

# Life: A Huge Archive Electronic Archive Has Become an Urgent Necessity in the Face of Enormous Technological Advances

Ahmed Yousef Laarfi

School of Engineering and Science, Florida Institute of Technology (FIT), Melbourne, FL, USA

Email: [alaarfi2015@my.fit.edu](mailto:alaarfi2015@my.fit.edu)

**How to cite this paper:** Laarfi, A.Y. (2020) Life: A Huge Archive Electronic Archive Has Become an Urgent Necessity in the Face of Enormous Technological Advances. *Journal of Computer and Communications*, 8, 1-10.

<https://doi.org/10.4236/jcc.2020.84001>

**Received:** January 29, 2020

**Accepted:** March 28, 2020

**Published:** March 31, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Smart devices collaborate in increasing the quantity of information hugely. Therefore, traditional archiving couldn't support this explosion of information. The electronic archive will lead to many advantages such as faster access, privacy and security, data accessibility, and distribution of authorities between administrative levels. This paper will discuss electronic archiving and whether it is a necessity or an administrative luxury. This research shed some light on three main topics, which are general archiving, distributed database systems, and a case study.

## Keywords

Archiving, Distributed Database Systems, Indexing And Retrieving Data, Electronic Archiving

---

## 1. Life: A Huge Archive

### 1.1. Introduction

In this time, information becomes an expensive commodity. Who owns the information owns many valuables. Two types of information are used; one of them is confidential and private, like the military, political, and economic documents, and the other can be widespread such as books, publications, and education in general. Mankind is concerned about information because it is vital to life issues, such as commercials regarding competitive, scientific, especially secret research, and even personal information can be kept. Internet of things has huge information that is available in the world since the beginning of creation to the 1970s is equivalent in size to the information published on the Internet in the last thirty

years until the beginning of the third millennium [1] [2] [3] [4]. The amount of information available that year is around 500 billion Gigabyte. These amounts of information are unimaginable!

## **1.2. Background**

In the era of empires like Great Britain and France, especially after the tremendous colonial expansion in the 18<sup>th</sup> century, the need for the existence of archives for each colony had emerged. On the other hand, some historians went back with the archive to a historical period preceding the age of Empires when touched the Assyrian civilization in Iraq, for example, the Library of Babel that considered as the oldest archive on the planet [5]. Not only imperial expansion in the eighteenth century was the unique reason for the establishment of archives but also the urgent need to build the archives was found before those centuries, which represented in the presence of civilizations such as Pharaonic, which had had their archives to manage their states. In those historical periods, the archive was a place compiling historical documents. Books, manuscripts, sculptures in museums, internal and external correspondence, including information on population numbers and their whereabouts were also essential components of the archive [5]. The archive can be defined as the collective memory of the nation. In our contemporary history, inventing the computer and the Internet of things turned a look of technicians to the digitalization of archives [6]. Although some people believe that electronic archive is an administrative luxury, the Distributed Database Systems (DDBSs) make it a big necessity; JOWFE Oil company automated their archive and manipulated their database as everyday use.

## **1.3. Documenting, Indexing, and Archiving**

### **1.3.1. Documentation**

Documentation concepts are the same, whether in regular or electronic archiving, even if different methods are used. The scope and characteristics can distinguish the process of documentation entirely and the outcome of the process. In other words, the conclusion which excluded from what was written that the documentation includes all activities that fall within the scope of any documentary process by the methodology used in the documentation and archiving, for instance: Compilation of information from the data sources that used to support the reliability and how can convince that this data is correct and can be trusted. "Data is a symbol set that is quantified and/or qualified. Information is a set of significant signs that has the ability to create knowledge [7]." Therefore, data is one of the components of the information. Documenting this information obtained is the first step of archiving.

Classification of Documents or Records: Records classified depending on their activities into [8] [9] (**Table 1**).

