

ESTUDOS DE CUSTOS PARA PRESERVAÇÃO DIGITAL E REPOSITÓRIO DIGITAL CONFIÁVEL: O CASO DO ARQUIVO NACIONAL DO BRASIL

COST STUDIES FOR DIGITAL PRESERVATION AND TRUSTED DIGITAL REPOSITORY: THE CASE OF THE BRAZILIAN NATIONAL ARCHIVE



iD

Maximiliano Martins de Faria

Arquivo Nacional do Brasil

E-mail: maxfaria@an.gov.br

Rio de Janeiro – Rio de Janeiro – Brasil

iD

Tiago Cesar da Silva

Arquivo Nacional do Brasil

E-mail: tiagocesar@an.gov.br

Rio de Janeiro – Rio de Janeiro – Brasil

**Resumo**

Introdução: O Arquivo Nacional do Brasil está se dedicando à modernização de seu Repositório Arquivístico Digital Confiável (RDC-Arq) para fortalecer sua presença online diante dos desafios da era digital. Este esforço reflete a complexidade do ambiente digital, exigindo investimentos consideráveis. **Objetivo:** O objetivo principal deste estudo é apresentar um modelo de precificação para o ambiente de preservação digital, visando aumentar a compreensão dos gastos envolvidos e o custo por objeto digital preservado, contribuindo para uma maior transparência pública e facilitando futuras decisões relacionadas à preservação. **Metodologia:** A metodologia empregada combina ação prática, investigação e pesquisa de experiências em instituições semelhantes, aplicadas ao contexto do Arquivo Nacional. **Resultados:** Os resultados incluem a construção de um modelo de RD através de um mapa conceitual, identificando funções e relacionando-as com controles de TIC conhecidos, como a ISO 27k, para estimar o custo de manutenção do ambiente de preservação. **Conclusão:** Este estudo busca compreender modelos de cálculo de custo de ambientes de preservação e sua aplicação específica ao Arquivo Nacional.

Palavras-chave: Repositório digital. Custo de repositório digital. Preservação digital. RDC-Arq.

Abstract

Introduction: The National Archives of Brazil is dedicated to modernizing its Trusted Digital Archival Repository (RDC-Arq) to strengthen its online presence in the face of digital era challenges. This effort reflects the complexity of the digital environment, requiring significant investments. **Objective:** The main objective of this study is to present a pricing model for the digital preservation environment, aiming to increase understanding of the involved expenses and the cost per preserved digital object, contributing to greater public transparency and facilitating future preservation-related decisions. **Methodology:** The employed methodology combines practical action, investigation, and research of experiences in similar institutions, applied to the context of the National Archives. **Results:** The results include the construction of an RD model through a conceptual map, identifying functions and relating them to known ICT controls, such as ISO 27k, to estimate the maintenance cost of the preservation environment. **Conclusion:** This study seeks to understand models for calculating preservation environment costs and their specific application to the National Archives.

Keywords: Digital repository. Costs of a digital repository. Digital preservation. RDC-Arq.

LICENÇA DE USO

Os autores cedem à [Revista Brasileira de Preservação Digital](#) os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a Licença Creative Commons Attribution (CC BY) 4.0 International. Esta licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHERS

Universidade Estadual de Campinas – Sistema de Bibliotecas / Instituto Brasileiro de Informação em Ciência e Tecnologia – Rede Brasileira de Serviços de Preservação Digital – Cariniana. As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Gildenir Carolino Santos, Miguel Angel Márdero Arellano.

CREDIT

RECONHECIMENTOS: Não aplicável.

FINANCIAMENTO: Não aplicável.

CONFLITOS DE INTERESSE: Os autores certificam que não têm interesse comercial ou associativo que represente um conflito de interesses em relação ao manuscrito.

APROVAÇÃO ÉTICA: Não aplicável.

DISPONIBILIDADE DE DADOS E MATERIAL: Não aplicável.

CONTRIBUIÇÕES DOS AUTORES: Conceituação, Curadoria de Dados, Análise Formal, Investigação, Metodologia, Administração de Projetos, Supervisão, Validação, Visualização, Redação – rascunho original, Redação: FARIA, M.; Análise Formal, Investigação, Metodologia, Validação, Visualização: SILVA, T.C.; Redação - revisão & edição: LACERDA, T.

CRÉDITO DA IMAGEM: Extraída do site do [Arquivo Nacional](#).

