



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO – UNIRIO
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS – CCH
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO DE DOCUMENTOS E
ARQUIVOS – PPGARQ

GABRIELA RONDADO FERREIRA

PARÂMETROS PARA A MANUTENÇÃO DA AUTENTICIDADE DO
DOCUMENTO DIGITAL A PARTIR DA CONVERSÃO DE FORMATOS NA
PRESERVAÇÃO DIGITAL

RIO DE JANEIRO

2025

GABRIELA RONDADO FERREIRA

**PARÂMETROS PARA A MANUTENÇÃO DA AUTENTICIDADE DO
DOCUMENTO DIGITAL A PARTIR DA CONVERSÃO DE FORMATOS NA
PRESERVAÇÃO DIGITAL**

Dissertação apresentada ao Programa de Pós-graduação em Gestão de Documentos e Arquivos, do Centro de Ciências Humanas e Sociais, da Universidade Federal do Estado do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Gestão de Documentos e Arquivos.

Orientadora: Profa. Dra. Brenda Couto de Brito
Rocco Coorientadora: Profa. Ma. Cláudia
Carvalho Masset Lacombe Rocha

RIO DE JANEIRO

2025

FICHA CATALOGRÁFICA

Catálogo informatizada pelo(a) autor(a)

R RONDADO FERREIRA, GABRIELA
PARÂMETROS PARA A MANUTENÇÃO DA AUTENTICIDADE DO
DOCUMENTO DIGITAL A PARTIR DA CONVERSÃO DE FORMATOS NA
PRESERVAÇÃO DIGITAL / GABRIELA RONDADO FERREIRA. -- Rio de
Janeiro : UNIRIO, 2025.
8

Orientadora: BRENDA Couto de Brito Rocco .
Coorientadora: CLÁUDIA Carvalho Masset Lacombe Rocha.
Dissertação (Mestrado) - Universidade Federal do Estado
do Rio de Janeiro, Programa de Pós-Graduação em Gestão de
Documentos e Arquivos, 2025.

1. PRESERVAÇÃO DIGITAL. 2. PROPRIEDADES SIGNIFICATIVAS.
3. CONVERSÃO DE FORMATOS. I. Couto de Brito Rocco ,
BRENDA, orient. II. Carvalho Masset Lacombe Rocha,
CLÁUDIA, coorient. III. Título.

GABRIELA RONDADO FERREIRA

**PARÂMETROS PARA A MANUTENÇÃO DA AUTENTICIDADE DO
DOCUMENTO DIGITAL A PARTIR DA CONVERSÃO DE FORMATOS NA
PRESERVAÇÃO DIGITAL**

Dissertação apresentada ao Programa de Pós-graduação em Gestão de Documentos e Arquivos, do Centro de Ciências Humanas e Sociais da Universidade Federal do Estado do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Gestão de Documentos e Arquivos.

Aprovado em: __/__/__

BANCA EXAMINADORA:

Profa. Dra. Brenda Couto de Brito Rocco
Universidade Federal do Estado do Rio de Janeiro

Profa. Ma. Cláudia Carvalho Masset Lacombe Rocha
Arquivo Nacional

Profa. Dra. Diana Vilas Boas Souto Aleixo
Universidade Federal do Espírito Santo

Prof. Dr. João Marcus Assis
Universidade Federal do Estado do Rio de Janeiro

SUPLENTE:

Prof. Dr. Bruno Ferreira Leite
Universidade Federal do Estado do Rio de Janeiro

Dr. Eloi Juniti Yamaoka I
Serviço Federal de Processamento de Dados - Serpro

Dedico a mim mesma, pela força inabalável e pela coragem que encontrei em momentos de adversidade. Reconheço a persistência que me guiou através de desafios e a resiliência que me manteve firme, mesmo quando as tempestades pareciam intermináveis.

Celebro a capacidade de me abrir para processos difíceis, abraçando o crescimento pessoal e emocional.

Esta jornada é uma prova do meu compromisso com a autodescoberta e com a superação, lembrando-me sempre do poder e da sabedoria que reside dentro de mim.

AGRADECIMENTOS

Agradeço primeiramente a DEUS, às minhas orientadoras, Brenda Rocco e Cláudia Lacombe, pela paciência e pela caminhada ao meu lado.

Agradeço também à equipe do Arquivo Nacional/Copra/Dipad, que compreendeu todos os meus momentos e necessidades.

À Renata Barbato e à Thiara Alves, que gentilmente leram meu projeto inicial antes do ingresso no Programa e me deram dicas valiosas de como melhorá-lo.

Agradeço às minhas amigas Ari, Andressa e Natalia, pelo respeito e apoio incondicional nessa jornada, estando sempre ao meu lado, escutando minhas lamentações sobre a dificuldade e os desafios que este processo traz.

Agradeço à Lena, minha psicóloga, que me acompanhou durante todo o processo, mostrando-me que eu era capaz de enfrentar o “monstro” que é o processo do mestrado e ajudando-me a entender o porquê essa jornada é tão árdua.

Agradeço aos colegas do PPGARQ e aos professores pelas discussões durante as aulas, foi muito enriquecedor.

Agradeço também a Maria Luiza Liberato, Thalia de Melo e ao Carlos Eduardo Amand, que me ajudaram muito na etapa final dessa pesquisa.

Agradeço aos que gentilmente participaram da minha banca de qualificação, por suas indicações, e agradeço aos professores que aceitaram me avaliar no processo final da defesa, pois todos vocês foram muito importantes para o meu aprendizado.

Gratidão pelo tempo de vocês.

Uma fase difícil não determina a vida, não determina o grau de conhecimento, nem tampouco a habilidade de fazer coisas boas, ser competente e levar esse processo até o final.

RESUMO

Esta pesquisa aborda um tema de crescente relevância na Arquivologia: a preservação de documentos arquivísticos digitais, com enfoque na estratégia de preservação de conversão de formatos digitais aplicado aos documentos pertencentes ao fundo da Comissão Nacional da Verdade recolhidos ao Arquivo Nacional do Brasil. Tem como objetivo geral investigar, diante da necessidade de garantir a presunção de autenticidade, se as propriedades significativas, são preservadas durante o processo de migração formatos dos documentos arquivísticos digitais. Esta pesquisa possui um caráter exploratório e qualitativo, baseada em um estudo de caso do acervo digital da Comissão Nacional da Verdade. Visa também verificar se as propriedades significativas dos documentos foram mantidas durante a conversão de formatos no processo de submissão à plataforma de preservação denominada Archivematica. O estudo parte da presunção da autenticidade dos documentos arquivísticos digitais e da necessidade de identificar e preservar suas propriedades significativas. Para garantir a autenticidade digital, foi analisada com base na metodologia do projeto *Investigating the Significant Properties of Electronic Content over Time* (Inspect). Foram examinados os documentos de áudio e imagens matriciais recolhidos. Após a comparação das propriedades significativas de objetos digitais antes e depois de sua submissão a uma plataforma de preservação e conversão para formatos de longo prazo. A pesquisa alcançou resultado satisfatório. Sendo assim, sugere-se que outros trabalhos, a partir dos resultados aqui apresentados, possam ser elaborados para que auxiliem na decisão sobre quais propriedades significativas são importantes a partir dos formatos oficiais de objetos digitais que serão recolhidos ao Arquivo Nacional pela Administração Pública Federal.

Palavras-chave: Propriedades significativas; conversão de formato; CNV; Projeto Inspect.

ABSTRACT

This research addresses a topic of increasing relevance in Archival Science: the preservation of digital archival documents, focusing on the preservation strategy of digital format conversion applied to documents belonging to the National Truth Commission's collection collected by the National Archives of Brazil. Its general objective is to investigate, given the need to guarantee the presumption of authenticity, whether the significant properties are preserved during the format migration process of digital archival documents. This research has an exploratory and qualitative nature, based on a case study of the digital collection of the National Truth Commission. It also aims to verify whether the significant properties of the documents were maintained during the format conversion in the submission process to the preservation platform called Archivematica. The study is based on the presumption of authenticity of digital archival documents and the need to identify and preserve their significant properties. To guarantee digital authenticity, it was analyzed based on the methodology of the project Investigating the Significant Properties of Electronic Content over Time (Inspect). The audio documents and matrix images collected were examined. After comparing the significant properties of digital objects before and after their submission to a preservation platform and conversion to long-term formats, the research achieved satisfactory results. Therefore, it is suggested that other studies, based on the results presented here, can be developed to assist in the decision on which significant properties are important based on the official formats of digital objects that will be collected by the National Archives by the Federal Public Administration.

