v. 4, n. 28, 2024

International Journal of Human Sciences Research

# PRESERVATION OF DIGITAL DOCUMENTS AS A NATIONAL STRATEGY TO ENSURE TRANSPARENCY AND MEMORY

***Gabriela Castillo Solano***
Professor of Archival Science,
Universidad de Costa Rica, Costa Rica

***Raquel Umaña Alpízar***
Professor of Archival Science,
Universidad de Costa Rica, Costa Rica

**Keywords:** digital documents, digital preservation, preservation policy, digital preservation strategies.

The fourth industrial revolution has reached us, bringing with it innovations, challenges and even new paradigms, where technological advances play a major role in terms of access to information and shortening distances and times. To this end, organizations are betting on the use of technology and with it, electronic documents to maximize their effectiveness and at the same time be more efficient.

In this scenario, the immaterial is becoming the only evidence of our knowledge and our footprint as a society, which makes it urgent to implement techniques and methods to ensure the preservation of organizational and collective memory, but also to ensure evidence of legality and transparency in the management of organizations.

ISO 13008:2012 defines preservation as: "The processes and operations performed to ensure the intellectual and technical permanence of authentic documents over time" (p. 8); therefore, digital preservation must ensure access to digital information in the long term, maintaining the value of its significant properties, in a constantly evolving environment; reason why the challenges must consider technological obsolescence.

In addition to the above, the number of documents in digital format is growing in the absence of regulations, making it difficult to preserve essential documents for administration and decision making, it is urgent to have a national strategy to ensure the authenticity, durability and access to digital documents; one of the most pressing challenges of the time.

The preservation of evidence and history contained in digital documents strengthens transparency, administrative management in general, ensures citizens to enjoy and maintain their rights. Developing a national policy and digital preservation strategies that ensure access to documents, while guaranteeing their authenticity and integrity over time, is a duty to the homeland, citizens and new generations.

## OBJECTIVES

### GENERAL OBJECTIVE

Encourage the implementation of digital preservation through the development of national policies and strategies that ensure access to documents and guarantee evidence of legality and transparency in institutional management and organizational and collective memory.

### SPECIFIC OBJECTIVES

1. Identify the minimum elements required for the preservation of information assets in the digital environment that are the product of the functions of the organizations.

2. To provide guidance for the development of digital preservation policies and strategies for the management and preservation of authentic, integral and reliable digital documents over time.

## METHODOLOGY

It is based on the methodology used for the Final Applied Research Work of the Professional Master's Degree in University Administration of the University of Costa Rica, called "*Preservation Model of Digital Documents in the University Administration. Case study: Universidad Nacional*"; which was developed with the objective of proposing an applicable model that guarantees the administration and preservation of digital documents produced in organizations, as well

as their authenticity, durability and access to the information contained in these digital documents over time.

The research is qualitative, so an analysis of international requirements models was carried out to determine the best practices applicable to the context in which it was developed. It is also an applied research that indicates the strategies for the preservation of digital documents, in order to face the challenges of technological obsolescence and long-term preservation of digital information assets.

The bibliographic and documentary review is the technique by means of which information is collected through books, articles, theses, guides, brochures, among others; therefore, for the development of the research, primary and secondary bibliographic and documentary sources were used, such as: international standards and models and applicable legislation, scientific articles on digital archives, full text databases, among others.

Regarding the treatment of digital preservation, we take as a reference the ISO 14721 Standard: Spatial data and information transfer systems. Open Archival Information System (OAIS); as well as the methodology used by Professor Jordi Serra Serra, Master of Document Management at the University of Barcelona (MGDIE), in the advice and training provided to the University of Costa Rica[1] , and the experience of the ARCA project[2] .

## ELEMENTS REQUIRED FOR DIGITAL PRESERVATION

### A GUIDE TO DEVELOPING DIGITAL PRESERVATION POLICIES AND STRATEGIES

#### 1. Digital Preservation Policy[3]

Preservation is not only about storage, but also strengthens organizational management; for this reason, a digital preservation policy should be established, which should be aligned and consistent with existing national and organizational policies.

The first element of this guide is the Digital Document Preservation Policy, which has as its purpose the preservation and access to digital documents resulting from the substantive functions of the organization, ensuring the integrity, authenticity and reliability of the information over time.

In this sense, preservation policies should be developed with a view to implementing the following aspects:

a. Promote the proper management of documents to ensure reliable information for decision making, efficiency, accountability, risk management and preservation of organizational memory.

b. Preserve the records that are produced in the organizations in digital format and that must be preserved over time, ensuring the authenticity, integrity and access to the information assets; as well as the clear designation of those responsible for the preservation of the records.

c. Establish the strategic, legal, technical, technological and functional requirements to ensure the continuity, preservation and integrity of the Digital Archive.