Submetido em: 23/11/2023 – **Aceito em:** 02/05/2024 – **Publicado em:** 28/06/2023

1 Introdução

Nas últimas décadas, têm sido observados crescentes investimentos na transformação digital do patrimônio cultural, com ênfase na digitalização de livros e manuscritos históricos, visando à sua preservação ao longo do tempo (Parziale *et al.*, 2020). O Arquivo Nacional do Brasil (AN) desempenha um papel fundamental nesse contexto, com a missão de recolher, custodiar e disponibilizar documentos que registram diversas atividades do Poder Executivo Federal, além de acervos de instituições privadas declarados de interesse público (BRASIL, 1991, p. 159).

Nos últimos anos, o AN tem enfrentado desafios significativos relacionados à gestão de seu acervo físico e à expansão do acervo digital. Investimentos substanciais têm sido direcionados para a preservação de objetos digitais, uma vez que sua proteção é vital para a custódia da memória cultural do país. Segundo Levy (LÉVY, P., 2010), a informação digital está transformando o ciberespaço no principal canal de comunicação e suporte de memória da humanidade.

A preservação digital se tornou uma disciplina essencial nas instituições arquivísticas, compreendendo um conjunto de atividades voltadas para garantir o acesso contínuo, e a longo prazo, à informação (Ferreira, 2006). Essa abordagem é fundamental para assegurar que os documentos digitais permaneçam acessíveis e interpretáveis ao longo do tempo (Ditadi *et al.*, 2015), exigindo o desenvolvimento contínuo de técnicas robustas de preservação digital.

A implementação da preservação digital implica em mudanças gerenciais e técnicas significativas, envolvendo uma infraestrutura tecnológica sofisticada, recursos humanos qualificados e a gestão eficaz de objetos digitais (Abrams *et al.*, 2012). Nesse contexto, os repositórios digitais (RD) desempenham um papel fundamental, fornecendo uma solução informatizada para o gerenciamento de materiais digitais ao longo do tempo (Ditadi *et al.*, 2015).

O Arquivo Nacional tem dedicado esforços para aprimorar sua capacidade de preservação digital, culminando na implementação de seu repositório digital em 2019. O modelo adotado busca alinhar-se com padrões reconhecidos, como o modelo *Open Archival Information System* (OAIS), garantindo a integridade e o acesso aos acervos digitais da instituição (Santos; Flores, 2019). Ao repositório digital compete armazenar, garantir a integridade e possibilitar acesso aos acervos digitais de uma instituição. É importante que as casas arquivísticas mantenham seus repositórios digitais alinhados às boas práticas, aos princípios da proveniência, da manutenção à custódia e autenticidade como, também, estar ajustado às funções arquivísticas (produção, aquisição, classificação, avaliação, descrição, conservação e difusão) (Santos; Flores, 2019).

Neste sentido, compreender os custos da preservação digital é crucial para o planejamento e a gestão eficaz dos recursos institucionais, embora precificar esse ambiente seja desafiador devido à diversidade de atividades envolvidas e à falta de modelos consolidados (Beagrie, 2017). Nas últimas décadas, tem havido um debate

significativo sobre como calcular esses custos, mas ainda existem lacunas nessa área. Apesar da proposição de vários modelos de precificação, a natureza heterogênea do campo torna a tarefa de calcular esses custos uma empreitada contínua (Boté; Fernandez-Feijoo; Ruiz, 2013).

Conforme apontado por Boté (Boté; Fernandez-Feijoo; Ruiz, 2013), é fundamental para qualquer organização saber o valor econômico de seu ambiente de preservação. A sustentabilidade econômica da preservação digital é um fator chave a ser considerado, juntamente com outras atividades de custo da organização. Assim, este estudo analisou diversos estudos na literatura especializada e explorou experiências de instituições arquivísticas e bibliotecas que mantêm serviços de preservação digital, aplicando esses conhecimentos ao caso do Arquivo Nacional (AN).

A partir das informações acima, nosso trabalho propõem um modelo de precificação para um ambiente de preservação digital, com foco no Repositório Digital da instituição.

O propósito é oferecer uma compreensão mais aprofundada dos gastos relacionados ao RD e ao custo por objeto digital preservado, visando promover maior transparência pública e facilitar as decisões futuras relacionadas ao ambiente de preservação.

Como resultado, foi desenvolvido um mapa conceitual, do RD do Arquivo Nacional, representando as interligações entre as perspectivas de Pessoas, Gestão Organizacional e Tecnológica. A partir desse mapa, foi possível identificar as funções contidas e relacioná-las aos controles de modelos conhecidos de Tecnologia da Informação e Comunicação (TIC), como a ISO 27k. Com base nesses controles, foi possível precificar os ativos de TIC, incluindo pessoal especializado, sistemas de apoio e contratos de serviços. O estudo estimou um custo de manutenção do ambiente de preservação do Arquivo Nacional em torno de R\$ 2.300.000,00 por ano, além de um custo de R\$ 0,013 por objeto digital preservado de 12 MB em média.