Keywords: significant properties; CNV; format conversion; Inspect Project.

LISTA DE QUADROS

Quadro 1 - Conceito de Documento	21
Quadro 2 - Conceito de Objeto Digital.....	24
Quadro 3 - Conceito de Autenticidade	28
Quadro 4 -Propriedades significativas de registro sonoro.....	51
Quadro 5 - Propriedades significativas de registro sonoro no formato BEXT.....	51
Quadro 6 - Propriedades significativas de mensagens de correio eletrônico	52
Quadro 7 -Propriedades significativas de imagem matricial.....	52
Quadro 8 - Propriedades significativas de texto estruturado	53
Quadro 9 - Áudio 0001: extensão .wma.....	66
Quadro 10 - Audio 0002: extensão .wma.....	67
Quadro 11 - Áudio 0003: extensão .wma.....	67
Quadro 12 - Áudio 0004: extensão .wma.....	68
Quadro 13 - Áudio 0001: extensão .mp3.....	68
Quadro 14 - Áudio 0002: extensão .mp3.....	69
Quadro 15 - Áudio 0003: extensão .mp3.....	69
Quadro 16 - Áudio 0004: Extensão .mp3	70
Quadro 17 - Imagem matricial 0001: extensão .bmp	71
Quadro 18 - Imagem matricial 0002: extensão .bmp	72
Quadro 19 - Imagem matricial 0003: extensão .bmp	72
Quadro 20 - Imagem matricial 0004: extensão .bmp	73
Quadro 21 - Imagem matricial 0001: extensão .psd.....	74
Quadro 22 - Imagem matricial 0002: extensão .psd.....	75
Quadro 23 - Imagem matricial 0001: extensão .tiff	76
Quadro 24 - Imagem matricial 0002: extensão .tiff	76
Quadro 25 - Imagem matricial 0003: extensão .tiff	77
Quadro 26 - Imagem matricial 0004: extensão .tiff	78
Quadro 27 - Imagem matricial 0001 – extensão .jpeg.....	79
Quadro 28 - Imagem matricial 0002: extensão .jpeg	79
Quadro 29 - Imagem matricial 0003: extensão .jpeg	80
Quadro 30 - Imagem matricial 0004: extensão .jpeg	81

LISTA DE FIGURAS

Figura 1- Cadeia de Interpretação do nível físico ao conceitual	26
Figura 2 - Modelo Funcional OAIS.....	39
Figura 3 -Quantitativo de extensão em formato áudio	62
Figura 4 -Quantitativo de extensão em formato imagem matricial	63

LISTA DE ABREVIATURAS E SIGLAS

AIP - *Archival Information Package*

NA - Arquivo Nacional do Brasil

CCSDS - *Consultive Committee for Space Data Systems*

CNV - Comissão Nacional da Verdade

CONARQ - Conselho Nacional de Arquivos

DIP - *Dissemination Information Package*

HD - *Hard Disk*

JISC - *Joint Information Systems Committee*

INSPECT - *Investigating the Significant Properties of Electronic Content over Time*

ISO - *International Organization for Standardization*

METS - *Metadata Encoding & Transmission Standad*

SIGA - Sistema de Gestão de Documentos de Arquivo

SIGAD - Sistemas Informatizados de Gestão Arquivística de Documentos

OAIS Open - *Archival Information System*

PREMI - *Preservation Metadata: Implementation Strategies*

RDC-Arq - Repositórios Arquivísticos Digitais Confiáveis

SAAI - Sistema Aberto de Arquivamento de Informação

SINAR - Sistema Nacional de Arquivos

SIP - *Submission Information Package*

SO - Sistema Operacional

TI - Tecnologias da Informação

SUMÁRIO

1 INTRODUÇÃO	13
2 A EVOLUÇÃO DOS REGISTROS DOCUMENTAIS E A PRESERVAÇÃO DAS PROPRIEDADES SIGNIFICATIVAS EM DOCUMENTOS ARQUIVÍSTICOS DIGITAIS.....	18
3 PRESERVAÇÃO DIGITAL: CONCEITO, DESAFIOS E ESTRATÉGIAS FRENTE AOS DOCUMENTOS ARQUIVÍSTICOS	30
4 ARQUIVO NACIONAL: EM FOCO SEU REPOSITÓRIO.....	41
5 PROJETO INSPECT: A DEFINIÇÃO DE PROPRIEDADES SIGNIFICATIVAS ...	45
6 PROCEDIMENTOS METODOLÓGICOS	59
6.1 ETAPAS DA PESQUISA.....	59
6.1.1 Identificação e seleção dos documentos.....	59
6.1.2 Processamento e submissão ao Sistema de Preservação	60
6.1.3 Análise comparativa dos resultados.....	60
6.1.4 Ferramentas utilizadas.....	61
7 RESULTADOS	62
7.1 DOCUMENTOS DA COMISSÃO NACIONAL DA VERDADE RECOLHIDOS PARA O ARQUIVO NACIONAL;.....	62
7.2 ANÁLISE DAS PROPRIEDADES SIGNIFICATIVAS DOS DOCUMENTOS DIGITAIS COM BASE NA METODOLOGIA DO PROJETO INSPECT.....	63
7.3 ESTUDO DAS VARIAÇÕES NAS PROPRIEDADES SIGNIFICATIVAS ENTRE OS DOCUMENTOS DIGITAIS RECOLHIDOS E PRESERVADOS NO REPOSITÓRIO DIGITAL CONFIÁVEL DO ARQUIVO NACIONAL	64
8 CONSIDERAÇÕES FINAIS.....	83
REFERÊNCIAS	85

1 INTRODUÇÃO

Este trabalho está centrado na análise da manutenção da autenticidade dos documentos digitais a partir de suas propriedades significativas, da forma como são preservados e de como as estratégias de preservação influenciam essa autenticidade. Além disso, avaliamos como os documentos chegam à idade permanente mantendo sua autenticidade e propriedades significativas ao longo de todo o ciclo de vida do documento digital.

Com a disseminação das Tecnologias Digitais da Informação e da Comunicação (TDICs) a sociedade passou por mudanças em diversos aspectos profissionais e pessoais. Destaca-se, nesta pesquisa, a forma de produzir e registrar informações em formatos digitais. O uso de computadores pessoais (PCs) facilitou e acelerou diversas atividades, resultando na produção de documentos digitais.

A Arquivologia também foi influenciada pelas TDICs, uma vez que os documentos arquivísticos passaram a ser produzidos, tramitados e armazenados digitalmente, trazendo novos desafios para a área. Com a produção de documentos arquivísticos digitais, surgiu a necessidade de preservar uma série de informações sobre eles, referenciada como metadados. É importante registrar e preservar o contexto tecnológico e de produção dos documentos para que futuros usuários possam entendê-los.

A preservação dos documentos continua sendo determinada por sua capacidade de atender às utilizações imputadas, garantindo que estas permaneçam satisfatórias para usos futuros.

Em observância ao compromisso de preservação, os documentos arquivísticos digitais devem ser amparados por políticas, normas e regulamentações para garantir a longevidade de acesso aos usuários. Esse compromisso está intrinsicamente ligado ao Poder Público na perpetuação da memória, às instituições custodiadoras desse patrimônio documental e aos profissionais atuantes nesse campo.

O Arquivo Nacional do Brasil (AN), instituição pertencente à administração pública federal e órgão central do Sistema de Gestão de Documentos e Arquivos (Siga), publicou diversas recomendações para a preservação de documentos arquivísticos no Brasil. Destaca-se a Política de Preservação Digital do AN, de 2016, que define estratégias para o recolhimento, preservação e acesso aos documentos arquivísticos digitais recebidos no Arquivo Nacional, empenhado em atender ao compromisso com a comunidade interessada (Arquivo Nacional, 2016).