### **1.3.2. Indexing**

It is a crucial to retrieve the information that has been stored. The known primitive

**Table 1.** Three types of documents.

Active	Also, they called “Current records,” which Documents mean the documents that in use currently.
Inactive Documents	They called “Noncurrent records,” which means they are no longer used and preserved somewhere for returning back to them from time to time. They also include legal or historical documents. for example
Semi-active Documents	Also, called “Semi current Records,” which means the Documents are not being used on a daily basis or periodically and are rarely resorting to it and placed nearby.

way of indexing was cataloged. Now, many electronic ways are followed. [10] The indexing, therefore, is the most critical step for the classification and putting the documents in the archive. However, most systems currently reverted indexing is the Dewey Decimal System and the Library of Congress System. Regarding the Distributed Data System, and several methodologies based on the types of databases and programming languages used [8] [9]. The comparison between ordinary and electronic archives shown in **Table 2**.

### 1.3.3. Archiving

After Documentation and indexing, the stage, were setting the record or document in an accessible place come. This stage should take place at the required speed by the time it should be required to achieve the desired result. “Archives’ [have] three [earnings] [11]. [Archives] can be: 1) the physical records which are collected; 2) the physical space which houses archival records; 3) the department or unit which manages the records.” The definition leads to that archiving is the last step to the process where the documents should be found.

## 1.4. Distributed Database Systems

### 1.4.1. History and Improvement

Before inventing the networks and the Internet of things, the information had stored in computers isolated from each other’s, which lead to significant difficulties in the exchange of them. Improving networks leads to relying on the exchange of information across networks. The emergence of the divergent networks and manipulating databases from different sources has been needed; therefore. Up to now, this subject has been developing over time.

### 1.4.2. Data Everywhere

Two things have become important characteristics of the data. First is spread all data over the places and the second is the possibility of access to the most data that allowed to reach. There is security data which are not approved for each one to have a look at. These data might be put either in remote computers or internal networks which access has to be restricted to. In general, access to this data will administrate through the database management systems, for example, a user at any University can access to existing data in any library in Japan or Brazil. This issue is directly related to the electronic archiving and distributed database management systems concepts together.

**Table 3** Distributed Database Management Systems, retrieved from [12].

**Table 2.** A comparison between traditional and electronic archiving.

Comparison	Traditional archive	Electronic Archive
Document Recovery	minutes to hours	Seconds
Missing Documents	Sometimes	Not at all
Document Saving	Minutes to hours	Seconds
Document Sending	Long time and sometimes resent back. The fax machine gives unclear copies	Easy and faster to be sent resent back. The fax machine gives unclear copies.
Document Sharing	A lot of copies	Just taking the authority
Dedicated places	Large and expensive	Small and cheap
Customer Services	Takes a time	Less time

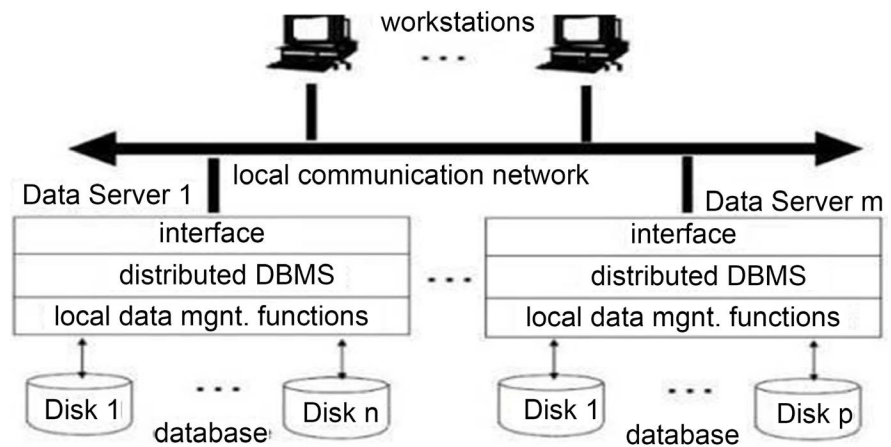
**Table 3.** DBMS advantages and problems.