Apresentamos, então, o estudo a partir das seguintes seções: uma breve revisão da literatura, buscando a experiência de outras instituições no cálculo de custo de preservação digital; apresentação da metodologia do estudo, onde temos como objetivo de expor o método seguido para se alcançar os resultados; e, por fim, os resultados do modelo conceitual usado para se encontrar os custos do ambiente de preservação do AN.

2 Revisão e Literatura

A análise dos estudos recentes revela um aumento significativo nos investimentos para a transformação digital do patrimônio cultural. Bibliotecas, arquivos nacionais e outras instituições dedicadas à preservação da memória, têm digitalizado uma grande quantidade de documentos para garantir sua preservação ao longo do tempo e facilitar o acesso público (Parziale; Capriolo; Marcelli, 2020).

Observa-se um esforço contínuo dessas instituições na preservação de seus acervos físicos, ao mesmo tempo em que investem cada vez mais na preservação de acervos digitais, seja através de estudos e técnicas específicas, seja na implementação de infraestrutura tecnológica adequada (Lévy, 2010).

Os estudos relacionados à preservação digital tornam-se cada vez mais relevantes para instituições que buscam garantir o acesso contínuo aos seus documentos digitais (Ferreira, 2006). A preservação digital compreende um conjunto de atividades destinadas a assegurar o acesso e interpretação da informação pelo tempo necessário (Ditadi *et al.*, 2015), exigindo o desenvolvimento de técnicas para garantir a integridade e recuperação dos documentos digitais em qualquer momento.

A implementação de uma política de preservação digital envolve escolhas gerenciais e técnicas complexas, incluindo a implementação de uma infraestrutura sofisticada e o investimento em recursos humanos especializados (Abrams *et al.*, 2012). Um dos elementos centrais desse processo é a implementação de um repositório digital, que armazena e gerencia os ativos digitais da instituição (Ditadi *et al.*, 2015).

Em 2017, o Arquivo Nacional iniciou um Grupo de Trabalho para estudar e aprimorar sua preservação digital, resultando na implantação de seu repositório digital em 2019. O repositório digital do AN busca alinhar-se com modelos reconhecidos, como o OAIS (Santos; Flores, 2019), mantendo-se em conformidade com boas práticas arquivísticas e princípios de preservação.

A análise dos trabalhos existentes revela que muitos deles propõem modelos de custos internos para a preservação digital, alguns dos quais podem ser adaptados para outras instituições. Este estudo visa a desenvolver um modelo simples e acessível para precificar o ambiente de preservação digital do Arquivo Nacional, com base em experiências anteriores.

Apesar da busca por modelos de custos para a preservação digital ao longo dos anos, essa tarefa permanece desafiadora devido à complexidade e particularidades da área (Boté; Fernandez-Feijoo; Ruiz, 2013). Estudos sobre custos de preservação digital têm aumentado desde 2010, contribuindo para o entendimento dos fatores que influenciam esses modelos, embora a aplicabilidade desses estudos em diferentes contextos ainda seja limitada (Palaiologk *et al.*, 2012).

Para este trabalho foram analisados quatro artigos mais significativos em termos de citações, relacionados à precificação de ambiente de preservação digital. O artigo mais relevante é o da NASA: *Cost Estimation Toolkit* (CET), ferramenta para estimar custos do ciclo de vida para implementação, operação e manutenção de um sistema de preservação digital de dados científicos. O estudo se baseia no modelo OAIS para mapear os processos de negócio, e se utiliza de uma abordagem de custeio baseado em atividades (ABC) para encontrar os valores (Fontaine *et al.*, 2007).

Outro trabalho muito referenciado é o projeto LIFE, desenvolvido pela *University College London* (UCL). Esse projeto é uma metodologia que modela o ciclo

de vida de um objeto digital, dentro do ambiente de preservação, a fim de se estimar seus custos. A modelagem também se baseia no modelo OAIS para encontrar os processos e, em seguida, as atividades envolvidas na preservação. Segundo seus autores, a metodologia pode ser aplicada a outras instituições de patrimônio cultural de qualquer dimensão (Hole *et al.*, 2010).

O terceiro estudo analisado neste trabalho foi o projeto *Keep Research Data Safe* (KRDS). O KRDS também foi baseado no modelo OAIS, para identificação de funções e processo, e no modelo ABC para precificação das atividades. O modelo divide o ciclo de vida da preservação digital em fases, em atividades e subatividades, cada uma das quais representando uma variável de custo (Beagrie; Voie; Woollard, 2010).

Por fim, o último estudo a ser analisado foi *Cost Model for Digital Archiving* (CMDA). A proposta é utilizar o modelo OAIS para encontrar os processos e atividades e, em seguida, o modelo ABC para precificá-las. Da mesma forma, o modelo estabelece vinculações com outros sistemas gerenciais, como o *Balanced Scorecard* (Palaiologk *et al.*, 2012).