No conteúdo publicado dessa política, especificamente na seção 3, o Arquivo Nacional assumiu o compromisso da preservação dos documentos arquivísticos digitais recolhidos bem como, a preservação das características desses documentos. O compromisso está ancorado nos tipos de documentos e formatos aceitos e ainda a implementação de estratégias de preservação. E reafirma o compromisso com a garantia de que os documentos poderão ser recuperados e não haverá perda das propriedades significativas. Essas propriedades significativas são as características que devem ser preservadas desses objetos digitais, elas foram definidas para cada tipo de formato estudado pelo Projeto *Investigating Significant Properties of Digital Objects* (Inspect).

O projeto Inspect é uma metodologia pouco difundida no Brasil, que tem como propósito estudar o conceito de propriedades significativas e estabelecer as características dos objetos digitais estudados para aproximar a aplicação na execução das estratégias de preservação digital.

Entre 2007 e 2009, o projeto britânico, denominado, projeto Inspect realizou estudos que culminaram na proposição de uma metodologia e na identificação de propriedades significativas para quatro tipos específicos de objetos digitais: áudios, e-mails, imagens matriciais e textos estruturados (GRACE *et al.*, 2009).

A investigação resultou na criação de uma listagem com as características técnicas que devem ser observadas em diferentes tipos de objetos digitais. Essas listas serviram como norteador para lidar com a observação da conversão desses objetos perante as transformações recorrentes do mundo digital.

Campos (2002, p. 9) reflete que “[...] o mundo digital é, essencialmente, um mundo de mudança e de constante dinâmica onde a tecnologia está em permanente mutação, o ambiente legal é sujeito a revisões e atualizações e os próprios objetos digitais são, também eles, dinâmicos.” Considerando o autor apresentado, a produção de documentos nato-digitais sem critérios mínimos de preservação pode estar fadada à perda. A obsolescência tecnológica é um desafio significativo na produção de documentos arquivísticos digitais, ligada diretamente à evolução tecnológica, formatos de arquivos ultrapassados e sistemas operacionais em desuso, tornando a preservação um grande desafio.

Tal preocupação advém de observações práticas, como o comportamento da comunidade interessada e a rapidez com que novas ferramentas e *softwares* se tornam obsoletos, impactando o acesso aos conteúdos digitais, para a manutenção dos acessos a esses documentos, apoiando nas estratégias de preservação digital. E ainda, à medida que cresce a produção e o

uso de documentos arquivísticos digitais, aumenta a preocupação entre estudiosos da Arquivologia, entidades custodiadoras, usuários e profissionais da informação.

Este pode ser um meio que a comunidade encontrou para lidar com a obsolescência tecnológica. Considerando que este trabalho centra nas estratégias de migração dentro das estratégias de preservação digital, esta permite acompanhar as atualizações de versões de arquivos e formatos e mitigar o avanço da obsolescência tecnológica.

A migração é basicamente a conversão de um formato X em Y para que seja possível permanecer acessível a longo prazo, sendo assim, quando um registro digital passa por uma conversão de formato para fins de preservação digital, suas características podem ser analisadas para se verificar a manutenção das propriedades significativas. Essa abordagem, de verificação da conversão, se mostra especialmente pertinente para o acervo da Comissão Nacional da Verdade (CNV), que contém registros digitais essenciais para essa análise, esses documentos foram produzidos ao longo da atuação a comissão.

A Comissão Nacional da Verdade foi um colegiado instituído pelo Estado brasileiro para investigar graves violações de direitos humanos ocorridas entre 1946 e 1988, com foco especial nas violações durante a ditadura militar (1964-1985), foi criada pela Lei nº 12.528/2011 e oficialmente instalada em 16 de maio de 2012. A comissão teve como objetivo examinar e esclarecer essas violações, ouvindo vítimas e testemunhas, convocando agentes da repressão para depoimentos, e promovendo eventos públicos para apresentar relatórios preliminares. Em 10 de dezembro de 2014, a CNV entregou seu relatório final à presidente Dilma Rousseff, concluindo que as práticas de detenções ilegais, tortura, execuções e desaparecimentos forçados eram parte de uma política estatal contra a população civil, esse relatório serviu para documentar e divulgar os fatos para que a sociedade conheça a verdade sobre esse período.

A CNV e o AN têm uma relação importante na preservação da memória e na documentação das violações de direitos humanos ocorridas durante a ditadura militar no Brasil.

A Comissão da Verdade existiu para completar as peças e esclarecer o que aconteceu e para que nossa história se complete, e o Arquivo nacional tem como papel fundamental trabalhar para que essa memória seja preservada para o acesso a longo prazo possa contribuir para que sociedade se esqueça e para que nunca mais aconteça.

Nesse sentido, com vistas a apoiar a presunção de autenticidade de documentos arquivísticos digitais, produzidos advindos da CNV e recolhidos ao Arquivo Nacional, indaga-se: como o processo de migração de formatos de documentos arquivísticos digitais impacta a preservação das propriedades significativas e a presunção de autenticidade desses documentos,

especialmente no momento do recolhimento para a entidade custodiadora? Diante desse problema, buscamos atingir os objetivos apresentados.

A presente dissertação teve como objetivo geral investigar como as propriedades que garantem a presunção de autenticidade são preservadas durante o processo de migração de formatos dos documentos arquivísticos digitais no momento de seu recolhimento pela entidade custodiadora. Este estudo de caso foi realizado utilizando o acervo da Comissão Nacional da Verdade. Para o alcance do objetivo geral, essa pesquisa percorreu os seguintes objetivos específicos:

- 1) Identificar os documentos recolhidos pela Comissão Nacional da Verdade (CNV) para o Arquivo Nacional;
- 2) Analisar as propriedades significativas desses documentos digitais, com base nos relatórios e demais documentos relacionados ao recolhimento do Arquivo Nacional tendo como apoio a metodologia adotada pelo Projeto Inspect;
- 3) Analisar as variações nas propriedades significativas de documentos digitais antes e após sua conversão para formatos de preservação digital.

A pesquisa desenvolvida visou contribuir para a problemática da conversão de documentos nato-digitais quando submetidos a plataforma de preservação digital do Arquivo Nacional, do ponto de vista social, a justificativa está ancorada na manutenção da memória de Estado Brasileiro, na relevância desses documentos produzidos, na importância dessa comissão para a sociedade, a fim de preservar a memória para promover a justiça e fortalecer a democracia.

Do ponto de vista científico, justifica pela celeridade da obsolescência tecnológica tornando-se uma causa que incide na verificação da manutenção da autenticidade desses documentos nato-digitais. Também na tentativa de difundir um pouco mais a metodologia do projeto Inspect. Trazer elementos para a verificação desses dados se colocando com agente observador frente as dificuldades que podem aparecer no campo da preservação digital e identificar possíveis ações para mitigar riscos futuros. A partir dessa verificação de manutenção também é possível elaborar diretrizes e normas para orientação da comunidade acadêmica e profissionais e entidades custodiadoras atuantes na área de preservação digital.

E do ponto de vista pessoal, essa pesquisa se justifica por se tratar de uma necessidade de observação do campo empírico, *in loco*, que permite verificar a conversão na prática dos formatos e colaborar para a construção de métodos e recomendações a partir dessa pesquisa, da metodologia defendida pelo projeto Inspect, pouco difundida no Brasil, que enriquece o cenário

de análise das propriedades significativas desses objetos digitais e permite que estes sejam acessados pelo público a partir do Arquivo Nacional que disponibiliza esses documentos para consulta pública, permitindo que pesquisadores, familiares de vítimas e o público em geral tenham acesso às informações sobre as violações de direitos humanos. E para atender a transparência na promoção da justiça, permitindo que as violações sejam conhecidas e que os responsáveis sejam identificados. E por último, porém não menos importante, a preservação desses documentos ajuda a educar as novas gerações sobre os horrores da ditadura e a importância de defender os direitos humanos e a democracia e a manutenção da memória de um país.