1. Activities organized by the Universidad de Costa Rica in 2012, 2013 and 2017.

2. Digital Preservation Digital Repository used in the National Archive of Costa Rica and developed by the company *Business Integrator Systems* (www.bis.co.cr), from which advice was received during the research process.

3. The structure or format of digital preservation policies may vary depending on the national and organizational context, as well as existing regulatory provisions .

d. Develop a metadata schema to ensure the preservation of contextual information and meaningful properties that enable the continued accessibility, usability, significance and evidential value of the information contained in digital objects, over time and across technological change.

e. To implement the necessary preservation strategies and the provision of new services for the preservation of heritage in digital format, through the use of information and communication technologies; to have the means and resources required to facilitate, preserve and guarantee access to information over time.

**2. Preservation Strategies[4]**

Preservation strategies, the second element of this guide, are intended to establish a proposal of requirements for the development and implementation of a Digital Archive; this proposal has a strategic and operational scope at the organizational level.

The proposed strategies are based on the Open Information Systems Reference Model, OAIS Model (ISO 14721), as a conceptual verification within the principles of reliable digital repositories, defined by international standards and best practices.

***a. Administrative aspects***

• Digital Archive: is responsible for preserving and guaranteeing access, authenticity and integrity of documents in digital format. The Digital Archive (AD) shall be part of the entity's archive system and shall be in charge of the Archive Unit[5], which shall be responsible for managing the economic, technological and human resources for its proper operation.

In addition, the Technology Unit[6] must assume responsibility for the management and assurance of the technological infrastructure and platform, in compliance with the operational requirements of the Archive Unit, the improvement proposals and the updating of the technological components involved.

• Institutional Digital Preservation Commission (CIPD): is the governing body in charge of coordinating, issuing and updating institutional policies, strategies and guidelines regarding digital preservation and the development of the Digital Archive. Among the main functions of the CIPD are: to issue the guidelines and directives to be applied in the area of digital preservation; to establish the necessary strategies for the proper functioning of the AD; to provide guidance for the use of budgetary, human, equipment and service resources; and to intervene in the resolution of conflicts involving the producers, users and managers of the Digital Archive.

4. The design of digital preservation strategies is based on the PLATTER model, DPE Repository Planning Checklist and Guidance. DPED 3.2 April 2008; on the OAIS Model, Standard UNE-ISO 14721: 2015 and on the consultancy carried out by Jordi Serra Serra at the Universidad de Costa Rica in 2013.

5. This is understood as the body of the organization responsible for the management of documents and archives; whether it is the Central Archive, Institutional Archive, Archive Department, Archive Office, Records Management Unit, Records Management Directorate, etc.

6. Organizational entity responsible for information and communication technologies; depending on the name: management, unit, office, department, etc.

• Human Resources: the organization must make available to the DP the human resources necessary for its optimum operation, and ensure that the personnel in charge have the appropriate academic training and are constantly updated to perform their duties in line with technological evolution; to this end, an annual training program must be in place.

The employee responsible for the AD will be known as Senior Information Risk Owner (hereinafter SIRO), and will be the official in charge of managing the AD and applying the preservation strategies, and must have the ability to direct, lead, create policies and procedures that allow the preservation and security of the information.

• Budget: the CIPD must define on an annual basis the necessary resources for the implementation of the Digital Preservation strategies, as well as for the maintenance, updating and growth of the platform and the infrastructure that supports it, for storage, maintenance and updating of the AD solution and for the training program.

• Requirements for transfers between systems: the AD repository must be able to receive documents and information from the different systems existing in the organization, for which there must be approval from the Archive Unit; these transfers must be carried out in accordance with the provisions of a Transfer Protocol, previously agreed between the parties.

In addition, the Technology Unit will be in charge of supervising and authorizing the services, applications and integrations for transfers and/or communication between systems, according to the needs defined by the CIPD, guaranteeing the interoperability, authenticity and integrity of electronic documents during their transfer, migration and storage.

***b. Data Plan***

The definition of the characteristics of the contextual information and that contained in the documents to be preserved is the responsibility of the Archive Unit, and must be executed in accordance with the guidelines of the Institutional Committee for the Selection and Disposal of Documents (CISED)[7] and the CIPD. The exchange and storage of information should be managed in accordance with the following types of information packages/containers, as established in the OAIS model, namely:

• **Submission** *Information Package* (***SIP***): used for the transfer, the *Submission Information Package* (*SIP*) is the content and description information of the documents (provenance, context, references, access rights, authenticity, integrity, etc.) that is delivered by the organizational instances to the AD, for use in the construction of the *Archival Information Package* or *AIP*.

