

Cloud Computing and Digital Archive Preservation

Fiqru Mafar^{1*}, Budhi Santoso², Nur Azizah Maulidyah³

* Lead Presenter

^{1*} UIN KHAS Jember, Indonesia, and mafarfiqru@gmail.com

² UIN Raden Fatah Palembang, Indonesia

³ UIN KHAS Jember, Indonesia

This study aims to describe the use of cloud computing in digital archive preservation. The method used in this study is a qualitative method with documentation techniques to collect the data. Results shown that conventional archive components such as cabinets, folders, and archives can be applied in the cloud computing systems. The cabinet is realized through the selection of cloud computing media. The map is realized in the form of storage folders. Meanwhile, archives are realized in the form of digital entities stored and managed in a cloud computing system.

Keywords: archive; digital; cloud computing.

1. INTRODUCTION

From time to time, the world of technology has experienced rapid development. This development is inseparable from the industrial development that accompanies it. As has been understood, the industrial world's development began with the industrial revolution.

The industrial revolution, which was initiated with the invention of the steam engine, was later known as the industrial era 1.0. As time goes by, the progress of the industrial world today has entered the industrial era 4.0. Some experts have stated that today's society has entered the era of society 5.0. The term Industry 4.0 is a term that the German government originally pioneered to introduce the computerization and digitization of manufacturing (Yahya 2018). In the 4.0 era, the keyword that is used as the driving force in digital technology (Suherman et al. 2020).

The existence of digital technology as a driving force has a tremendous impact on the industrial world. Digital technology has also had a considerable influence on society at large. People seem to be forced to adapt to the digital world in their daily lives. They are also increasingly aware of the importance of the existence of technology in various activities. If in the 1990s electronic mail (email) and the like could only be accessed using computing media such as personal computers (PCs) and laptops, nowadays people only use gadgets such as cellular phones, people are free to use technology according to their needs.

Almost all levels of society can access digital technology. The ease of obtaining a gadget as a digital device has encouraged the high use of technology in everyday life. They can easily access email and social media through the devices they have.

Ease of access also occurs to information provided by government agencies. Several government agencies have applied digital technology to the broader community as a service. The existence of available digital services is expected to provide service acceleration for those in need.

In carrying out its daily activities, a government institution turns out to produce many vital archives. Some of them are essential archives with legal value, so they require different treatment from other archives. One of the important treatments that need to be considered is maintaining the durability of the archive so that it can be accessed when needed. At this stage, the archival institution has a vital role.



The importance of archive management by archival institutions is increasingly felt in managing essential archives related to events in an area. Various events in various regions in Indonesia will produce various forms of archives that must be appropriately managed. As a result, if these archives are not appropriately managed, the historical sequence is lost in the archives. In addition, the historical proof process has become a bit hampered. On the other hand, some research activities carried out by related parties sometimes require essential archives. This requires the archive preservation process to be carried out carefully.

Preservation activities can be carried out through a series of activities to ensure the availability of archive access when needed. The development of the industrial world and technology in society has encouraged archive preservation activities based on digital technology. Through digital technology, preservation activities can be done quickly. Archives can be transferred into digital form to last longer than archives in their physical form. In addition to providing easy access, digital archives can also maintain the information contained in them. Transferring archive media is carried out to produce other forms of archives with different media without changing the contents.

Digital archive preservation activities do not only stop at transferring media into digital form. Preservation is also closely related to the availability of storage media. The form of storage that can be done is the digital storage (Mafar 2021). This form of storage has developed quite rapidly. If digital archive storage was initially done through storage media such as a set of computers, it has now developed into a form of cloud computing-based storage.

Some archive managers, such as the Bendey Historical Library at the University of Michigan (UM), have taken advantage of the cloud computing era to preserve their archives in digital form (West 2007). The same thing, of course, can be done by other archival institutions and even the general public to manage their archives. People can store and manage by uploading their archives into available storage media based on cloud computing.

Based on this background, this paper will discuss using cloud computing in the digital archive preservation process. As a follow-up, this paper will focus on applying archival concepts in a cloud-based storage form

1. THEORITICAL FRAMEWORK

Many previous researchers have carried out studies on cloud computing. Previous studies that the author found were articles written by Sontana, Rahmatullah, and Rachman. They discuss the Google Picker API in cloud computing-based archive storage in their writing. According to them, the application of the Google Picker API in electronic archive storage can facilitate the process of storing, accessing, and reducing archive management costs (Sontana, Rahmatulloh, and Rachman 2019).

Louk conducted the second study. In an article published in the journal *Teknika*, Louk examines the development of a cryptographic system to protect medical data security in cloud computing. The results show that the use of a functional re-encryption formula will give multiuser confidence in using a series of cloud computing services (Louk 2018).

The third study was produced by Yahfizham et al. In 2018, they conducted a study on cloud computing and learning management systems. The study results indicate that the cloud computing-based learning system based on Software as a Service will develop quite rapidly in the future. They predict that in the coming years, several developed countries will no longer use applications by installing them on computer devices but will prioritize the use of virtual operating systems (Yahfizham et al. 2018).