Distributed DBMS	
Advantages	Problems
Improved performance	Complexity design and of implementation
Efficiency	Data consistency
Extensibility (addition of new nodes)	Safety
Transparency of distribution	Failure recovery
<ul style="list-style-type: none"> <li>● Storage data</li> <li>● Query execution</li> </ul>	

No one of users concerns about what stranded behind the interface. According to what is seen in the above graph (1), the users have not paid attention to the technical complexities as much as what to watch and use as an interface. All that user interest in is how can get the appropriate information, regardless of what it conceals a range of complex technical links. Moreover, **Figure 1** shows only where the information is stored and a user who wants to use it.

Three facts should be mentioned. First, it is not known for the user what operations are done and how they work. This disappeared behind the interface. Second, electronic indexing is very complex, and in seconds can conduct any data which will be needed. Third data storehouses are collections of electronic archives scattered in multiple locations. The three points together might be called data transparency which defined [13] as “Degree to which system, may remain unaware of the details of how and where the data items are stored in a distributed system.”

Many people agree that the Archive is an ancient technique to save documents, valuables, and records in a safe place where can refer to later and restore the required data. Upon inventing it, the computer has been exploited in many areas of public life, especially while computer networks have been introduced. A vital usage has applied when the computer involved in distributed databases, electronic archives which are represented in databases deployed. Consequently,



**Figure 1.** A client with distributed server architecture adapted from [12].

another question will appear which is what the reasons for all of these efforts to build electronic archives are as long as the traditional ones still found. Despite the fact that papers and archives are increasing daily, an electronic archive does not solve the congestion problem! They even work in parallel. As soon as an obvious necessity to electronic archive has arisen, and, at the same time, the traditional archive with its large spaces, difficulties of retrieving documents, unsafe places, and its dust, are originated, then a paradox is set up! In fact, the electronic archive does not necessarily cancel the traditional one, but it made it just like the museum, where the original copies of pictures and statues are collected. In the age of speed, it cannot be spent hours or days to retrieve essential data, which may lead to losing essential businesses. The electronic archive is quickly retrieving any data requested without any need to leave the place or to know where the data are. Nowadays, in the being thereof an electronic archive, no need to store original copy, the emails became identities, and what they contain, in case of presence digital signature, all that these emails can be guaranteed. While digitalization everything, all these data become sources for the electronic archives. Many methodologies can help to transfer from traditional to electronic versions of an archive.

#### 1.4.3. Reach and Use

In all stages of building archives, there is an essential rule -linkage between archives and the storehouse of information in all its forms-. These procedures are known as indexing, which involved in all the operations regarding all partials of work. Accessing any information, whether the used system is an electronic archive or traditional, has to make a link either virtually or real connection to be able to recover this information. Regardless of information uses, and whatever they are, the critical concern is when and how to apply. It is significant to choose the appropriate time to prepare information to be worked on, in parallel, at what time this work should be involved. In other words, selecting the proper time to retrieve and use information is substantial for obtaining a good outcome in the

form of a decision. This will be served by a quick retrieving of information.

## **2. Case Study: Payroll and Personnel Information Systems in JOWFE Oil Technology (One of Libyan National Oil Corporation)**

### **2.1. Purpose**

These two computer systems in JOWFE Company have represented a shared database. New data has been added, modified, and retrieved daily. Important information has been taken from its database, and some reports have been issued using the data entered. Calculating the salaries of employees has been the most crucial operation that these systems have done. The processes include deducting the days of absence, collection of the number of times of being a delay, and the deduction of social insurance for retirement payments. At the same time, they already have an archive, including all the data. These represent one of many electronic archives in this company. Three decades ago, this operation was manually calculated. By that time, it had taken a long time to be achieved, more employees, and more efforts, in addition to the noisy environment which was easy to miss any relevant documents. The feedback came from other departments as a set of papers which create workstation filled by packages of papers.

### **2.2. Studying the Systems**

It was essential to study and analyze the structure and installation of the data and processes involved in the work. For example, knowing all that has been used like types of documents, photocopies, images, video or audio recordings, and digital data which entered as records, etc. Appropriate forms of expected data incorporated in general were designed. Applications and programs, which could be applied as input data, were developed. Data the two users for input, process, and output. On the other hand, the development of management systems regarding survey data is done.

