1 The evolution of data archiving 5

2.2 Data archiving methodology 5

2.3 Data archiving 5

2.4 Benefits of data archiving 6

2.5 Automated data archiving 6

2.6 Database archiving for long-term data retention 6

2.7 Open archival information system 6

2.8 Archive information system 7

Chapter Three 8

Discussion 8

3.1 What is archiving? 8

3.2 The importance of effective archiving 8

3.3 The benefits of archiving for growing businesses 8

3.4 Data archiving defined 9

3.5 Importance of data archiving 11

3.6 Data archiving considerations 11

3.7 Best practices for creating a strong data archiving strategy 12

3.8 Dos and don’ts for data archiving 13

3.9 Typical data archiving tools 17

Chapter Four 18
Conclusion 18
4.1 Conclusion 18
Reference 19

CHAPTER ONE

INTRODUCTION

1.0 INTRODUCTION

Many businesses and individuals hold onto a lot of data, especially after years of business. In some industries, it's vital that this information is organized and stored for later retrieval. The process of storing and organizing this data is called data archiving.

Organizations give a lot of emphasis and spend considerable time to build processes for entering data into an ERP (Enterprise Resource Planning) system. Organizations also spend a lot of time in developing processes for archiving manuals, drawings, specifications, and other hard-copy documents. But the question is -do organizations spend sufficient time on data archiving? In most cases, the answer would be either no or not yet ready.

Irrespective of its size and revenue, organizations that run Baan applications extensively should consider doing data archival. Depending on the data transactions volumes, typically an organization should consider data archiving after 4-5 years of going live. Archiving in general is the process of moving historical data from the operational environment to a special archive environment. Archiving historical data is an irreversible process. After data is moved to the archive company, the data can no longer be uploaded back into the operational company. Therefore, an organization must define a robust archiving strategy. Baan system administrators who are mostly responsible to undertake such project should have a strategy and address three major topics: What, When and Who.

1.1 IMPORTANCE OF DATA ARCHIVING

An organization has to decide what must be stored and for how long. Every finance manager must decide how long what data must be stored, in other words, kept in an operational environment for quick use. The prime reasons why organizations should consider archiving are listed below:

- **Improving the run time for generating financial reports.**

  In the operational system, the data used for running the reports is humungous. The system has to identify the relevant report from a vast selection and print the same which consumes a lot of time. By migrating the data there is ample space created in the operation system, which makes it all the easier for the system to pick up the relevant report in less time.

- **Improve MRP run times.**

  Material Requirements Planning (MRP) is a complex process imbibing all the elements of planning, procurement and manufacturing, from start to finish. The process is so intricate, that it eats up a considerable amount of time and also occupies sizeable space. Archiving proves to be a savior by migrating a huge chunk of the data and thereby creating a window for MRP runs to be processed in less time.

- **Make space for live server memory.**

  By deleting the obsolete data there is increased longevity, coupled with an enhanced efficiency.

- **Improve performance of job processing.**

  As data in the tables grow, running processes tend to slow down. This is where archiving makes its presence felt by unburdening the system and fueling it up to perform at high speed.

- **Data accessible anytime.**

  Historical data can be accessed for reporting purposes with read-only option.

1.2 DATA ARCHIVING STRATEGY

With growing data volume, increasing data security breaches, and complex application-performance issues, most organizations today face significant data management challenges. Keeping inactive data online not only creates security risk, but also increases infrastructure and database cost. Although there are several options available to control data growth, data archiving is the best long-term solution to significantly improve application performance and reduce the cost of servers and storage.

An organization's data archiving strategy is a plan to ensure that inactive data is stored, managed, secured, accessed, and destroyed according to the policies you define. Building a data archiving strategy requires the participation of several stakeholders, including data architects, enterprise architects, database administrators (DBAs), developers, information security and infrastructure professionals, business analysts and business users.

A data archiving strategy should focus on archiving inactive data periodically. After retaining that archived data for another period, consider moving it to longer-term offline storage such as tapes — though an organization's policies may differ depending on the need to keep data online versus offline. During archival of data there could be slow response on certain Baan processes- MRP, job processing etc. It is advisable that organizations must plan for the archiving service during "lean" business periods, for example during long weekends. The time to complete archiving depends on the data to be moved out of live company to the archiving company.