Several previous research results that have been presented show that studies on archives and cloud computing are still rarely carried out. On the other hand, cloud computing is

currently developing so rapidly. Today's society has a wide choice of cloud computing storage media.

Cloud computing is a term used to describe the combination of computers and cyberspace. It is known that the internet is a combination of several information technology infrastructures connected into a super extensive network. The word cloud describes a very complex infrastructure that is hidden behind the internet network. The word cloud also illustrates that in the era of cloud computing, users are not required to be experts in network infrastructure development because they access it with a set of computers.

In this era, users no longer need to worry about the storage capacity of a computer. Cloud computing is an era where data is stored on virtual servers on the internet. Users use computers as temporary storage media (Kho 2009). This allows each user to have their virtual server without having to depend on the capacity of the computer in front of them. In fact, with current technological developments, users do not have to use a set of computers as a medium to access the information they have on the server.

Cloud computing has made information owners more flexible. They can mix and match the content they have without being hampered by technological infrastructure problems or costs that will be incurred (Kho 2009). This flexibility is a distinct advantage for people who want to manage their information in the cloud.

There are many benefits that people get from the development of cloud computing, including the following (Setiawan 2011).

- a. The budgeting process can be carried out more efficiently
- b. Managers can easily do development
- c. Management and other operational processes can be done easily
- d. Provide easy access to collaboration processes
- e. Operational costs can be reduced

2. RESEARCH METHODS

This study uses a qualitative approach with data collection through documentation techniques. The analysis carried out is descriptive analysis. Through this analysis, the author presents the concepts of conventional archive management and compares them in the form of application to cloud computing systems.

3. FINDINGS

Mark Matienzo, an archivist in Washington DC, argues that many archivists either do not know or do not care about technology (West 2007). On the other hand, information technology has affected the community's activities so that it requires fast access to the information needed. This shows that there is a mismatch between the demands of society and the current state of the archives.

The above statement is not entirely accurate, in the author's opinion. Several archive managers have been able to keep up with technological developments. They are starting to realize that if they stick to conventional archival forms in their original form (generally in paper form), the original form will gradually be damaged. Therefore, some of them began to apply technology in managing their archives. However, for those who are not careful in understanding the concept of applying technology, archives in their original form are often destroyed after producing archives in digital form. The misunderstanding made the archive in its original form disappear. This principle is, of course, different from the principle of digital libraries.

In principle, digitizing activities in a digital library does not mean eliminating the physical form of a document but as a part that complements the library in its physical form itself. Digitization is ideally carried out as one of the steps to maintain its information. The

preservation of the original form is essential so that the digital form can be used to complement the conventional form.

Apart from the word 'destruction' above, in order to be able to preserve technology-based archives in the cloud computing era, an archival organization first needs to determine the design of their digital archive preservation to ensure the sustainability of their preservation (Su-Shing 2007). In this case, the archive organization has three important functions:

1. Monitor the information needs of users and interact with them and information producers to determine changes in the form of services and technology used.
2. It is responsible for the sustainability of the technology used so that the continuity of access to digital archives can be maintained.
3. Run the standard and the migration process that has been carried out.

The intended standards include format standards, metadata, and documentation. For this reason, archive managers need to understand the components related to archive preservation. Archive preservation using cloud computing technology is not much different from the conventional archive preservation concept. The main components such as cabinets, folders, and archives are also applied. Furthermore, an explanation of the three components is as follows.

1. Cabinet

In conventional preservation, cabinets are realized by the presence of shelves or filing cabinets. In the era of cloud computing, the cabinet concept is realized in a virtual form in the form of a server in cyberspace. Currently, many companies provide virtual servers with various capacities. Call it Onedrive, DocStoc, Scribd, Google Drive, and others.

The selection of this virtual server is essential because it is related to how many archives will be stored. This is because each server has a different storage capacity. As an illustration, Google Drive provides a free capacity of approximately 15 GB and an additional capacity of up to more than 1 TB for those willing to use it for a fee. The picture shows that the selection of a server that functions as a 'cabinet' will indirectly affect the costs incurred.

In addition to costs, this election is also related to access rights to the archives stored. Archives need to determine whether the archive is intended for the public or limited to certain circles. Determination of a server with adequate privacy facilities is essential in this case. Some virtual servers are servers based on social cataloging. It causes each user to access the archive even with limited rights, such as being able to only read without being able to download the archive in question. Therefore, archival institutions must choose a 'cabinet' related to this privacy issue.

2. Folders

In conventional archive storage, folders are used to store and group files in cabinets. Folders can be realized in the form of a storage folder on a virtual server. Each archive that is stored will be organized with a different storage system. Some use a file number system as the order of storage, while others are based on the contents of the archive itself.