Para alcançar os resultados este trabalho é composto por cinco seções, começando com uma introdução que contextualiza o cenário do objeto de estudo. A introdução também apresenta o problema de pesquisa, os objetivos (geral e específicos), a justificativa e a estrutura do trabalho. Com base nesse contexto, buscou-se alcançar os objetivos delineados nesta dissertação, que está organizada em cinco seções.

A Seção 1 é a Introdução, que apresenta a pesquisa, bem como um breve histórico da produção documental, da evolução tecnológica e suas instabilidades quanto à produção do documento digital, além da necessidade de manutenção da autenticidade desses documentos a partir de conversões para lidar com a questão da obsolescência tecnológica que leva ao problema de pesquisa.

A Seção 2 apresenta uma revisão bibliográfica, partindo dos conceitos de documento, documento arquivístico e documento arquivístico digital, identificando suas características para a manutenção da autenticidade dos documentos. Também apresenta os principais conceitos de preservação digital e a conceituação das propriedades significativas.

A Seção 3 é a apresentação da metodologia do Projeto Inspect para entendimento de como o projeto alcançou as propriedades significativas para diversos formatos, que embasa o procedimento de análise para alavancar o resultado desta pesquisa.

A Seção 4 apresenta a metodologia utilizada nesta pesquisa e os caminhos escolhidos para a coleta dos dados e análise do material coletado.

A Seção 5 contém a conclusão da pesquisa, as considerações finais sobre a dissertação e algumas sugestões para trabalhos futuros.

2 A EVOLUÇÃO DOS REGISTROS DOCUMENTAIS E A PRESERVAÇÃO DAS PROPRIEDADES SIGNIFICATIVAS EM DOCUMENTOS ARQUIVÍSTICOS DIGITAIS

Desde os seus primórdios, a humanidade desenvolveu mecanismos de registro para armazenar informações e facilitar a comunicação. Os primeiros vestígios da prática remontam às pinturas rupestres, criadas há mais de 30.000 anos, nas representações artísticas, os povos pré-históricos registravam cenas de caça, rituais e a vida cotidiana.

Com o surgimento da escrita, aproximadamente em 3.200 a.C., na região da Suméria, ocorreu uma significativa transformação nos mecanismos de transmissão de informações, que permitiu a codificação e o registro de conhecimentos de forma mais eficaz e duradoura, com destaque a escrita em argila, a escrita divina na Grécia (hieróglifos), além de outras formas que proporcionaram meios de comunicação e armazenamento das informações. Os egípcios, por sua vez, desenvolveram o papiro como suporte para a escrita, que permitiu maior mobilidade dos documentos. Silva *et al.* (2002)

No Antigo Oriente as tábuas de argila datadas de 4.000 a.C. foram conservadas e armazenadas em palácios e templos, demonstrando a organização daquela sociedade. Os palácios foram estudados por arqueólogos, que traduziram a história dos arquivos desde o quarto milênio até o século IV a.C.

Na antiga civilização grega, os atenienses guardavam documentos valiosos na corte de justiça e nos templos, onde eram conservados e transmitidos desde os tempos primitivos até por volta do século III. Com o desenvolvimento administrativo da sociedade grega, cada magistratura tinha seu próprio arquivo, contendo documentos expedidos pelas autoridades.

Durante o governo dos Césares, a organização do Império Romano evoluiu, resultando na criação de depósitos para arquivos que abrigavam documentos legislativos, de jurisprudência e administrativos das províncias. Essa burocracia contribuiu para a criação de um arquivo público, vinculado ao Estado e conferindo fé pública aos documentos. Ruiz Rodríguez (1995)

Com a queda do Império Romano na Idade Média, a administração construída pelo governo dos Césares foi perdida e a ideia de arquivo público deixou de existir.

Durante a Alta Idade Média, a concentração de documentos passou a ocorrer nas catedrais e monastérios, pois a igreja precisava dos documentos para respaldar a aquisição de bens e propriedades, conferindo fé pública aos documentos eclesiásticos custodiados.

Na Idade Moderna, com as revoluções burguesas no final do século XVIII, surgiu um

governo caracterizado por uma monarquia absolutista centralizadora, que levou à criação de um Arquivo do Estado. Esse arquivo concentrava os documentos gerados pela Coroa e se tornou um elemento fundamental da máquina administrativa, com função predominante jurídico-administrativa.

A Revolução Industrial, datada do século XVIII na Inglaterra e expandida mundialmente no século XIX, substituiu a manufatura por máquinas, aumentando consideravelmente a produção de produtos. Isso resultou em um aumento do contingente de pessoas, da economia e da sociedade, além de um crescimento significativo na produção de documentos, apontando a necessidade de organização dos arquivos.

Para Souza (2003) A centralização dos arquivos reacendeu o debate sobre a organização dos documentos. As formas tradicionais de ordenação foram desafiadas por um novo contexto.

Desde o início da Revolução Industrial, o aumento do consumo de novas tecnologias resultou em uma maior produção documental, o que levou à expansão das atividades governamentais. Com o progresso da sociedade, houve uma explosão documental, transformando uma pequena quantidade inicial em uma vasta massa de documentos. Como relata Schellenberg (2006, p. 65),

Esse conteúdo descreve como o aumento da população levou à expansão das atividades governamentais, o que, por consequência, impactou a produção de documentos. Com a implementação de métodos tecnológicos modernos na preparação dos documentos, o volume de documentos aumentou significativamente nas últimas décadas. Esse crescimento foi tão expressivo que seguiu um padrão geométrico, ao invés de aritmético, evidenciando um aumento exponencial em comparação com o crescimento linear.

O século XX apresentou mudanças significativas e entre elas se destacam: a crescente urbanização e concentração das indústrias; a aglomeração de trabalhadores; o surgimento e crescimento das cidades; o aumento demográfico, em parte devido às técnicas agrícolas que aumentaram a oferta de alimentos; o desenvolvimento da produção em massa com a introdução de máquinas; o rápido aprimoramento tecnológico e a maior divisão do trabalho; e a adoção de políticas econômicas. A Revolução Industrial é descrita como “[...] a mais radical transformação da vida humana já registrada em documentos escritos” (Hobsbawm, 2000, p. 13).

Ainda em meados os do século XX, principalmente após a II Guerra Mundial, o progresso científico, o progresso tecnológico e a expansão das atividades do Estado contribuíram significativamente para o aumento da produção de documentos. Tal crescimento foi tão elevado que excedeu a capacidade de controle e organização das instituições e demandou

a implementação de novos sistemas de gestão documental e tecnologias de informação para garantir a eficiência e a segurança na administração desses registros.

Considerando a Revolução Industrial como a maior transformação dos registros humanos, observa-se uma nova era informacional devido à produção acelerada de documentos. A Revolução da Informação, comparada à Revolução Industrial, apresenta um cenário similar, mas de forma diferenciada, sendo impulsionada por máquinas de produção de informação em vez de máquinas a vapor. Para Ducker (2000, p. 1), “A premissa é que a Revolução da Informação será como foi a Revolução Industrial no final do século XVIII e início do século XIX. E é exatamente assim que tem sido a Revolução da Informação nestes seus primeiros 50 anos”.

Os primeiros computadores surgiram nos anos 1940, marcando o início da Revolução da Informação. Com o advento dos computadores, vários mecanismos de produção de documentos e comunicações foram criados, evidenciando a ligação entre a Revolução Industrial e a Revolução da Informação. São duas grandes transformações em dois séculos, conforme abordado por Ducker (2000, p. 1):

[...] E a máquina a vapor foi para a primeira Revolução Industrial o que o computador foi para a Revolução da Informação – seu gatilho e, acima de tudo, seu símbolo. Quase todos atualmente acreditam que nada na história econômica evoluiu tão depressa nem teve tanto impacto quanto a Revolução da Informação. No entanto, a Revolução Industrial evoluiu, no mínimo, tão depressa quanto a Revolução da Informação no mesmo intervalo de tempo, e provavelmente teve um impacto igual, se não maior [...]

Comparando essas mudanças, observa-se a evolução da sociedade em relação à informação. A necessidade de comunicação forçou indiretamente os registros, provocando uma mudança de hábito na sociedade que veio acompanhada pela evolução tecnológica.