Os artigos acima mencionados foram a base da proposta ora aplicada. Este estudo também se inspirou no modelo de referência OAIS, para encontrar as principais funções em um ambiente de preservação digital para, em seguida, detalhar os processos relacionados e estabelecer ligações com outros modelos de TIC. Dessa forma, foi possível precificar os processos/controles não havendo necessidade de detalhamento das atividades, pelo menos não nesta primeira versão.

3 Materiais e métodos

A metodologia adotada neste estudo está fundamentada na pesquisa-ação, a qual se caracteriza por combinar a realização de uma ação ou mudança com uma investigação. No contexto organizacional, a pesquisa-ação é uma forma de investigação interna que visa resolver problemas ou aprimorar práticas por meio da reflexão crítica e da implementação de medidas de intervenção (Nogueira *et al.*, 2013) (Altrichter *et al.*, 2002).

A pesquisa-ação realizada neste estudo envolveu a revisão dos artigos citados na seção de revisão de literatura. Durante essa etapa, foram examinados dados, normas e modelos pertinentes ao cálculo de custos em ambientes de preservação digital. A metodologia adotada baseou-se em uma abordagem qualitativa, que possibilitou a análise dos estudos encontrados nos artigos e documentos selecionados. Essa abordagem qualitativa permite uma interpretação mais profunda da realidade, considerando a perspectiva dos indivíduos ou do contexto pesquisado (Silva; Gobbi; Simão, 2005).

A seleção dos materiais de estudo foi feita de forma não probabilística e intencional, alinhada aos objetivos da pesquisa e aos modelos de cálculo de custo em questão. Foram escolhidos quatro artigos científicos que abordam modelos de custo,

todos eles de alguma forma relacionados ao modelo OAIS e aos processos de negócios da organização estudada. Essa seleção teve como objetivo principal levantar informações sobre os custos envolvidos no ambiente de preservação digital.

A partir da análise dos estudos selecionados, tornou-se claro como os elementos de um ambiente de preservação são precificados. Além disso, ao estudar os modelos de Arquivologia e identificar os processos de negócios que afetam os custos do ambiente de preservação, foi possível determinar as principais funções e correlacioná-las com modelos de Tecnologia da Informação e Comunicação (TIC), facilitando assim sua precificação. Com base nessas conclusões, o estudo elaborou um mapa conceitual do Arquivo Nacional como etapa preliminar para precificar seu ambiente de preservação.

4 Resultados e Discussão

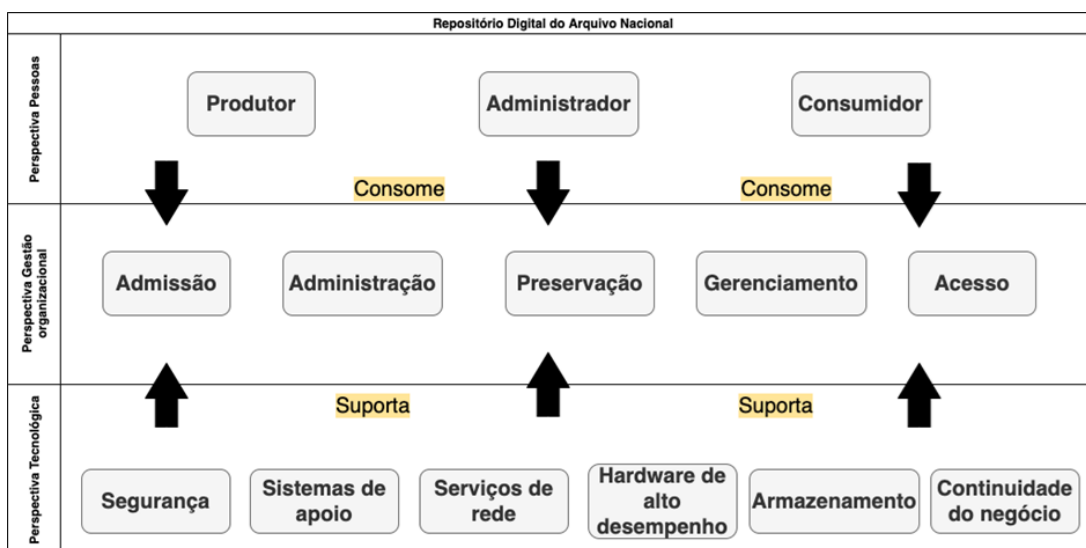
Aqui são apresentados os resultados encontrados a partir da aplicação da metodologia mencionada na seção anterior. Em seguida, o estudo propõe um mapa conceitual e uma planilha com os custos do ambiente de preservação, mais especificamente o repositório digital do Arquivo Nacional. Importante mencionar que a visão final do estudo está relacionada com a tecnologia, ou seja, segundo as premissas de um repositório digital, em especial a ISO 16363:2012 (*International Organization for Standardization* 2012, p. 16). Toda a organização está envolvida no processo. De acordo com esta norma, o processo está dividido em três seções principais: Infraestrutura Organizacional, Gestão de Objetos Digitais e Infraestrutura Tecnológica (Rezende; Cruz-Riascos; Hott, 2017), sendo esta última o foco deste trabalho.