1.3 COMMON DATA ARCHIVING CHALLENGES

- **Allocation of resources:** On the flip side there are certain challenges to be overcome in archiving.
- **People:** Archiving is a separate world in itself, and to conquer it requires a certain amount of know-how and expertise. Users often need to be trained so as to operate it at full efficiency.
- **Budget:** Since we are creating a separate archiving company, it needs a certain level of investment and a separate budget allocation. This creates a certain amount of reluctance on the part of the management. However, this can be tackled by creating awareness about the merits of archiving, which overshadow its imperceptible imperfections. This is after all a small price to pay in view of the future returns on the investment.
- **Hardware:** Another impediment is the hardware set-up. As a general practice, the data is stored in different companies and each archiving company has at least a year worth of data. This calls for additional space in the server.
- **Timeline:** Timeline is another crucial factor playing truant. As we are migrating a huge volume of data there can be no accurate estimation of time that the system will take to run and complete the whole process. In addition, any error arising in the due course only serves to prolong the already tedious process.

PRE-REQUISITES

The objective of this seminar, is to create awareness about archiving data and how, consultants do this job meticulously.

ERP LN requires logistical and financial processes to be completed for the data to be archived. Logistics and finance are interwoven. Whether a logistic process is a purchase order or sales order, it has to be in essence wholly completed for it, to reflect in finance. The whole process can move ahead only if all the individual processes are in a completed status.

Archiving does not encourage any deficient data. Not understanding the recommended prerequisites give scope for confusion as to why some data is still found in live company

Certain data, like master data is essential for operational company and hence is retained in the operational company. However, the master data is copied (without deleting) to the archiving company to maintain data consistency.

1.4 RATIONAL APPROACH TO DATA ARCHIVING

Some organizations are not clear on what inactive data they must retain and what they may archive or purge. As a result, these organizations avoid archiving for fear of failing to meet compliance requirements. In addition, too many organizations manage storage reactively as opposed to proactively, hiding the cost implications of failing to act. Once established, an archiving strategy helps clearly define what data you must retain, what you can archive or purge, and how users should access archived data. When an organization is not properly aware about data archiving and

how to initiate the process, it is advisable to consult experts who have wide knowledge on the subject. Consultants who possess deep knowledge of financial and logistics process can be trusted to assign such task. Knowledge on archiving process is also a mandatory criterion a consultant must have. Listed below are the benefits that an organization will derive by hiring an expert consultant.

- **Assessment as to recommend archiving is needed or not**

  An expert consultant will decide on the necessity of archiving, depending on the company size and turnover. Small companies operating on a small database and small budget may not require archiving, as their database is relatively smaller and easier to manage. Archiving however is a must for companies operating on large database. A consultant would identify the companies that need archiving.

- **Providing assistance to educate users on the Baan/ERP LN process prerequisites**

  Imparting training to the users would help them understand the setup, its importance and speed up the performance of both the users and the system.

- **Provide help in executing the archiving process**

  The users would be provided continuous support and back up in executing the archiving process from the beginning.

- **Provide tools to help users measure the archiving process success**

  The time taken to run certain reports/processes before archiving and after archiving is compared and recorded in a log. This enables the management to compare and analyze the time factor involved in archiving and how the processes are considerably speeded up after archiving.

- **Project management support**

  A consultant with wide expertise and knowledge can foresee which activities take time and accordingly advise the management on the sluggishness of the system during the archiving process. As a result, the organization can prepare itself for the delay. The schedule can be adjusted accordingly by the management based on the delayed performance of the system.

Having a good understanding of the process, methodology, tools coupled with the knowledgeable advice provided by expert consultants would make your management of data a simple task, thereby achieving greater operational efficiency and resource productivity. After you define an archiving strategy that suits your requirements, you can define the archiving plan with the help of an expert. In this plan, you translate the archiving strategy into a more operational level.

CHAPTER TWO

LITERATURE REVIEW

2.1 THE EVOLUTION OF DATA ARCHIVING

According to Ortiz (2009) over the last few years there has been an increasing amount of interest in data archiving. Factors such as the explosive growth of data quantities on corporate networks along with the need to retain more and more of this data for longer periods in order to meet various legal and corporate governance requirements, and the need to reduce costs wherever possible, are driving this growing interest. Along with this growing interest the evolution of data archiving that has now culminated in Cloud Based Archive Solutions. The most common method of achieving this goal is through the process of archiving, which, unlike backups, moves inactive data from primary disk-based storage to an easily accessible, less expensive secondary storage tier then deleting it from the disk source locations. This can help realize the objective of cutting costs by freeing up expensive primary storage, reducing backup windows thereby increasing operational efficiency and providing reliable protection of the data for the long term. A viable and effective data archive should provide: scalability, cost effectiveness, availability, and secure long-term protection for the data.