Rapidamente nos adaptamos a essas novidades e passamos – em geral, sem uma percepção clara nem maiores questionamentos – a viver na Sociedade da Informação, uma nova era em que a informação flui a velocidades e em quantidades há apenas poucos anos inimagináveis, assumindo valores sociais e econômicos fundamentais (Takahashi, 2000, p.1).

A sociedade acompanha essas transformações, apesar dos desafios impostos pela gigantesca massa de informação produzida. A adaptação às novas tecnologias, a velocidade e a quantidade de informações também facilitam a disseminação de notícias falsas e desinformação. Além disso, a falta de igualdade no acesso às tecnologias da informação pode criar desigualdades sociais e econômicas. A Revolução da Informação proporciona uma nova mercadoria: a informação, que se valoriza a cada dia, marcando a transição de uma revolução de máquinas para uma revolução de informação e conhecimento. Assim, diante do breve histórico analisado até aqui, torna-se fundamental para a presente pesquisa a análise de alguns

conceitos basilares. Esta investigação se baseou em referências da Arquivologia e da Ciência da Informação, assim, para apresentar alguns conceitos fundamentais para o desenvolvimento deste estudo, para a conceituação de documento tomamos como base o livro “O Documento Arquivístico ante a Realidade Digital”, Rondinelli (2013) apresenta um levantamento de diversos conceitos acerca de documento definidos por autores das áreas de ciência da informação, tal como sintetizado no quadro 1.

Quadro 1 - Conceito de Documento

Ano	Autor	Conceitos
1934	Otlet	As coisas materiais em si (objetos) podem ser consideradas documentos quando se colocam como elementos de estudo ou provas de uma demonstração.
1951	Briet	Todo índice concreto ou simbólico, conservado ou registrado com a finalidade de representar, reconstruir ou demonstrar um fenômeno físico intelectual.
1989	Cortes Alonso	A era testemunho da atividade do homem fixado em um suporte.
1991	Buckland	Informação-como-coisa: livros, textos, fotografias, edifícios, pessoas, eventos apontam para alguma informação, logo constituem evidência física, documento, ou seja, coisa a partir da qual se pode aprender.
1991	Heredia Herreira	Documento como testemunho da atividade do homem, fixado em um suporte perdurável que contém informação.
1996	Martin Pazuelo Campillos	Ferramenta indispensável para transmitir conhecimento, ideias e dar testemunho dos fatos.
1998	Duranti	Evidência produzida em um suporte.
1999	Redan Rojas	Objetivação do pensamento. Serve de orientação ao projeto social, humano. Só é documento quando submetido a certas estruturas pelo cientista da informação.
2002	Rodriguez Bravo	Todo o suporte portador de uma mensagem emitida com intenção comunicativa e potencialmente informativa para o receptor.
2002	Duranti	Informação registrada.

Fonte: Adaptado de Rondinelli (2013, p.54).

Ao analisar as definições expostas no Quadro 1, observação, essa, realizada pela autora, foi possível notar uma similaridade entre os conceitos de materialidade (conteúdo registrado em um determinado suporte) e funcionalidade (que abrange ensino, aprendizagem, registro e comunicação de informações, além de servir como testemunho de eventos e ações).

O Dicionário de Terminologia Arquivística (Dibrate) define documento como “uma unidade de registro de informações, qualquer que seja o suporte ou formato” (Arquivo Nacional, 2005, p. 73). A partir de tais definições pode-se compreender que um documento é uma informação registrada em um suporte material, utilizado para consulta, estudo, prova e

pesquisa, servindo como testemunho de fatos, fenômenos, formas de vida e pensamentos de uma época ou lugar específicos; é registro de informações que serve como prova ou evidência de atividades, transações ou eventos como esses apresentados nesta breve linha do tempo.

Conforme indicado nas Diretrizes do Produtor, que faz parte do Projeto *The International Research on Permanent Authentic Records in Electronic Systems InterPARES*, da Universidade de British Columbia, Canadá, em acordo técnico com o Arquivo Nacional e aprovadas pela Resolução nº 38 do CONARQ, um documento é definido como a representação afixada de um meio em uma forma imutável, ou seja, o conteúdo ou a apresentação desse documento se mantém inalterado sempre que é acessado, garantindo estabilidade e consistência ao longo do tempo. A relação entre tais características e o objeto digital é essencial para assegurar a autenticidade dos documentos digitais.

Pode-se observar com base no conceito de documentos duas características fundamentais: a forma fixa e o conteúdo estável. No Glossário da Câmara Técnica de Documentos Eletrônicos (CTDE) do Conselho Nacional de Arquivos (CONARQ, 2020), é possível enxergar tais conceitos como os que seguem:

- Conteúdo estável: característica de um documento arquivístico em que os dados e a mensagem nele contidos se mantêm inalterados e inalteráveis, i. e., não podem ser sobrescritos, alterados, apagados ou acrescentados (Conselho Nacional de Arquivos, 2020, p. 18); e
- Forma fixa: característica de um documento arquivístico que assegura que sua aparência ou apresentação documental permanece a mesma cada vez que o documento é manifestado (Conselho Nacional de Arquivos, 2020, p. 30).

Corroborando com as definições supracitadas, Rondinelli (2013) menciona que um documento digital é considerado de forma fixa e conteúdo estável quando sua exibição na tela do computador permanece constante, mesmo que a cadeia de codificação seja alterada. Isso significa que o mesmo documento digital pode ser visualizado através de diferentes codificações.

De acordo com Carta para a preservação do patrimônio arquivístico digital: preservar para garantir o acesso, emitida pelo Conselho Nacional de Arquivos – CONARQ, em 2006 Os documentos arquivísticos são criados e preservados por organizações e indivíduos com o objetivo de registrar suas atividades e atuar como fontes de evidência e informação.

O Glossário da CTDE do CONARQ (2020) define documento arquivístico como

“documento produzido (elaborado ou recebido), no curso de uma atividade prática, como instrumento ou resultado de tal atividade, e retido para ação ou referência” (CONARQ, 2020, p. 24).

A estabilidade do documento arquivístico é um conceito fundamental na preservação de registros históricos e na garantia da fidedignidade da informação ao longo do tempo. No contexto arquivístico, a estabilidade é entendida como a capacidade do documento de manter suas características essenciais e de resistir a alterações que possam comprometer sua integridade.

A estabilidade não se configura apenas como uma característica física, mas também como um atributo vinculado à credibilidade do conteúdo do documento, refletindo sua autenticidade e relevância para o patrimônio documental. No contexto diplomático, a estabilidade está associada à autenticidade formal do documento, garantindo que ele seja um registro fiel, elaborado e preservado em conformidade com práticas estabelecidas e oficialmente reconhecidas.

Do ponto de vista arquivístico, “a estabilidade está implícita no conceito de documento arquivístico à medida que este é predominantemente identificado como documento, e não um simples dado ou informação” (Rondinelli, 2013, p. 235). No caso de documento arquivístico digital, “essa vinculação se dá entre documento dentro e fora do sistema, nos chamados ambientes híbridos, que abrangem documentos digitais e não digitais ao mesmo tempo” (Durante; Thibodeau, 2008, p. 236 apud Rondinelli, 2013).

A diferença entre documentos arquivísticos e documentos arquivísticos digitais está na codificação de bits. Documentos arquivísticos são informações registradas em qualquer formato, autênticas, representando fielmente os fatos registrados, com identidade e integridade, ou seja, não são alterados ou adulterados.

Os documentos arquivísticos digitais, por sua vez, são codificados em dígitos binários, acessíveis e interpretáveis por meio de sistemas computacionais. Nesse sentido, Rondinelli (2013, p. 235) apresenta o seguinte conceito de documento arquivístico digital:

[...] é uma unidade indivisível de informação constituída por uma mensagem fixada num suporte (registrada), com uma sintática estável; e ainda, produzido e/ou recebido por uma pessoa física ou jurídica, no decorrer das suas atividades. Codificado em dígitos binários e interpretável por um sistema computacional em suporte magnético, ótico ou outro.