*SIP* processing can be carried out in two ways: the *pull* method, which implies that the repository has a "broker" agent that "collects" the material; or the *push* method, whereby the *SIP* is delivered to the repository for processing by the original producing system or organization. Regardless of the input method, it is essential to verify the absence of *malware* and the integrity of the transferred document, data validation, metadata generation, extraction of significant properties and quality assurance of the package.

7. According to Article 33 of Law No. 7202 of the National Archives System of Costa Rica, CISED is the competent body to evaluate and determine the administrative and legal validity of the organization's documents.

- **Archival** *Information* **Package (*AIP*)**: The *Archival Information* Package (*AIP*) is the package in which the preservation processes are executed and which contains the content, descriptive, relationship, and associated preservation information, including the modifications necessary to be stored in the AD, as well as the file itself self-contained in the package structure.

- *Dissemination Information Package* (***DIP***): generated for consumption as a product of a query, the *Dissemination Information Package* (*DIP*) is the information package created by the AD and derived from the *AIP*, to distribute the digital content that is consulted by users or by systems that have access to the Digital Archive.

The *DIP* will take the form of a single file, which will be considered an authentic copy of the documents stored in the AD; in addition, the container index (electronically signed), with the metadata of the document unit to which each file corresponds and with a cryptographic summary; or by means of a hyperlink to the relative location of each file, or it may be contained within the package in the form of a consultation copy.

The files may be provided in a format other than the one kept in the *AIP* or may be reproduced at the time of consultation, without the document having to be kept in the format for consultation; it is necessary to emphasize that an access evaluation must be carried out beforehand to establish whether a specific user may consult a specific document.

***c. Income Plan***

This plan establishes the services and functions to receive the *SIP*, is the process of transferring those digital documents that, as part of a preservation strategy, become part of the Digital Archive.

The entry of documents to the AD can be done by means of automatic or assisted transfers, but always through the defined channels, and can be done on a regular basis within an agreed period of time, as determined by the Transfer Protocol, for which the detail and method (*push* or *pull*) must be specified, which can be done by Document Series[8] or by Institutional System, in order to establish the structure, content and form.

For this task, a document must be formalized that sets forth the conditions and requirements for the transfer of the documents to the AD, which will be the *Transfer Protocol*, which is prepared jointly by the agency that produces the documents and the Archiving Unit.

The transfer protocol must contain at least the following information: scope, transfer methods, *SIP* structure, supported file formats, metadata schema, control mechanisms and validations, rights to perform preservation actions on the documents, preservation period, significant collectible properties, use and access rights, effective date of the protocol.

***d. Conservation Plan***

The preservation plan provides the services and control functions of the Digital Archive environment and includes the technical strategies with the objective of maintaining, over time, the characteristics of integrity, authenticity and access to the information in custody. Therefore, the following aspects are contemplated in the preservation plan:

- Storage: it is required to provide the services and platform for the secure storage, maintenance and retrieval of *AIPs*.

8. The series that materialize in information systems will be considered recapitulative.

• Document retention conditions: documents retained by the AD must have a current technological representation, encoded in the updated format established by the SIRO, and must maintain the structural relationships inherent to such documents.

• Risk monitoring and detection: changes in service requirements and technological changes (threats, media, formats, systems and computer platforms) should be monitored to identify emerging technological trends that could cause risk or obsolescence in the technological environment of the Digital Archive and potentially impede access and preservation over time of documents and information held by the AD; this periodic review function is the responsibility of the SIRO.

• Migration process: migrations may involve changes to content data objects and/or representation information to new formats, and may involve the development of new *AIPs*, test plans, implementation and reviews for automatic execution of the transformations, in order to reduce the likelihood of information loss.
This process is usually complex and costly, so it should be performed only in justified cases, either because of format and/or media deterioration, because new end-user requirements have been identified and/or for better cost-effectiveness, due to the constant evolution of technology.

***e. Access Plan***

This plan establishes the services to meet the needs of end users. Access to documents must be managed using automated processes; to this end, it is necessary to determine the existence, description, location and availability of the information stored in the AD. Therefore, it is up to the Archive Units to analyze the use of the documents and identify the corresponding use and access permissions.

*Permission of use* refers to the ability to establish and authorize a configuration of permissions and roles available to a user to perform the various actions provided by a system; it must be performed in accordance with the best practices set out in ISO/IEC 27001 Information Security Management Systems. The *access permission* refers to the security component of the Information and Query Package.