2.2 DATA ARCHIVING METHODOLOGY

Described below are the critical steps adopted in data archiving in ERP LN:

- Allocate space in the RDBMS/Server for the archiving activity
- Creation of archiving company as a mirror image of operational company with all the required activities performed in company setup
- Maintain common master data and the parameters
- Start archiving process using the standard sessions in the application after completing the pre-requisites for each session individually
- Straightening out the errors if any and re-process the same session
- Check the data in archiving company after completion of all the sessions
- Transfer/relocate the data to external device or another server

2.3 DATA ARCHIVING

According to Rouge (2010) data archiving is the process of moving data that is no longer actively used to a separate data storage device for long-term retention. Data archives consist of older data that is still important and necessary for future reference, as well as data that must be retained for regulatory compliance. Data archives are indexed and have search capabilities so that files and parts of files can be easily located and retrieved. Data archives are often confused with data

backups, which are copies of data. Data backups are used to restore data in case it is corrupted or destroyed. In contrast, data archives protect older information that is not needed for everyday operations but may occasionally need to be accessed.

2.4 BENEFITS OF DATA ARCHIVING

Archiving is one of the best ways to manage uncontrollable data growth because it will help you identify which files are no longer active within your organization and, thus, not required on your primary storage tier. The requirements of your data (value to the organization, speed of access, security, etc.) will change over time so both the media and location of this data should also change to more cost effectively meet those requirements. Archiving software helps you achieve this.

2.5 AUTOMATED DATA ARCHIVING

Data archiving is essential for organizations that accumulate new information but still need to retain older information. The trends of corporate and agency policy, legal precedent, and government law and regulation are for longer retention, more information, and faster retrieval. Automated data archiving helps organizations to achieve these capabilities at lower costs.

2.6 DATABASE ARCHIVING FOR LONG-TERM DATA RETENTION

According to Craig S. Mullins (2006), organizations are generating and keeping a more data now than at any time in history before. This is so for many reasons. First of all, the amount of data in general is growing. According to industry analysts, enterprise databases are growing at the rate of 125% annually. Even more interesting is that as much as 80% of the information in those databases is not actively used (in other words, it is ready for archiving). Data may need to be retained for both internal and external reasons. Internal reasons are driven by company needs. If an organization business requires the data to conduct business and make money then that data will be retained. Today's modern organizations are storing more data for longer periods of time for many internal reasons. Typically, data is stored longer than it used to be to enable analytical processes to be conducted on the data. Data warehousing, data mining, OLAP, and similar technologies have delivered more and better techniques for extracting information out of data. So, businesses are inclined to keep the data around for longer periods of time. But external reasons, typically driven by the mandate to comply with legal and governmental regulations are another significant factor driving the need to store more data.

2.7 OPEN ARCHIVAL INFORMATION SYSTEM

An Open Archival Information System (or OAIS) according to Wikipedia.org (2013) is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. The term OAIS also refers, by extension, to the ISO OAIS Reference Model for an OAIS. The information being maintained has been deemed to need "long term preservation", even if the OAIS itself is not permanent. "Long term" is long enough to be concerned with the impacts of changing technologies, including support

for new media and data formats, or with a changing user community. “Long term” may extend indefinitely. OAIS archive must abide by: Negotiate for and accept appropriate information from information Producers; the community should be able to understand the information without needing the assistance of the experts who produced the information; Make the preserved information available to the Designated Community.

2.8 ARCHIVE INFORMATION SYSTEM

The Archive Information System (AS) is an SAP data archiving environment-integrated generic tool for indexing data archives. The indexes that are created, called the Archive Information Structure, let you display the archived data. The archive information structure includes the following components: Archive Retrieval Configurator (ARC)-The ARC enables you to create archive information structures with the help of field catalogs, and to fill the structures with data from the archive. The archive information structure, which represents a kind of archive index, provides the basis for archive data reporting; and Archive Explorer-The Archive Explorer enables fast searches of archived data. It does this by accessing the archive information structures that have been created and stored in transparent database tables using the Archive Retrieval Configurator. Furthermore, the Archive Explorer allows direct accesses of individual data objects in the archive, which can then be displayed in both technical and application-specific views.

CHAPTER THREE

DISCUSSION

3.1 WHAT IS ARCHIVING?

Archiving is the process by which inactive information, in any format, is securely stored for long periods of time. Such information may – or may not – be used again in the future, but nonetheless should be stored until the end of its retention schedule.

It should be emphasized that archived, inactive data can be made active again, as the implication of not being able to access information again sometimes dissuades businesses from archiving their records. Information should still be readily available if required at short notice.

Archived data can be stored in a number of different formats, on a variety of devices.

When businesses archive data they tend to archive entire collections of files. The word archive in itself suggests the retention of multiple records. Data that is archived might need to be because of legal rulings or because it is vital information that is likely to be used again in the future.

3.2 THE IMPORTANCE OF EFFECTIVE ARCHIVING

Archiving is vitally important for information management and can give an organization, institution, and business greater control over their information processes.