O primeiro resultado é um mapa conceitual com as principais funções e atores envolvidos em um repositório digital (RD). A ideia através desta visão é esclarecer quais funções são utilizadas para dar suporte técnico ao RD. A utilização destas funções faculta o estabelecimento de conexões com os modelos mais utilizados nos sistemas de informações computacionais, destacando os elementos físicos que possam ser precificados. Lembramos que, a partir das funções de tecnologia encontradas, o caso em questão é a estrutura tecnológica do Arquivo Nacional. Abaixo é apresentado o mapa conceitual seguido de um pequeno detalhamento.

Para se estabelecer um modelo de custos, conforme já mencionado, a perspectiva utilizada foi a tecnológica. Nesta perspectiva, foram levantadas seis funções principais retiradas da literatura de TIC (*International Organization for Standardization*, 2013). São elas: **Segurança**, **Sistemas de Apoio**, **Serviços de rede**, **Hardware de alto desempenho**, **Armazenamento** e **Continuidade do negócio**.

Essas funções são conhecidas da área de TIC e diversos modelos fazem referência a elas. Para este trabalho o foco foi a ISO 27001 e a ISO 27002 (*International Organization for Standardization*, 2013).

Figura 1. Mapa conceitual com as principais funções do repositório digital do AN



Fonte: AN, 2023

A partir das funções elencadas, o estudo buscou por controles das normas ISO 27k que representassem elementos físicos que pudessem ser precificados, como por exemplo, a **“Função Segurança, que contém controles de segurança física, estes encontrados nas dependências do Datacenter do Arquivo Nacional, mantidos por contratos de manutenção e técnicos especializados”**. Por fim, após o levantamento dos controles e seus elementos, foi montada uma planilha de custos.

A seguir, breve apresentação das funções (mapa conceitual do RD do AN) e os controles (itens da ISO 27002), elencados no estudo:

1. Função Segurança

Esta função trata de todas as questões relacionadas à segurança física e lógica necessárias para manter o ambiente de preservação do AN. Abaixo, os principais controles da ISO 27k sobre o tema:

- a. Controle Segurança Física;
- b. Controle Segurança de equipamentos;
- c. Controles de acesso.

2. Função Sistemas de Apoio

Esta função trata dos sistemas necessários para implementar e manter o ambiente de preservação do AN. Aqui estão contidos os sistemas ligados diretamente e indiretamente, imprescindíveis de estarem implantados. Abaixo, o principal controle de apoio a esta função:

- a. Aquisição e manutenção de sistema de informação.

3. Função Serviços de Rede

Nesta função estão contidos os arcabouços necessários de rede para se manter um ambiente de preservação. Pode-se citar desde o serviço de rede local até os serviços de *Internet* e *Extranet*. Abaixo, o controle principal:

- a. Gerenciamento de operações e comunicações

4. Função Hardware de alto desempenho

Esta função, apesar de muito específica, se faz necessária por tratar-se de todo ambiente de estações servidoras de alto desempenho, onde estão instalados os softwares e serviços (como portais). Aqui, basicamente, trata-se dos ativos que sustentam todos os serviços de TIC da casa. Abaixo, o principal controle:

- a. Aquisição e manutenção de sistema de informação

5. Armazenamento

Nesta função são tratados elementos e processos de armazenamento dos objetos digitais. O item mais importante desta função são os storages, sejam eles transitórios ou permanentes. Esses são representados por entidades lógicas e físicas. A seguir, o principal controle:

- a. Aquisição e manutenção de sistema de informação.

6. Função Continuidade de negócio

Nesta função são tratadas as questões de continuidade do ambiente, desde *backups* até *sites recovery* e tolerância a falhas. Abaixo, o principal controle:

- a. Gestão da continuidade de negócios.

Importante ressaltar que muitos controles e sub controles, segundo a ISO 27k, podem ser encontrados mais de uma vez em funções diferentes. A ideia não é um detalhamento profundo das Funções X Controles (ISO 27k), mas estabelecer uma visão entre as funções tecnológicas de um RD com os elementos físicos (salas cofres, equipamentos, técnicos especializados, dentre outras) adquiridos pelo AN, e que possam ser precificados.