De acordo com Rondinelli (2013, p. 235), as características do documento arquivístico digital, do ponto de vista da diplomática, incluem: forma fixa, conteúdo estável, relação orgânica, contexto identificável, ação e o envolvimento de cinco pessoas (autor, redator,

destinatário, originador e produtor). A forma fixa e o conteúdo estável significam que o documento arquivístico digital deve manter a apresentação que tinha quando foi “salvo” pela primeira vez (Rondinelli, 2013, p. 235).

A forma fixa e o conteúdo estável são conceitos fundamentais para garantir a autenticidade (identidade + integridade) dos documentos ao longo do tempo. Isso significa que uma vez que um documento é “salvo” ou arquivado, sua aparência original e seu conteúdo devem ser mantidos inalterados, permitindo que seja um registro fiel do momento de sua produção.

De acordo com o Glossário de Documentos Arquivísticos Digitais da Câmara Técnica de Documentos Eletrônicos (CTDE) do CONARQ (2020), um documento arquivístico digital é definido como “um documento digital reconhecido e tratado como um documento arquivístico” (CONARQ, 2020, p. 25).

Ao preservar documentos arquivísticos digitais, é primordial identificar e manter suas propriedades significativas, garantindo que as representações dos objetos digitais permaneçam compreensíveis e úteis para os usuários, mesmo diante das mudanças tecnológicas.

A convergência entre documentos arquivísticos digitais e propriedades significativas visa preservar a autenticidade, o acesso e a compreensão desses documentos ao longo do tempo. Logo, um documento deve manter sempre seu conteúdo e sua forma inalterados, a exemplo cita-se um documento textual com determinados elementos de forma e conteúdo deve continuar sendo apresentado como tal, assim como os documentos que contêm imagem, áudio, vídeo, entre outros, devem preservar suas respectivas apresentações e conteúdo informacional. Tais propriedades estão diretamente relacionadas ao conceito de objeto digital (Ferreira, 2006), e podem ser mais bem compreendidas por meio do Quadro 2, onde são apresentados conceitos de objetos.

Quadro 2 - Conceito de Objeto Digital

Ano	Autor	Conceitos
2004	Márdero Arellano	Objeto digital é algo que é criado em um computador, podendo ser um original ou uma versão digitalizada.
2006	Ferreira	Objeto digital é qualquer informação que pode ser expressa através de uma sequência de dígitos binários.
2020	CONARQ	Objeto digital é uma unidade de informação em formato digital, composta por uma ou mais sequências de bits e metadados que a identificam e descrevem suas características.

Fonte: A autora (2024).

Assim, é possível observar que os autores expressos no quadro acima tratam do objeto digital enfatizando a sua criação em um ambiente digital e sua codificação em dígitos binários.

Podemos notar uma ampliação do escopo, pois o conceito inicial de Mádero Arellano (2004) enfatiza a criação digital e a possibilidade de digitalização, mas não menciona sua estrutura técnica. Ferreira (2006) avança ao definir o objeto digital como qualquer informação representada por dígitos binários, ampliando a visão para qualquer dado digitalmente codificado.

Outro ponto é a inserção dos metadados a partir da definição do CONARQ, que reflete sua importância, além de assegurar a identificação e garantir autenticidade, acessibilidade e preservação do objeto digital ao longo do tempo.

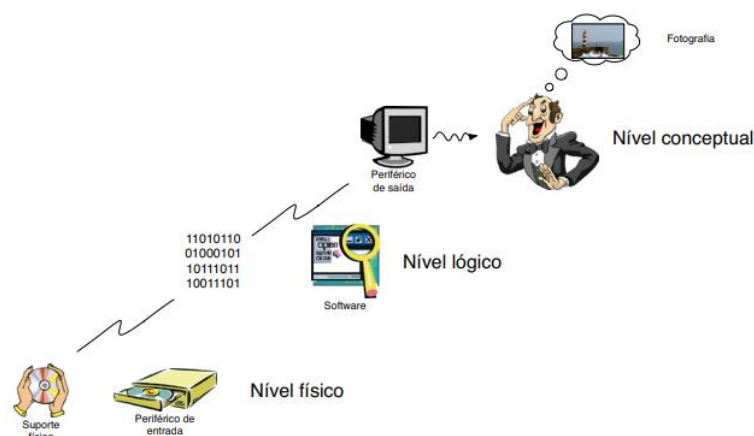
De acordo com o glossário CTDE, 2020, p36, a definição de metadados pode ser compreendido da seguinte forma “dados estruturados que descrevem e permitem encontrar, gerenciar, compreender e/ou preservar documentos arquivísticos ao longo do tempo.”

Já para Souto (2014), metadados são informações que descrevem características de um recurso. Eles desempenham várias funções, como localização, descoberta, documentação, avaliação e seleção. Metadados fornecem o contexto necessário para compreender os dados ao longo do tempo. São dados associados a objetos que ajudam os usuários a aproveitarem ao máximo o conhecimento sobre sua existência ou características. Metadados são essenciais para transformar dados brutos em conhecimento.

O Glossário da Câmara Técnica de Documentos Eletrônicos (CTDE) do CONARQ (2020) apresenta uma definição abrangente e detalhada dos objetos digitais, conceituando-os como unidades de informação em formato digital compostas por sequências de bits e acompanhadas de metadados que fornecem informações sobre suas propriedades e características. A definição destaca a complexidade e a estrutura multifacetada dos objetos digitais, considerando tanto os dados em si quanto os metadados relacionados, os quais fornecem contexto e significado ao documento digital.

Já Ferreira (2006, p. 24) aponta que o objeto digital apresenta níveis de abstração, como pode ser observado na Figura 1, apresentada a seguir.

Figura 1- Cadeia de Interpretação do nível físico ao conceitual



Fonte: Ferreira (2006).

De acordo com a imagem acima, Ferreira (2006) propõe uma distinção fundamental entre três níveis de abstração dos objetos digitais: físico, lógico e conceitual. O nível físico se refere à estrutura material dos objetos digitais, que envolve os suportes de armazenamento, bem como os dígitos binários registrados neles. O nível está diretamente relacionado à tecnologia de hardware, que é responsável pela gravação, leitura e preservação dos dados.

O nível lógico trata da organização e da representação dos dados de forma estruturada, que apresentam significado para a máquina. No nível, os dados são organizados segundo uma lógica de interpretação que permite seu processamento e seu uso dentro de um sistema específico. A preservação lógica busca manter a integridade e a funcionalidade dos dados, incluindo a manutenção de metadados que descrevem as propriedades e o contexto dos objetos digitais.

Por fim, o nível conceitual abrange o significado e a contextualização dos objetos digitais para os sujeitos, ou seja, o conteúdo dos dados, além de seu formato e sua estrutura. O nível lida com a interpretação semântica dos objetos digitais, assegurando que o valor original do conteúdo seja preservado.

A compreensão clara dos conceitos e definições relativos a objetos digitais é essencial para facilitar a verificação da autenticidade dos documentos. Além disso, a habilidade de rastrear a produção, as alterações e os acessos a um documento são fundamentais para assegurar que ele seja genuíno e não tenha sido alterado.

Duranti (2014), considera fundamental que os profissionais que lidam com documentos digitais tenham conhecimento sobre a estrutura e a composição dos objetos digitais, bem como sobre a importância dos metadados e da análise de autenticidade, pois isso se torna

especialmente relevante em contextos legais, administrativos e históricos nos quais a autenticidade e a exatidão dos registros são cruciais.

Conforme Indolfo (2013), os documentos arquivísticos são produzidos e/ou recebidos por entidades no exercício de suas atividades, formando um conjunto orgânico que reflete as funções e atos de seus produtores, podendo ser textuais, cartográficos, registros sonoros, imagens fixas ou em movimento, dependendo do tipo de informação e do suporte em que estão armazenados.

Esses documentos são fundamentais para a preservação da memória institucional e para a transparência administrativa, pois registram as atividades de entidades públicas e privadas, documentando fatos, decisões e refletindo as funções dessas instituições. Além disso, mostram-se essenciais para os pesquisadores, pois são fontes primárias de informações e atuam como ponte entre o passado e o futuro, fornecendo informações vitais para a tomada de decisões administrativas, legais e de pesquisa.