As an organization, institution, and business grows it will create more data – data that needs to be meticulously managed and monitored in order for it be utilized properly. Keeping tabs on this data can prove problematic for businesses that never put an archiving system in place.

Data that is not archived is harder to locate, secure and appropriately disseminate if stored in a local environment – such as storage devices – and thus will be inaccessible to other users. This will eventually have a negative effect on an organization, institution, and business functions and productivity.

3.3 THE BENEFITS OF ARCHIVING FOR GROWING BUSINESSES

- **Prevent Data Loss**

Information that hasn't been archived on a central, secure repository could be lost forever. The chances of an employee accidentally deleting or misplacing a file is slim, but it does happen. In some cases, data recovery experts might be able to retrieve this information, but this is likely to take some time, cost a lot and is rarely 100% accurate.

Archived data allows employees to retrieve back up information independently without having to rely on the expertise third parties.

- **Legal Requirements**

Archiving is important for legal reasons too. Many businesses accidentally dispose of documents that they legally should be keeping.

An effective archiving system will ensure company-specific retention schedules are adhered to, regardless of each employee's knowledge of these schedules. Data protection authorities are enforcing more severe penalties on businesses so employees should be made aware that ignoring these policies could lead to hefty fines or even prison sentences in some cases.

- **Increase Security**

Archiving is also important for security reasons, especially at a time when cyber-attacks and data breaches are becoming more frequent.

By securely archiving documents businesses can keep track information and increase protection from unauthorized third parties. Even the most cautious of businesses are now targeted by very adept hackers.

Paper records in open circulation can easily be taken from crowded offices or stolen by bitter employees. A reliable offsite archiving system will reduce this risk by warning senior staff when files are missing.

Archiving is vital for business continuity and ensuring the highest level of performance in a competitive marketplace.

3.4 DATA ARCHIVING DEFINED

Data archive is simply a collection of data or information that is stored in an organized manner. This information can be quickly accessed for later retrieval or compliance requests. In the tech world, an archive is different than a backup. Backups are more focused on speed. This is information that can quickly be retrieved to be edited or updated. Archives, on the other hand, are filled with data that are no longer being updated or revised. These are static documents kept on hand for documentation and record keeping.

Data archiving is the practice of identifying data that is no longer active and moving it out of production systems into long-term storage systems. Archival data is stored so that at any time it can be brought back into service.

Data archiving strategy optimizes how necessary resources perform in the active system, allowing users to quickly access data archive storage devices or data archiving plans for easy retrieval and more cost-effective information storage. It also clarifies how users should move data for best performance within applicable regulations and the law.

Data Archiving is the practice of shifting infrequently accessed data to low-cost storage repositories. It is an important part of a data management strategy. The goal is to reduce costs on

warm storage while retaining old data needed for future reference or analysis, and information needed for regulatory compliance.

A Data Archive is often built using cold storage tiers, which can hold large amounts of data at a low cost. There are many types of archival architecture and features, each meeting different needs. However, the minimum requirements are indexing and searching capabilities, which ensure files remain easily accessible.

3.4.1 ONLINE VERSUS OFFLINE DATA ARCHIVING

When starting your own data archiving system, you'll need to decide what kind of storage you'd like to use. There are several different options for both online and offline data storage. Let's review a few of them.

3.4.2 MOBILE MEDIA

Many people choose to store their data on mobile media devices such as USBs, external hard drives, tapes, CDs, DVDs and the like. There are benefits and drawbacks to each but the main thing you need to realize is that data archiving via mobile media gives you the ability to take those documents wherever you go. For instance, if you're a photographer and you're meeting with a client for the first time, you may want to take your external hard drive with you. If you've consistently archived your photos and stored them on that drive, you'll be able to quickly access and show off your work to potential clients. They'll also be happy to see that you are so organized, giving you more credibility and authority in your field.

3.4.3 CLOUD STORAGE

Another popular choice for data archiving is cloud storage. Cloud storage is the act of storing your data within an interface that is hosted and managed by another company. One simple example of cloud storage that many people use on a daily basis is Google Drive. This tool is synched with your Gmail account so if you have that, you have access to Google Drive as well. This tool enables users to efficiently store, add, share and edit documents, spreadsheets, photos, pdfs, etc. and for the everyday person, this is enough storage. Google Drive also has apps for Android and iOS so files can also be managed from mobile devices. Professionals may want to look into other heftier cloud storage providers such as Carbonite. Carbonite is a cloud storage provider that offers unlimited cloud storage and a variety of different plans for both businesses and individuals. Services like this often have a monthly or yearly fee, but if you need secure cloud storage, the fee is worth protecting your data.