### **2.3. How to Execute the New System**

As a database for the two systems is designed using the same methodology and database model, it is not hard to reorganize that the two databases have many similarities between both; these similarities could be used while re-designing tables are required. Each system kept the same screen design unique. Ultimately, the project needed just a system analyst and a programmer to design tables and modify and/or create some programs. The database will be the archive, but it is just for company use. Most of the output of this system are mainly used to be submitted to the authorized departments and the decision-maker. Externals may access other data. Furthermore, this system has its archive to retrieve old-dated data or that data which regarded to employees who left the company or retired. The system has kept track of any employee records. However, there is an external backup which referred to in case of emergency *i.e.* losing some data from the

server, accidentally.

## 2.4. Testing and Documenting

Inputting real data to the system is the only way to test your modifications. This is to test the new and modifying programs. If they regularly worked, without mistakes, that would be good results; otherwise, the database and programs should be revised. While passing this step, the system analyst and programmer will confirm the effectiveness of the job, then they refer the work to be documented. The system analyst will document the work finally. All the tables, the methodology used designs, programming syntaxes, and screens that appear in each stage. The documentation is vital for the future because it allows anyone else to revise and understand.

*B.V.* Another type of data that can be accessed by anyone is found on the website. Most of them are accessible unless they are confidential information. This data is installed to servers, computers with the huge amount of data that can be stored there with many facilities such as mirroring which means copying versions from this data to be recovered in case of system corruption. This project can be considered as a model for most of what has been shown as steps. In addition to being linked to the traditional archives, it used Distributed Database Systems.

## 2.5. The Novelty

Although there are large specifications for today's computer machines, especially memory, this does not prevent the use of better management of available resources. The programmer can store the databases in the clouds. If there are massive databases, they are hard to upload. This highlights the importance of splitting the single database. The programmer's task is to load into the cloud an empty data file containing only fields. There are files with tens of millions of records. It's hard to load it all at once—especially those files with pictures and maybe videos. The programmer designs a loop so that it allows for specific numbers of records from loading at the same time. The process is repeated quickly until the entire database is fully uploaded or downloaded. In each transmission, the authentication and notification of full data arrival are carried out. Fearing that all or some of what has been sent will remain, dynamic loading technology is used, a feature that is available in most operating systems. The dynamic loading brings the data from its storage locations and puts it temporarily in memory. Then after the job is done be traced to their original positions and clean the memory of these contents.

## 3. Conclusions and Recommendations

To sum up, Archive is “A place where documents and other materials of public historical interest are preserved” [14]. It is essential to organize our “memory”, the archive, to make things that are easier to be retrievable and reachable, and

the ancient nations had done it before. The archive has improved over time. Today, electronic archives have been used because they offer many facilities such as speed, safe, secure, and reliability. In their book “digital archive the new challenge”, connected between the inventing of writing and inventing of the archive.

There is a difference between naming archives as digitalization and electronic archive. Digitalization specializes in converting documents to be accepted as a part of electronic archives and DDBSs. The archive was explored of historical background and some definitions had been discussed. Three main components are asserted: documentation, indexing and archiving, distributed databases, and finally, a case study.

Perhaps some people do not be aware of the great benefits provided by the archiving systems in general, although they touch all the particulars of daily life. Once a man wakes up in the morning, enters the bathroom looking for particular things, all that finds are organized and easily accessible. Going to wearing clothes, he finds everything arranged in cabinets, shelves, and drawers. He goes out to find his car parks somewhere. Documents and things are organized, up to the train station, it is known to leave on a schedule—streets, where to find a place of work, which floor of the building. There is a desk, papers, and documents that organize the duty. Also, helped tools in performance work were found. Navigating computers, searching and retrieving information, logging on e-mail to reaching emails, responding, and archiving are essential tasks in work, etc. If the words (find, order, organize, etc.) are mentioned this means retrieving and hosting that is the archiving. Thus, the importance of archives will be seen.