Conforme mencionado por Rondinelli (2013), os documentos arquivísticos desempenham um papel essencial na construção da identidade e da cultura organizacional, pois refletem os valores, as práticas e as contribuições de uma instituição ao longo do tempo. Produzidos no contexto de atividades institucionais, esses documentos são fundamentais para a administração pública, uma vez que possuem valor administrativo probatório e informativo, o que exige uma gestão adequada para garantir sua preservação e acessibilidade.

Neste contexto, a autenticidade dos documentos arquivísticos digitais torna-se um elemento central, pois está diretamente relacionada à credibilidade do objeto digital. Para compreender esse conceito, inicialmente, serão apresentados os diferentes entendimentos de autenticidade segundo distintos autores.

O conceito de autenticidade dos documentos arquivísticos digitais envolve uma interseção entre a Arquivologia e a Tecnologia, visto que ambas as áreas oferecem contribuições fundamentais para assegurar a confiabilidade e a integridade desses registros. Assim, compreender as propriedades significativas que garantem essa credibilidade é essencial para a gestão documental no ambiente digital.

Na Arquivologia, a autenticidade se refere à garantia de que os documentos são genuínos, íntegros e criados ou enviados pela pessoa ou entidade que alegadamente os produziu. Isso inclui assegurar que os documentos não foram alterados de maneira não autorizada ao longo do tempo. Na área de Tecnologia, a autenticidade envolve métodos e procedimentos técnicos para verificar e preservar as propriedades essenciais dos documentos digitais. Isso

inclui a utilização de metadados, certificação digital e outras técnicas que garantam que os documentos permanecem confiáveis, acessíveis e utilizáveis ao longo do tempo. Conforme o quadro 3, apresentamos algumas observações:

Quadro 3 - Conceito de Autenticidade

Fonte	Definição
Glossário do Tribunal de Justiça de Santa Catarina	autenticidade é um atributo essencial relacionado à segurança da informação, que garante a correspondência entre o criador de uma informação e a entidade, processo ou sistema a que se atribui a autoria.
Glossário de Segurança da Informação da Secretaria de Segurança da Informação e Cibernética, vinculada à Presidência da República pela Portaria nº 93/2021	autenticidade se refere à característica que garante que a informação foi gerada, enviada, alterada ou eliminada por uma pessoa física, equipamento, sistema, órgão ou entidade específica.
Resolução nº 37, de 19 de dezembro de 2012, do Conselho Nacional de Arquivos (CONARQ) aprova as Diretrizes para a Presunção de Autenticidade dos Documentos Arquivísticos Digitais	a autenticidade de um documento arquivístico digital é definida como a qualidade de um documento ser exatamente aquele que foi produzido, sem ter sofrido alterações, corrompimento ou adulteração
Glossário da CTDE	autenticidade é a credibilidade de um documento enquanto tal, ou seja, a qualidade que assegura que um documento é o que afirma ser e que está livre de adulteração ou qualquer forma de corrupção. A autenticidade é composta por dois elementos fundamentais: identidade e integridade

Fonte: A autora (2024).

Para analisar esses conceitos de autenticidade, é importante compará-los e entender como cada definição aborda a segurança e a credibilidade das informações. Vamos examinar cada uma delas.

Cada definição aborda a autenticidade de maneira distinta, refletindo diferentes aspectos e prioridades. Enquanto o TJSC se concentra na correspondência entre autor e informação, a Portaria nº 93/2021 abrange o ciclo de vida da informação e a responsabilidade dos agentes envolvidos. Por outro lado, o CTDE destaca a credibilidade e integridade do documento, considerando também o contexto em que ele está inserido.

Apesar das diferentes abordagens conceituais, todas convergem para a garantia da confiabilidade da informação. Na Segurança da Informação, o foco está na autoria e no controle de acessos. Por outro lado, na Arquivologia e Gestão Documental, a principal preocupação é manter a integridade e a confiabilidade do documento ao longo do tempo, o que é fundamental para sua função probatória e administrativa.

A identidade de um documento se refere à garantia de que este é exatamente o que afirma ser, estabelecendo uma relação clara entre o documento arquivístico e a pessoa ou entidade responsável pela sua produção. Ela diz respeito ao conjunto de características, metadados e atributos que a caracterizam como única e a distinguem de outros documentos

arquivísticos, como data, autor, destinatário, assunto, número identificador e número de protocolo, entre outros.

A integridade assegura que o documento arquivístico permaneça inalterado desde sua criação, preservando seu conteúdo original e garantindo que não houve modificações não autorizadas. Refere-se à capacidade do documento de transmitir exatamente a mensagem que levou à sua produção, sem alterações de forma e conteúdo, atingindo seus objetivos.

A integridade de um documento se refere ao fato de que ele é completo e não sofreu adulterações. No contexto não digital as alterações podem ser visíveis, como rasuras ou manchas. No ambiente digital, garantir a integridade envolve o uso de técnicas para assegurar que o documento não sofreu modificações não autorizadas. Isso pode incluir manter um registro de alterações, realizar verificações de logs e utilizar outros mecanismos para garantir que o conteúdo permaneça inalterado ao longo do tempo. Um documento íntegro é aquele que preserva todas as suas partes e propriedades originais, sem alterações não autorizadas.

Conforme as diretrizes do produtor a elaboração e a manutenção de materiais digitais: diretrizes para indivíduos, esses conceitos são fundamentais para assegurar a autenticidade e a confiabilidade dos documentos digitais. A identidade permite a identificação e diferenciação dos documentos, enquanto a integridade garante que os documentos permanecem intactos e completos ao longo do tempo.

3 PRESERVAÇÃO DIGITAL: CONCEITO, DESAFIOS E ESTRATÉGIAS FRENTE AOS DOCUMENTOS ARQUIVÍSTICOS

A preservação digital é um conjunto de atividades e processos que garantem o acesso contínuo e a longo prazo a materiais digitais. Inclui a manutenção da autenticidade dos documentos, mesmo com a evolução tecnológica e a obsolescência dos formatos e suportes. A importância da preservação digital é fundamental, considerando a enorme quantidade de informação produzida e consumida digitalmente. Sem medidas adequadas, muitos dados podem se perder ou se tornar inacessíveis com o tempo, conforme mencionado por Márdero Arellano (2004)

Ao cotejar esses conceitos, é possível entender a importância da autenticidade para que um documento possa ser utilizado como fonte de informação, bem como a necessidade de procedimentos e sistemas que assegurem que os documentos sejam mantidos de forma genuína e confiável ao longo do tempo. A manutenção da autenticidade dos documentos arquivísticos digitais representa um desafio complexo e variável que envolve a preservação de suas propriedades significativas ao longo do tempo. Isso requer uma abordagem abrangente que leve em consideração não apenas a forma e o conteúdo dos documentos, mas também o ambiente em que são produzidos, mantidos e utilizados.

Cada vez que se realiza o envio de um documento ou um recolhimento, há um risco de perda da sua autenticidade. Segundo a Diretriz do Conselho Nacional de Arquivos, a autenticidade dos documentos arquivísticos digitais é ameaçada sempre que eles são transmitidos através do espaço (entre pessoas, sistemas ou aplicativos) ou do tempo (armazenagem contínua ou atualização/substituição de *hardware/software* usados para armazenar, processar e comunicar os documentos) (CONARQ, 2012).

Nesse contexto, esta pesquisa foca em como definir parâmetros para a presunção de autenticidade desses documentos arquivísticos digitais. Sendo assim, a autenticidade dos documentos arquivísticos digitais é assegurada pela combinação de práticas arquivísticas e técnicas tecnológicas, sendo as propriedades significativas um elemento-chave para manter essa credibilidade. Vamos explorar esses conceitos em relação à preservação e à autenticidade de documentos arquivísticos digitais.

Para Lacombe e Silva (2007) a autenticidade de documentos arquivísticos digitais não abrange apenas a identidade e a integridade. A preservação adequada e contínua é essencial para manter a autenticidade dos documentos. Isso envolve confiar nas entidades que mantêm a

custódia dos documentos e aplicar técnicas que garantem que os documentos permaneçam completos e inalterados.