3.4.4 GETTING ORGANIZED

Once you've determined what kind of data archiving storage system you prefer, it's time to get organized and come up with a file naming system. It's one thing to dump all your data into a folder

on Dropbox or your external hard drive, and it's another thing entirely to clearly depict and separate that data into logical chunks.

For instance, let's say you have 2,000 photos on your computer that you want to archive with cloud storage. Before you dump them all onto the cloud, create individual folders with a clearly defined format. Create individual folders labeled "2016 PHOTOS" and inside that folder, break your data down even further into separate folders such as "FAMILY REUNION" or "GRADUATION."

There's no wrong way to use a file naming system – just use one that works best for you!

3.5 IMPORTANCE OF DATA ARCHIVING

The main benefit of archiving data is the ability to remain organized. Whether you're a self-employed professional, a major executive for a large corporation, or simply trying to keep track of your family history, archiving your data will ensure that you keep photos, documents and other media safe for the long haul while still being able to easily find and access them for later use.

Data archiving is primarily used to eliminate the need for large backups. While it's wise to backup data you use regularly, data that hasn't been used in forever can be moved into the archive and stored. Additionally, storage costs for archiving data are typically less expensive than primary storage because it's focused on large capacity, not speed.

A data archiving plan is an important part of your data lifecycle management policy, providing you with a way to retain information while staying within a reasonable storage budget. A data archiving implementation typically involves supporting tooling and automation, which help drive efficiency into the process. Here are key functions of a data archiving solution:

- **Data Discovery:** A data archiving solution can help admins and end-users to easily find files, including spreadsheets, documents, and presentations.
- **Data Management:** A data archiving solution can help you locate redundant data and remove it, or remove aging files from your file's servers.
- **Data Visibility:** A data archiving solution analyzes, classifies and indexes data before storing it on your servers. This process ensures you can easily perform searches and gain insights.
- **Data Compliance:** A data archiving solution automatically prepares responses to requests, including compliance audits, business queries, and litigation.

3.6 DATA ARCHIVING CONSIDERATIONS

The main advantages of data archiving are the ability to significantly reduce storage costs. However, a data archive is not ideal for all use cases. You should not, for example, use a data archive instead of a backup solution.

3.6.1 DATA ARCHIVING VS BACKUP

While data archives and data backups are used as secondary storage repositories, the two do not provide the same value.

A data backup is a copy of data created for the purpose of protecting and recovering data. Backups of data contain critical information needed for quick recovery during data loss or other disasters.

A data archive contains data kept for the purpose of retaining the information for the long term. Data archives often retain infrequently accessed data, which is not critical for recovery or needed for ongoing business continuity.

3.6.2 DATA ARCHIVING AND DATA LIFECYCLE MANAGEMENT

A data archive and a data lifecycle management perform different functions.

Data lifecycle management processes manage the entire lifecycle of data, from the time a piece of data is created and until it is deleted. Organizations create data lifecycle management policies, which are enforced by administrators and management tools.

Data archiving is often created as part of an overall data lifecycle management program. Organizations create data archiving policies, and the process is then fully automated by an archiving software.

3.6.3 DATA ARCHIVE AND COMPLIANCE

Organizations can use data archives to retain general information and to retain information needed for compliance.

A data archive can be used to retain information needed for compliance purposes. This data is often retained for future compliance audits. In this case, the data needs to be easily accessible, to ensure regulatory entities can quickly reference the data.

Organizations often use data archives to story aging data of all kinds. Some data, like corporate information, can be kept according to company policies. Other types of data, like private information, needs to be compliant with relevant regulations. In this case, there are certain regulations and standards the archive must meet in order to ensure compliance.

3.7 BEST PRACTICES FOR CREATING A STRONG DATA ARCHIVING STRATEGY

Here are several best practices to consider when creating your own data archiving strategy.

- **Identify and Sort Data Before Archiving**

  Take a look at your data and create an inventory. Categorize data into types and then prioritize, carefully considering which data is needed for ongoing operations and which can be moved to

the archive. If you need to archive both structured and unstructured data, determine whether you want them stored in separate repositories or in one centralized archive.

- **Synchronize Data Archival and Data Lifecycle Management**

  A data archive has an impact on the data lifecycle. When you create your archival plan, consider the lifecycle of the archived data. For example, ask yourself for how long you want to retain archival data and when archival data should be purged. Update your data lifecycle management strategy accordingly. Be as specific as possible.

- **Plan for Regulatory Compliance**

  Make sure that your archive strategy is fully compliant with relevant compliance regulations. If you need to enforce strict compliance standards on half of the archival data, but the other archival data can be easily stored, then you might want to consider creating two separate archives.