Generally, every virtual server provider has provided a search menu to simplify the process of retrieval of stored archives. However, it would be nice if the archive manager still divides the archive into several folders according to the system that has been set. It is done to show that each archive stored in the folder has a relationship with one another. In general, information preservation is divided into three types, namely:

Provenance. This type divides archives based on the history and origin of the archive.

Context. This type divides archives based on the relationship of information stored between archives with one another.

References. This type divides files based on the unique number that each file has.

Fixity. This type divides the archive based on the accessibility and level of change of the information stored.

3. Archive

Suppose, in conventional archive management, the storage is carried out in its original form in a cloud computing system. In that case, the archive needs to be transferred first before being managed and stored on a cloud computing system. Digital archives that are stored can be in image formats, office documents (word, excel, PowerPoint), or pdf format. Determining the archive format to be stored is an important thing that needs to be considered.

The format of the stored documents will indirectly affect the grouping of archives. As mentioned above, one type of archive grouping is fixity. In this case, some document formats do not allow users to make changes, such as image and pdf formats. In addition, the archive format will affect the server selection. Some formats, such as image and pdf formats, require a larger capacity than office document formats, so they will also require a larger storage capacity.

For the components that have been described to be appropriately implemented, archivist competencies are needed that are appropriate and in line with existing technological developments. Archivists and librarians have common challenges in terms of information preservation (West 2007). As a professional engaged in information management, archivists are required to:

- a. Understand and follow the developments needed to protect the integrity of archives in a new way
- b. Assure the community to obtain valid documents
- c. Securing and providing a documentary heritage for the society of today and tomorrow (Hadiwardoyo dalam Burhanudin DR. 2009).

To meet the above demands, an archivist needs to be equipped with specific competencies. The competence in question is related to educational issues and relates to other skills regarding archives. These competencies are as follows.

- Understanding of the organic context in the administrative structure and national accountability to future generations;
- Reviewing an administrative system and formulating a system for managing information on archives to ensure administrative efficiency and ensure security and preservation of the national cultural heritage in an accurate, targeted, and timely manner;
- Having the skills to manage the deposit of information and scientific insight that allows him to assess the culture that needs to be preserved (Jabatan Fungsional 2006).

In addition to the above competencies, archive managers also need to equip themselves with knowledge about the legality of archives that have been transferred. Therefore, archive managers should understand the principles of transferring media of a document as stated in Government Regulation Number 11 of 2008 concerning Procedures for Transferring Company Documents into Microfilm or Other Media and its Legalization.

4. CONCLUSION

Cloud computing has demanded archive managers to be able to apply information technology in the archive management process. Matters related to the digitization of archives are essential issues that need great attention among archivists. In addition, increasing competence, such as the aspect of legalizing the transfer of archival media, is another thing that should not escape the attention of archive managers.

REFERENCES

- Burhanudin DR. 2009. *Tantangan Arsiparis Dalam Menghadapi Era Paperless Office: Ke-Belumberhasilan Komputerisasi Di Kalangan Birokrasi*. Yogyakarta. <https://adoc.pub/tantangan-arsiparis-dalam-menghadapi-era-paperless-office-ke.html>. “Jabatan Fungsional.” 2006.
- Kho, Nancy Davis. 2009. “Content in the Cloud.” *EContent* 32(2): 26–30.
- Louk, Maya Hilda Lestari. 2018. “Sistem Kriptografi Di Komputasi Awan Untuk Kebutuhan Data Medis.” *Teknika* 7(1): 16–20.
- Mafar, Fiqru. 2021. “Pengelolaan Dokumen Kepangkatan Dosen Berbasis Komputasi Awan Pada Perguruan Tinggi Keagamaan Islam Negeri.” *JIEMAN: Journal of Islamic Educational Management* 3(2): 163–84.
- Setiawan, Deris. 2011. “Teknologi Cross Platform, Telecomputers & One Stop Solutions: Cloud Computing.”
- Sontana, Indra, Alam Rahmatulloh, and Andi Nur Rachman. 2019. “Application Programming Interface Google Picker Sebagai Penyimpanan Data Sistem Informasi Arsip Berbasis Cloud.” *JURNAL TEKNOLOGI DAN SISTEM INFORMASI* 5(1): 25–32.
- Su-Shing, Chen. 2007. “Digital Preservation: Organizational Commitment, Archival Stability, and Technological Continuity.” *Journal of Organizational Computing & Electronic Commerce* 17(3): 205–15.
- Suherman, Musnaini, Hadion Wijoyo, and Irjus Indrawan. 2020. *Industry 4.0 vs Society 5.0*. Banyumas: Pena Persada.
- West, Jessamyn. 2007. “Saving Digital History.” *Library Journal* 132: 1–6.
- Yahfizham et al. 2018. “Sistem Manajemen Pembelajaran Awan Berbasis Perangkat Lunak Sebagai Suatu Layanan Analisis Deskriptif.” *Techno.com* 17(3): 252–69.
- Yahya, Muhammad. 2018. *Era Industri 4.0: Tantangan Dan Peluang Perkembangan Pendidikan Kejuruan Di Indonesia*. Makassar.