A partir do levantamento acima é apresentada a planilha inicial dos custos. Esta planilha demonstra, inicialmente, o valor que o Arquivo Nacional paga pelo seu ambiente de preservação digital (na perspectiva tecnológica), onde o repositório está incluído. Vale reforçar que a estrutura construída para o ambiente de preservação é compartilhada por outros serviços de TIC da instituição. Isto não quer dizer que o custo não seja específico ou real, uma vez que a análise é feita para uma casa arquivística,

ou seja, a sua principal atividade é a **PRESERVAÇÃO** de seu acervo, seja ele físico ou digital.

Os valores encontrados na planilha são relativos aos custos com as empresas prestadoras de serviços e servidores públicos que atuam nas áreas de TIC, diretamente relacionados com as funções do ambiente de preservação do AN. Os valores são anuais e demonstram que o Arquivo Nacional gasta em torno de R\$ 2.300.000,00 (dois milhões e trezentos mil reais) por ano, para manter seu ambiente de preservação digital, relativo à infraestrutura tecnológica. Outro ponto é que esses valores se referem à manutenção do ambiente. Não foram incluídos, aqui, os custos de aquisição de equipamentos e de implantação dos serviços. Quanto ao tempo, os valores refletem o ano de 2020 em diante.

Quadro 1. Custos com empresas prestadoras de serviços e servidores públicos que atuam nas áreas de TIC

FUNÇÃO	CONTROLE	ATIVO	PRESTADOR	CUSTO ANUAL
Função Segurança				
	Controle Segurança Física	Datacenter	Prestador Extremo I	R\$ 500 k
	Controle Segurança de equipamentos			
	Controle de acesso			
Função Sistema de Apoio				
	Aquisição e manutenção de sistema de informação	Sistemas e Portais	Prestador Externo II	R\$ 1.200 k
			Prestador Interno I	R\$ 240 k
Função Serviços de Rede				
	Gerenciamento de operações e comunicações	Ativos e serviços de rede	Prestador Externo II	-
			Prestador Interno II	R\$ 160 k
Função Hardware de alto desempenho				
	Aquisição e manutenção de sistema de informação	Estações servidoras e estações de trabalho	Prestador Externo III	R\$ 120 k
			Prestador Interno III	R\$ 80 K
Função Armazenamento				
	Aquisição e manutenção de sistema de informação	Storages	Prestador Externo II	-
			Prestador Interno II	
Função Continuidade de Negócio				
	Gestão da continuidade de negócios	Datacenter	Prestador Externo I	-
TOTAL ANO				R\$ 2.300 k

Fonte: AN, 2023

A partir do custo anual, o próximo passo é estimar o valor por objeto digital. Para este trabalho, foi utilizada a definição de objeto digital, em que ele é definido como qualquer objeto de informação que possa ser representado através de uma sequência de dígitos binários (Ferreira, 2006). Na tabela abaixo, seguem os tipos de objetos digitais e seus possíveis formatos preservados no AN.

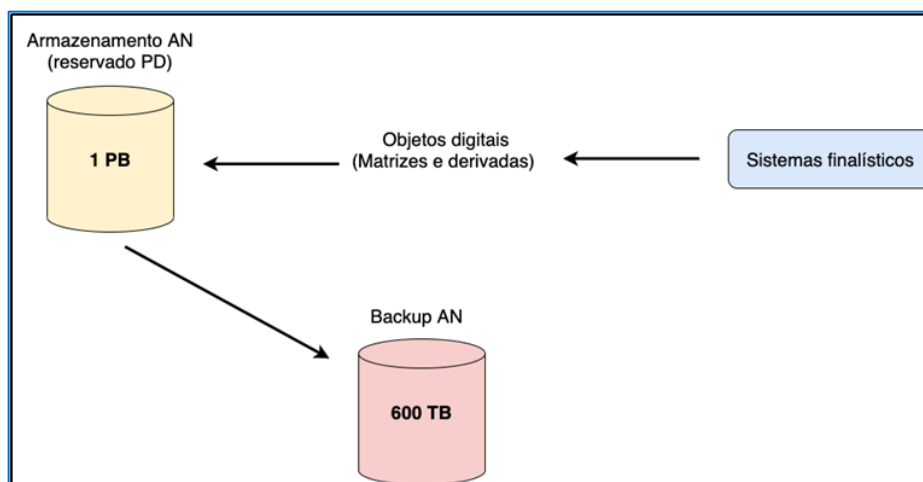
Tabela 1. Tipos de objetos digitais e seus possíveis formatos