- **Select the Right Archiving Tool**

  There are many types of solutions that can help support your data archiving strategy. Some tools are dedicated to data archiving, while other solutions offer end-to-end data solutions with built-in archiving capabilities. The most important feature of any tool is an efficient search engine.

- **Develop a Data Archiving Policy**

  Once you have assessed all the requirements for your data archive strategy, you can create a comprehensive data archiving policy that clearly documents all relevant processes and procedures. The policy should include all data archiving criteria and mechanisms, as well as specific roles and circumstances that determine when and who can access archived data.

3.8 DOS AND DON'TS FOR DATA ARCHIVING

- **DO - Clearly differentiate between your archive and your backup solutions.**

  During the past couple of years, some vendors have attempted to blur the line between backups and archives. Even so, backups and archives serve two distinctly different purposes and should be treated as such.

  Backups are created as a data recovery mechanism. In contrast, archives are designed for long-term storage of data and are not typically treated as a disaster recovery mechanism.

  Because of this distinction — and the fact that archived data should never be modified — some healthcare organizations do not back up or replicate their archives. However, archives can be prone to failures resulting in data loss, just as any other type of IT system can. Although traditional nightly backups probably aren't appropriate for archives, you do need a redundant

copy of archived data. For example, an organization that keeps its archives on storage area network (SAN) storage might use storage replication as a way of ensuring that a redundant copy exists, and is automatically kept current or can be replicated to the cloud to ensure that it is available in the event of a disaster.

- **DON'T - Configure archive storage in a way that could potentially become a single point of failure.**

  All too often, an organization's archives are treated as something of an afterthought. Data retention is an operational requirement, but because archives are filled with data that is rarely accessed, there might be pressure to use the least expensive storage solution possible.

  There is nothing wrong with using low-cost hardware for storing archived data, but the archives must not be constructed in a way that could result in a single point of failure. If an organization's archive system were to fail, a huge volume of data could be lost — and HIPAA regulations stipulate severe penalties for this type of data loss.

  Even if a component failure didn't cause any actual data loss, it would result in downtime for the archive system. Archive contents might be rarely accessed, but when someone does access the archives, there is usually an important reason for it. As such, it is important to construct the archive system in a way that prevents any component from becoming a single point of failure. Doing so can help prevent data loss and outages.

- **DO - Establish clear policies defining what data should be stored in which tier of storage.**

  The volume of data within a healthcare provider's archives can be huge, and there are costs associated with storing vast quantities of data. Organizations typically attempt to curb this cost by using multiple storage tiers. For example, archives might be spread among SAN storage, cloud storage and removable media such as tape, with each medium acting as a separate storage tier.

  It is important to have policies in place that stipulate what each storage tier is used for. These policies are usually based around the age of the data, data type or some combination of the two. For example, a healthcare provider may wish to store patient health data that is less than 5 years old on spinning disk so that it is readily accessible, while moving older data to tape or the cloud. This approach can help the organization keep the most current data readily accessible, while also reducing storage costs by placing the oldest data onto less expensive media.

- **DON'T - Treat all of your data the same with regards to archival.**

  Just as it is important to have policies in place that dictate how the various tiers of archival storage are to be used, it is also important to realize that it is a mistake to treat all of your data equally with regard to archiving. Data archiving requirements tend to vary widely based on data type.

These variances are often based on retention requirements. For example, a healthcare provider might have internal operational requirements that require it to retain patient data for 10 years after the patient has died. That same organization might need to keep financial data for only seven years.

It is also important not to overlook the fact that some data may not need to be archived. For example, a healthcare organization probably would not need long-term archiving of marketing materials or PowerPoint presentations unless there were operational requirements to do so.

- **DO - Take steps to automate data lifecycle management.**

  Regardless of whether a healthcare provider is large or small, it is unrealistic to expect that the data archiving process can be done manually. Manual archiving is impractical for a number of reasons.

  For starters, like any other manual process, manually archiving data is a labor-intensive process. Although this process might at first seem manageable, history has shown that data grows exponentially over time, and healthcare data growth rates today are nearing 40%. Therefore, what might be a somewhat manageable process today will likely become unwieldy and out of control later on. Even if the process were to somehow remain manageable for an indefinite period of time, manually archiving data is a poor use of your IT staff's time, especially when you consider that there are automated data lifecycle management products readily available.

  Another reason why you should automate data lifecycle management is that it is difficult to perform manual archiving on a consistent basis. Manually performing data lifecycle management probably isn't one of those tasks that an organization performs on a daily basis. As such, it is easy for the IT staff to forget to periodically manage the data.