A set of rules must be developed to identify access rights, as well as the system of permissions and restrictions applicable to documents, metadata, files and, in general, to all levels of grouping, always in compliance with current institutional and national regulations. The AD must have the capacity to provide information to internal and external users through a query interface, or through interoperability services to be consumed by other applications or systems.

***f. Technology Plan***

• Infrastructure and Technology Platform: The AD infrastructure should meet storage, processing, security and communications needs and be scalable and redundant, using state-of-the-art but reliable *software* and equipment.

• Security: the AD infrastructure must guarantee the authenticity, integrity and availability of the documents, information and data under custody, complying with the regulations and procedures established in the organization for this purpose. Likewise, the Digital Archive must have a program of authenticity and integrity verifications and tests of the recovery and service continuity processes; these reviews and tests will generate result reports, which are used as input to manage the protection against computer

attacks and for the configuration of the recovery processes in case of disasters.

- Backup copies: incremental backup copies should be kept on a non-alterable medium, performed on a daily basis (the incremental copies) and weekly for the base copy. Additionally, there should be a mirror storage at a remote site where the *AIP* packages are stored, using a different technology from the main site and including additional security mechanisms to protect the information.

- Audit trails: there must be a historical record that allows identifying and tracking all events or actions performed in the AD, to detect unauthorized changes, authenticity problems, records of document conversion operations and migrations; as well as proof of document destruction.

***g. Continuity Plan***

In order to guarantee the availability and uninterrupted operation of the Digital Archive, a "Continuity Plan" must be prepared and updated by the Technology Unit, with the purpose of being prepared in case of service interruptions caused by uncontrolled internal or external factors (natural disasters, events caused by third parties, etc.) and in order to restore services as soon as possible, to the greatest extent possible.

As part of the AD strategy, it is stated that, due to the importance of this service, the continuity plan must contemplate certain components in reference to the following aspects of the operation of the Digital Archive: service continuity, high availability, financial continuity, knowledge continuity and continuity in case of disaster; which must be considered vital for the continuity of the organization's business.

## CONCLUSIONS

With the fourth industrial revolution, it is necessary to implement best practices, standards, techniques and methods to ensure evidence of legality and transparency in the management of organizations, as well as to ensure the preservation of organizational and collective memory. Therefore, the development of digital preservation policies and strategies that ensure access to documents, while guaranteeing their authenticity and integrity over time, is a duty to the homeland, citizens and new generations.

Digital preservation represents one of the urgent challenges facing the information society, mainly due to technological obsolescence, so organizations must have a preservation policy and strategies that allow them to take advantage of information and communication technologies, but in a responsible manner, implementing a repository specialized in digital preservation and a robust technological infrastructure to manage and administer information assets in a digital environment and ensure their access over time.

Therefore, it is intended to encourage the implementation of digital preservation through the development of national policies and strategies that ensure access to documents and guarantee the evidence and memory of nations; by proposing a guide to guide the development of policies and strategies for the management and preservation of authentic, complete and reliable digital documents over time.

## REFERENCES

Asociación Española de Normalización y Certificación. (2013). *Norma UNE-ISO 13008:2013 Proceso de migración y conversión de documentos electrónicos. Información y documentación*. España: AENOR.

Asociación Española de Normalización y Certificación. (2015). Norma UNE-ISO 14721:2015 Sistemas de transferencia de datos e información espaciales. Sistema abierto de información de archivo (OAIS) Modelo de referencia. España: AENOR.

Castillo, M.G. y Umaña, R. (2018). *Modelo de Preservación de Documentos Digitales en la Administración Universitaria. Estudio de caso: Universidad Nacional.* (Trabajo final de investigación aplicada de Maestría Profesional en Administración Universitaria). Universidad de Costa Rica, Sede Rodrigo Facio.

Chaves, L. (2010). El paradigma cualitativo en la investigación educativa: una aproximación teórica. En Chaves, L., Díaz, M., García, J., Rojas, G., y Solís, N. (Ed.), *Investigación-Acción colaborativa: Un encuentro con el quehacer cotidiano del centro educativo para su transformación* (pp. 1-34). San José: INIE.

Ley N.° 7202 del Sistema Nacional de Archivos. La Gaceta: Diario Oficial de Costa Rica, San José, Costa Rica, 27 de noviembre de 1990.

Serra, J. (2014). Desarrollo de una Política de preservación digital como oportunidad de negocio para consultores y archiveros. Girona: Arxius i Indústries Culturals.

Serra, J. (2017). Taller de Preservación Digital para la Universidad de Costa Rica [Presentación]. San José.