Each business has to automate the work. This contributes to possible cooperation in information exchange with others, and to facilitate work organization. Furthermore, this also leads that information is transmitted between all, who have access to authority, easily. It is important to identify and restrict the powers of accessing information; therefore, each level of validity supposed to obtain its information. These controls differ between the lowest level in management and a higher level. Senior management and decision-makers are often no controls on their access to any information. Most electronic archiving systems depend on the networking systems which make the concept of e-government easier. This leads to the simplification of the services provided to the customer. In addition to that, it helps the ease of performing administrative work. Terminologies like e-management, e-government, e-marketing, e-commerce, and electronic archiving are all mostly, strictly, linked. Common factors are networks and distributed database systems. All of this is in the interest of globalization; the most critical elements are telecommunication and information exchanging across scattered data houses. This attributes to make the work more accessible in multinational companies, which are often technically complex. It reduces many of the expenses related to logistics. Yet, it is possible to be managed by the headquarters of the parent company. The United States compared to some European countries underserved these services, expensive in the prices of services and the lack of competitiveness because monopolizing in each group of states is on ei-



ther one or two service providers at the most, although it is the leading developer for most of these advantages. The primitivity of the US in this field leads the companies to monopolize these services. Thus, this imposes their consumer whatever prices and services, specifically regarding the Internet and communications. Being a continent consisting of fifty states, these services should be given priority and a higher degree of care by the decision-makers in the United States of America. The most important thing, in the country, although it is the richest in the world, there is high rates of unemployment compared to its wealthy—automated open more job opportunities. However, the archive couldn't be away from all that is mentioned.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- [1] Spencer, N. (2012) How Much Data Is Created Every Minute? Visual News. <http://www.futuretimeline.net/subject/computers-internet.htm>
- [2] Silverman, M. (2012) A Day in the Life of the Internet [Infographic]. Mashable. <http://mashable.com/2012/03/06/one-day-internet-data-traffic>
- [3] Meyer, D. (2009) Latest SD Card Format to Reach 2TB of Storage. ZDNT. <http://www.zdnet.com/latest-sd-cardformat-to-reach-2tbof-storage-3039589099>
- [4] Marsan, C. (2010) 10 Fool-Proof Predictions for the Internet in 2020. Network World. Silver Peak. <http://www.networkworld.com/article/2238913/wireless/10fool-proof-predictions-for-the-internet-in-2020.html?page=1>
- [5] Featherstone, M. (2006) Archive, Theory Culture Society (23) p. 591 Fredrick. FutureTimeline.net. <http://www.futuretimeline.net/subject/computers-internet.htm>
- [6] Cohen, D. (2004) Internet History. Computer History Museum. [http://www.computerhistory.org/internet\\_history](http://www.computerhistory.org/internet_history)
- [7] Zins, C. (2007) Conceptual Approaches for Defining Data, Information, and Knowledge. *Journal of the American Society for Information Science and Technology*, **58**, 479-493. <https://doi.org/10.1002/asi.20508>
- [8] Pearce-Moses, R. (2005) A Glossary of Archival and Records Terminology. SAA, Chicago.
- [9] Heinrich and Maurer (2000) Concept, Implementation, and Applications. *Journal of Universal Computer Science*, **6**, 1197-1202.
- [10] Harman, E. and Candela, L. (1990) Retrieving Records from a Gigabyte of Text on a Minicomputer Using Statistical Ranking. *Journal of the American Society for Information Science*, **41**, 581-589. [https://doi.org/10.1002/\(SICI\)1097-4571\(199012\)41:8<581::AID-ASI4>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-4571(199012)41:8<581::AID-ASI4>3.0.CO;2-U)
- [11] Hamill, D. (2009) Archives 101: Or, How Archives Are Organized. Northern Kentucky University. Kentucky Libraries, Summer 2009, Vol. 73, Issue 3, 22.
- [12] Goebel, V. (2011) Distributed Database Systems. Department Informatics, University of Oslo, Oslo.

- [13] Silberschatz, A., Korth, H. and Sudershan, S. (2010) Database System Concepts. Sixth Edition, McGraw-Hill, New York.
- [14] Manoff, M. (2004) Theories of the Archive from across the Discipline. *Portal: Libraries and the Academy*, 4, 9-25. <https://doi.org/10.1353/pla.2004.0015>