  A third reason why manual data lifecycle management is a bad idea is because the process is prone to human error. It is all too easy for data to accidentally be moved to the wrong storage tier or to be prematurely purged. In the case of patient health data, there can be severe consequences for this type of data mishandling. An upfront investment in data lifecycle automation software can save a substantial amount of money in the future in the form of labor costs and, potentially, in the avoidance of fines from the U.S. Department of Health and Human Services.

- **DON'T - Underestimate the importance of security — especially when archiving data to the cloud.**

  Even if your archives consist entirely of very old data, the content is sensitive. Organizations do not archive data unless it is potentially useful in the future or is needed to address compliance around long-term retention requirements. Archives may contain financial data, protected electronic health information, personally identifiable information about the

organization's employees, or just about anything else that can be imagined. Given the sensitivity of the archived data, it is critically important to make sure that the archives remain secure. This is especially true when archiving data to a cloud storage provider, since the cloud storage is outside of your direct control.

The Omnibus rule raises the stakes even further. Under the rule, cloud storage providers are treated as business associates. Furthermore, a covered entity shares the responsibility for security breaches with its business associates. This is why it is so critically important for healthcare providers to evaluate a cloud storage provider's security and to ensure that the provider is fully HIPAA-compliant.

- **DO - Assess the benefits and risks of various storage solutions.**

  There are a number of different options for storing archived data, and there are advantages and disadvantages to each type of storage. As such, it is important to examine the pros and cons of the available storage solutions before committing to an archive method. For example, tape has the advantage of being a highly reliable storage medium with a very low cost per gigabyte of storage. On the other hand, data that is stored on tape is less accessible (the tape must be mounted before the data can be read). Furthermore, the portable nature of tape makes it possible for vast quantities of archived data to walk out the door, so it is important to ensure your tapes are tracked and stored using a secure chain of custody

  On-premises online storage is another popular option. This option uses spinning disks to store archived data, usually on a SAN or network-attached storage device. The main advantages to using this type of storage are that the data is easily accessible and that fault-tolerant technologies can be used to safeguard against disk failures. Conversely, this type of storage tends to be significantly more expensive than removable media. Although on-premises online storage can be scaled, doing so consumes floor space and increases electric and cooling costs. Furthermore, the storage will eventually become obsolete, thereby leading to hardware replacement and data migration costs.

  Cloud-based storage is often thought of as a happy medium between on-premises disk storage and tape storage. Cloud storage is generally less expensive than on-premises disk storage, and storage capacity ceases to be an issue. However, the fact that your data is being stored outside of your direct control means that security and the cloud storage provider's reputation are of critical importance.

- **DON'T - Try to go it alone.**

  Finally, don't go it alone when implementing a data archiving system. Healthcare providers are subject to very strict data retention requirements, and there can be severe penalties for failing to adhere to those requirements. Given the exceedingly high stakes surrounding data lifecycle management in healthcare organizations, it is a good idea to enlist the help of an

organization that focuses on helping healthcare IT to properly implement data archiving and data lifecycle management solutions.

3.9 TYPICAL DATA ARCHIVING TOOLS

Different data archiving tools and data archiving plans have unique benefits and life expectancies. How much data is being processed is just one of the considerations that will control what the best archival data solution is for your organization.

Tapes, disks, flash storage, hard drives, and cloud data archiving are all possible storage mediums. For many larger companies, virtual archives such as cloud archival data sources or archive data software may be a better choice given the weaknesses and costs associated with storing other long term data archiving solutions. Cloud storage also offers high capacity with lower storage costs.

An additional issue to grapple with when choosing a long-term data archiving option is that today's interfaces eventually become obsolete. This is why updating your devices and conducting routine audits of your data archiving media interfaces is also a best practice. Using a cloud data archiving system automates that process.

CHAPTER FOUR

CONCLUSION

4.1 CONCLUSION

This seminar reports on the implementation of data archival management system. Informed and vital decisions are made on the basis of the information that is available to either individuals or organizations. Since information is important, it is needful to achieve and manage them properly for ease of their retrieval when needed. Nonetheless, the sorting, storing, searching, retrieving and moving of files has hitherto been done manually and with papers which is inefficient, tedious, and time-consuming. This research implements a data archival and management system in Moshood Abiola Polytechnic, Ojere, Abeokuta which solves the challenges and complexities in the manual system of archiving and managing of information in the institution.

This seminar aimed at the research of developing a complete, fully working data archival management system for Moshood Abiola Polytechnic, Ojere, Abeokuta. The system was researched to be designed primarily to handle the long-term storage and management of files in the institution as well as improve on already existing systems by implementing the feature/function of direct file transfer from one user to another. To the best of the researcher's knowledge, as of the time of this seminar, such a system had not been created. This system proves very useful and putting it to use will help create an environment with greatly reduced use of hardcopy files and a smoother and more efficient work flow as it solves a good number of practical problems. Although this seminar was developed specifically for Moshood Abiola Polytechnic, Ojere, Abeokuta, with little changes and modifications made, it can be used by anyone and in any organization. This system is a viable one, in the sense that it will go a long way in making work environments more convenient, efficient and productive for its users. Because of its importance the researchers strongly recommend the use of the system as its use is not restricted only to schools but also can be used in banks, hospitals and so on. This seminar achieved the initially set out goals and objectives.