Formato de origem	Possíveis formatos digitais	Unidade de armazenamento
Texto	PDF, DOC(x), ODT entre outros	Espaço em KB
Áudio	WAV	Espaço em KB
Áudio + imagem	MPEG	Espaço em KB
Imagem	TIFF, JPEG, BPM, PNG entre outros	Espaço em KB

Fonte: AN, 2023

Atualmente, o sistema de armazenamento do AN tem capacidade para 2 (dois) PB (Peta Bytes). Destes, 1 PB estão reservados para todo ambiente de preservação da instituição. Para simplificar o cálculo, pode-se estabelecer que todo ambiente de preservação é composto de entidades de Armazenamento, Backup, Sistemas finalísticos e objetos digitais preservados, sejam eles matrizes (documentos originais) ou derivadas (documentos convertidos). Na imagem abaixo é apresentada uma visão do sistema de preservação do AN, bem como as entidades acima citadas.

Figura 2. Visão do sistema de preservação do AN e as principais entidades funcionais



Fonte: AN, 2023

Os sistemas finalísticos são responsáveis por produzir e armazenar 24.200 milhões de objetos digitais, ocupando um total aproximado de 270 (duzentos e setenta) TB (Tera Bytes) de espaço. A partir de uma conta simples pode-se chegar ao tamanho médio do objeto: 283.000.000 MB / 24.200.000 objetos, obtendo-se um resultado próximo de 12 MB por objeto digital. A partir deste valor, quantos objetos digitais de 12 MB podem ser inseridos na área de armazenamento do AN (reservada para preservação digital), no total de 2 PB? O total é de aproximadamente 183 milhões de objetos digitais (de 12 MB), necessários para preencher todo o armazenamento. Dividindo pelo valor de manutenção anual, encontra-se o custo médio aproximado de R\$ 0,013 centavos por objeto digital preservado, de 12 MB em média. Essa lógica foi aplicada no biênio 2020/2021.

5 Conclusão

Este trabalho apresentou resultados de uma pesquisa-qualitativa baseada em uma pesquisa-ação com a finalidade de entender modelos de cálculo de custo de ambientes de preservação e de como estes modelos podem ser aplicados no caso do Arquivo Nacional. Para tal, foram coletadas informações sobre cálculos de custos a partir de estudos de artigos, normas, modelos e outras fontes de informação. Optamos por selecionar amostras não probabilísticas intencionalmente selecionadas ligadas a modelos de cálculo de custo.

Através das primeiras análises, foi possível realizar a criação de um mapa conceitual contendo as principais funções e os atores envolvidos em um repositório digital (RD). O foco foi esclarecer quais funções são utilizadas para dar suporte técnico ao RD. Ao identificar essas funções, é possível estabelecer conexões mais facilmente com os modelos mais comuns em sistemas de informações computacionais e, em seguida, identificar os elementos físicos que podem ser valorizados monetariamente.

A partir do mapa conceitual das funções foi possível estabelecer um modelo de custos para uma perspectiva tecnológica. Lembramos que o recorte deste artigo é a perspectiva tecnológica, e, através deste mapa, foram encontradas seis funções que dão apoio ao ambiente de preservação. São elas: **segurança, sistemas de apoio, serviços de rede, hardware de alto desempenho e continuidade do negócio**. Tais funções são da área de TIC e diversos modelos fazem referências a elas. Para este trabalho o foco foi a ISO 27001 e a ISO 27002 (*International Organization for Standardization*, 2013).

Por fim, através das funções tecnológicas de apoio ao sistema de preservação do AN encontram-se os controles baseados em modelos de TIC para, em seguida, relacioná-los com valores. O resultado foi uma planilha que demonstra o valor que o AN desembolsa pelo seu ambiente de preservação digital, sob a perspectiva tecnológica, onde o repositório está incluído. Os custos foram expressos no total gasto no ambiente de preservação. O primeiro valor de aproximadamente R\$ 2.300 K e o segundo, o custo gasto por objeto digital preservado, no valor de R\$ 0,013 centavos por objeto de 12 MB em média.

Importante ressaltar que este estudo apresenta algumas limitações podendo se destacar como principal a simplicidade da planilha que relaciona Funções x Controles x Valor. Esta simplicidade poderá ser otimizada em análises futuras com maior detalhamento dos controles, possibilitando melhor adequação de custos e, conseqüentemente, o aperfeiçoamento da sua qualidade.