REFERENCE

Aadamsoo, A. (2010). Web Based Project Management System [Scholarly project]. Retrieved June 29, 2019 from https://www.theseus.fi/bitstream/handle/10024/16996/Aadamsoo_Anne-Mai.pdf

Fiona, A.C. (2010). Development of a Data Management System for Students ‟ Final Year Projects Case Study: Department of Information Systems [Scholarly project]. Retrieved May 5, 2019 fromhttp://103.23.20.161/index.php/semnasif/article/view/1196

Garderen, P.V., Jordan, P., Hooten, T., Mumma, C., & McLellan, E. (2012) The Archivematica Project: Meeting Digital Continuity's Technical. Retrieved June 23, 2019, from http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/VC_VaGarderen_et_al_26_Workshop1.pdf

Hussain, A., &Mkpojiogu, E.O.C. (2017). Predicting the perceived worth of software products requirements with customer satisfaction. Advanced Science Letters. 23(5), 4269-4273.

Hussain, A., Mkpojiogu, E.O.C., & Abdullah, I. (2016). Requirements engineering practices in UUMIT centre: an assessment based on the perceptions of in-house software developers. Journal of Telecommunication, Electronic & Computer Engineering (JTEC). 8(8), 27-32.

Joshi, S.R. (2011). A Web Based Content Management System [Scholarly project]. Retrieved July 31, 2019 from http://sdsudspace.calstate.edu/bitstream/handle/10211.10/1113/Joshi_Satya.pdf

Kemppainen, T. (April, 2015). Data archive project [Scholarly project]. Retrieved May 5, 2019 from https://www.theseus.fi/bitstream/handle/10024/93069/Kemppainen_Timo Pekka. pdf? sequence=1&isAllowed=y

Malik, S., Al-youbi, S.A., Al-youbi, S.A, Hussein, R., Dweik, G., &Arif, S. (2018), Web-based Archive Management and Student Guidance for Final Year Projects, Journal of eLearning and Higher Education,2018 (2018), Article ID 871144, DOI:10.5171/2018.871144

Meimaris, M., Papastefanatos, G., &Pateritsas, C. (n.d). An Archiving System for Managing Evolution in the Data Web. Retrieved May 5, 2019 from http://ceur-ws.org/Vol 1377/paper2.pdf

Mkpojiogu, E.O.C. & Hussain, A. (2017). Can scoring rubrics be used in assessing the performance of students in software requirements engineering education? Journal of Telecommunication, Electronic & Computer Engineering (JTEC).9 (2-11), 115-119

Mkpojiogu, E.O.C., & Hashim, N.L. (2017). Quality-based prioritization: An approach for prioritizing software requirements. Journal of Telecommunication, Electronics & Computer Engineering (JTEC), 9 (2-2), 17-21. E-ISSN: 2289-8131.

Mutale, B.M., & Phiri, J. (2016). Web Based Document Archiving Using Time Stamp and Barcode Technologies – A Case of the University of Zambia. International Journal of Innovative Research in Science, Engineering and Technology, 5(4) 2319-8753. Retrieved July 5,2019 from https://pdfs.semanticscholar.org/076a/668fd87d0dd8e7fa5f5c4710b818d4e697bc.pdf

Nicholas, A.O., Ambrose, A.A., Adewole, A., &Edeh, E.E. (2014). Development of a File Tracking System for Tertiary Institutions [Scholarly project]. Retrieved July 31, 2019 from http://eprints.covenantuniversity.edu.ng/5567/1/1948.pdf

WEB REFERENCES

http://en.wikipedia.org/wiki/Open_Archival_Information_System

http://searchdatabackup.techtarget.com/definition/data-archiving

http://www.tdan.com/view-articles/4591

https://bluexp.netapp.com/blog/clc-blg-data-archiving-the-basics-and-5-best-practices

https://coservesolutions.com/how-to-adopt-a-strategic-approach-to-data-archiving/

https://getprostorage.com/blog/a-beginners-guide-to-diy-data-archiving/

https://www.inettutor.com/source-code/data-archiving-information-system-review-of-related-literature/

https://www.ironmountain.com/resources/whitepapers/d/6-dos-and-donts-for-data-archiving

https://www.securedatamgt.com/blog/what-is-archiving/