Referências

- ABRAMS, S. *et al.* **Total cost of preservation (TCP):** Cost modeling for sustainable services. Oakland, CA: UC Curation Center (California Digital Library), 2012.
- ALTRICHTER, H. *et al.* The concept of action research. **The Learning Organization**, United Kingdom, v. 9, n. 3, p. 125-131, 2002. Disponível em: <https://doi.org/10.1108/09696470210428840>. Acesso em: 28 jun. 2024.
- BEAGRIE, N. Are we there yet? Understanding digital preservation costs and benefits. **Digital Preservation Coalition**, 2017. Disponível em: <https://www.dpconline.org/blog/are-we-there-yet>. Acesso em: 28 jun. 2024.
- BEAGRIE, N.; LAVOIE, B.; WOOLLARD, M. **Keeping research data safe 2:** final report. [S. l.]: University of Essex, 2010.
- BOTÉ, J.; FERNANDEZ-FEIJOO, B.; RUIZ, S. Digital preservation cost: a cost accounting approach. **The Learning Organization**, United Kingdom, v. 20, n. 6, p. 419-432, 2013. DOI:10.1108/TLO-09-2013-0049
- BRASIL. Presidência da República. Casa Civil. **Lei n. 8.159, de 8 de janeiro de 1991.** Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Brasília, DF, 1991.
- CONARQ – CONSELHO NACIONAL DE ARQUIVOS. **Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis – RDC-Arq.** Rio de Janeiro, RJ, 2015.
- FAPESP – FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO. **Plano de gestão de dados.** São Paulo, SP, s.d. Disponível em: <http://www.fapesp.br/gestaodedados>. Acesso em: 04 fev. 2019.
- FERREIRA, M. **Introdução à preservação digital:** conceitos, estratégias e atuais consensos. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006.
- FONTAINE, K. *et al.* **Observations on cost modeling and performance measurement of long-term archives.** [S. l.]: NASA, 2007.
- GRÁCIO, J. C. A. **Preservação digital na gestão da informação:** um modelo processual para as instituições de ensino superior. São Paulo, SP: Cultura Acadêmica, 2012.
- HOLE, B. *et al.* The Life3 predictive costing tool for digital collections. **New Review of Information Networking**, London, v. 15, n. 2, p. 81-93, 2010. Disponível em: <https://doi.org/10.1080/13614576.2010.526014>. Acesso em: 28 jun. 2024.
- ISO – INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 16363.** Genebra, 2012.

LÉVY, P. **Cibercultura**. 3. ed. São Paulo, SP: Editora 34, 2010.

NOGUEIRA, J. M. *et al.* Leveraging the Zachman framework implementation using action-research methodology – a case study: aligning the enterprise architecture and the business goals. **Enterprise Information Systems**, Oxfordshire, v. 7, n. 1, p. 100-132, 2013. Disponível em:

<https://www.tandfonline.com/doi/full/10.1080/17517575.2012.678387>. Acesso em: 28 jun. 2024.

PALAIOLOGK, A. S. *et al.* An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS. **International Journal on Digital Libraries**, New York, NY, v. 12, p. 195-214, 2012. Disponível em:

<https://link.springer.com/article/10.1007/s00799-012-0092-1>. Acesso em: 28 jun. 2024.

PARZIALE, A.; CAPRIOLO, G.; MARCELLI, A. One step is not enough: a multi-step procedure for building the training set of a query by string keyword spotting system to assist the transcription of historical document. **Journal of Imaging**, Basel, Switzerland, v. 6, n. 10, 2020. Disponível em:

<https://pubmed.ncbi.nlm.nih.gov/34460550/>. Acesso em: 28 jun. 2024.

REZENDE, L. V. R.; CRUZ-RIASCOS, S. S. A.; HOTT, D. F. M. Em busca de repositórios digitais confiáveis no Brasil: análise da infraestrutura organizacional conforme a norma ISO 16363/2012. **Reciis – Revista Eletrônica de Comunicação Informação & Inovação em Saúde**, Manguinhos, RJ, v. 11, supl., 2017. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/1390>. Acesso em: 28 jun. 2024.

SANTOS, H. M.; FLORES, D. Responsabilidades de um Repositório Arquivístico Digital Confiável na perspectiva do Open Archival Information System. **Páginas a&b Arquivos e Bibliotecas**, Porto, Portugal, n. 11, p. 116-132, 2019. Disponível em:

<https://ojs.letras.up.pt/index.php/paginasaeb/article/view/5459>. Acesso em: 28 jun. 2024.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, João Pessoa, PB, v. 22, n. 3, p. 179-191, 2012. Disponível em:

<https://periodicos.ufpb.br/ojs/index.php/ies/article/view/12224>. Acesso em: 23 fev. 2013.

SILVA, C. R.; GOBBI, B. C.; SIMÃO, A. A. O uso da análise de conteúdo como ferramenta para a pesquisa qualitativa: descrição e aplicação do método. **Organizações Rurais & Agroindustriais**, Lavras, MG, v. 7, n. 1, p. 70-81, 2005. Disponível em:

<https://www.revista.dae.ufla.br/index.php/ora/article/view/210>. Acesso em: 28 jun. 2024.