Esses conceitos são essenciais para entender a preservação e a autenticidade de documentos arquivísticos digitais e garantir que eles permaneçam confiáveis e utilizáveis ao longo do tempo.

Conforme defende Ferreira (2006) a preservação digital exige o uso de diferentes estratégias, como, por exemplo, emulação, migração e encapsulamento, para evitar a perda de documentos devido à obsolescência tecnológica.

De acordo com Sayão (2010), é fundamental definir metadados robustos e padronizados, garantindo que os documentos digitais possam ser localizados, acessados e compreendidos no futuro. Esses metadados são essenciais, pois registram informações sobre o contexto, o conteúdo e a estrutura dos documentos, contribuindo para a sua organização, recuperação e conservação ao longo do tempo.

Uma das estratégias de preservação digital envolve a utilização de padrões de metadados, que são estruturas formais para descrever recursos de comunidades específicas. Esses padrões ajudam na gestão e preservação de objetos digitais em ambientes de informação, como repositórios institucionais (Formenton *et al.*, 2017)

De acordo com Gava e Flores (2021), com a crescente produção de documentos arquivísticos nato-digitais e a rápida evolução tecnológica, é necessário assegurar que esses documentos possam ser acessados a longo prazo. Isso requer um compromisso da comunidade envolvida na produção, gestão e guarda desses documentos, além de procedimentos adequados para a manutenção e a preservação dos documentos nato-digitais.

Esse compromisso envolve a comunidade que produz, gerencia e guarda os documentos arquivísticos nato-digitais. Para Besser (2010 *apud* Tavares, 2012), o reconhecimento do compromisso por parte da sociedade é que fará a manutenção dessa preservação, uma vez que “As próximas gerações necessitarão para o reconhecimento do seu passado do compromisso das sociedades atuais com a preservação digital [...]”.

No contexto da preservação de documentos, são aplicadas técnicas específicas para mitigar os impactos da degradação tecnológica. Essas abordagens podem alterar tanto as características quanto a forma de apresentação dos documentos, tornando indispensável o registro dessas modificações para assegurar sua autenticidade. Para esse fim, são utilizados padrões técnicos que viabilizam o registro detalhado dos eventos relacionados ao processo de preservação.

As políticas, as normas e os regramentos são essenciais para garantir a longevidade de acesso aos documentos arquivísticos digitais. A responsabilidade recai sobre o Poder Público, instituições custodiadoras e profissionais, exigindo vigilância constante e adaptação às novas tecnologias. A colaboração entre entidades custodiadoras e profissionais é fundamental para assegurar a continuidade das práticas de preservação, conforme mencionado por Gava e Flores (2021).

A conscientização sobre a importância da preservação digital deve ser promovida desde o início do ciclo de vida dos documentos, garantindo que as práticas de preservação sejam incorporadas desde a produção até a destinação final. A formação contínua dos profissionais é fundamental, pois as tecnologias e os desafios associados à preservação digital estão em constante evolução.

Instituições públicas têm um papel vital na conscientização e desenvolvimento de políticas eficazes para a preservação digital. Elas são responsáveis por criar normas, leis e recomendações que assegurem a longevidade dos documentos digitais. Arquivistas e profissionais da informação atuam como intermediários no ciclo de vida dos documentos digitais, aplicando práticas de preservação de forma consistente.

A Lei nº 8.159, de 08 de janeiro de 1991¹ estabelece a responsabilidade do poder público na gestão documental e proteção especial a documentos de arquivo, independentemente do suporte ou natureza (Brasil, 1991). A legislação já previa a preocupação com a manutenção de suportes diversos, incluindo documentos digitais. A gestão de documentos, conforme definida na lei, envolve um conjunto de procedimentos e operações técnicas contínuas e evolutivas, requerendo colaboração entre profissionais e entidades para garantir a longevidade dos documentos.

A Carta para a Preservação do Patrimônio Arquivístico Digital do CONARQ, de 2004, destaca a importância de preservar documentos digitais como patrimônio cultural e histórico, enfatizando a necessidade de políticas e ações para garantir autenticidade, confiabilidade e acessibilidade dos documentos digitais ao longo do tempo (Conselho Nacional de Arquivos, 2004).

O CONARQ (2020) define a preservação digital como um conjunto de ações gerenciais e técnicas para superar mudanças tecnológicas e a fragilidade dos suportes, garantindo acesso e interpretação dos documentos digitais pelo tempo necessário. É reconhecido que existem

¹ Lei nº 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8159.htm. Acesso em: 19 mar. 2025.

grandes desafios a serem enfrentados para alcançar a preservação de documentos nato-digitais.

No ambiente digital, esses documentos estão sujeitos tanto à deterioração quanto à obsolescência tecnológica. Para Innarelli (2015), preservá-los exige estratégias que assegurem o acesso contínuo ao seu conteúdo e às suas funcionalidades, utilizando os recursos tecnológicos disponíveis. Para isso, são empregadas técnicas que garantem a acessibilidade e a compreensão dos documentos ao longo do tempo, preservando sua autenticidade e usabilidade mesmo diante das constantes evoluções tecnológicas.

A preservação de documentos arquivísticos digitais é uma responsabilidade coletiva que requer atenção contínua, investimento em tecnologias e infraestruturas, e uma abordagem colaborativa entre arquivistas, tecnólogos, legisladores e sociedade. Somente assim o legado digital da humanidade será preservado para o benefício de todos.

A preservação digital é essencial devido à vulnerabilidade dos documentos digitais e à rápida mudança tecnológica, que apresentam desafios intrincados. Os principais obstáculos incluem a obsolescência tecnológica, a fragilidade dos dispositivos de armazenamento, questões de segurança, o risco de alteração ou corrupção das informações, custos, necessidade de garantir integridade e autenticidade dos dados, gestão de metadados, armazenamento por longos períodos e acesso mediado por hardware e software.

Santos e Flores (2015) aborda que com o avanço acelerado da tecnologia pode tornar formatos de arquivos e sistemas operacionais obsoletos rapidamente, exigindo medidas preventivas na gestão de documentos digitais. Para garantir sua autenticidade e confiabilidade ao longo do tempo, é essencial adotar metodologias adequadas. Além disso, esses documentos dependem de equipamentos e softwares que evoluem constantemente, tornando sua administração cada vez mais complexa.

A obsolescência tecnológica não afeta apenas os formatos de arquivos, mas também os meios de armazenamento, dificultando o acesso a informações importantes.

De acordo com Innarelli (2015) diante desse cenário, é fundamental que instituições elaborem estratégias de contingência para acompanhar as mudanças tecnológicas e reduzir os riscos de perda de dados essenciais. A dependência de softwares proprietários e a rápida evolução dos sistemas aumentam as chances de inacessibilidade, tornando a preservação digital um desafio contínuo.

No entendimento de Sayão (2010) a inclusão de metadados detalhados desempenha um papel essencial ao fornecer informações sobre o contexto, o conteúdo e a estrutura dos documentos, facilitando sua gestão, recuperação e conservação. O uso de formatos

padronizados e softwares de código aberto contribui para a interoperabilidade e a longevidade dos arquivos digitais.

Para Ferreira (2006) as estratégias eficazes de preservação devem incluir migração de dados, conversão para formatos sustentáveis e práticas de redundância, como backups regulares e armazenamento em locais distintos.

Diante desses desafios, profissionais e gestores da informação precisam atuar de forma proativa para garantir a preservação e o acesso contínuo aos documentos digitais. A implementação de políticas eficazes, investimentos em infraestrutura tecnológica e uma abordagem colaborativa são indispensáveis para assegurar a continuidade do patrimônio digital, beneficiando tanto a sociedade atual quanto as futuras gerações.

Para lidar com essas questões, diversas estratégias de preservação digital têm sido implementadas, como migração, preservação da tecnologia, emulação e encapsulamento, conforme destacado por Márdero Arellano (2008, p. 22).

O problema da preservação digital está no conteúdo dos objetos digitais, nas informações armazenadas e na maneira como foram armazenadas. Não são apenas sequências de zeros e uns ou cadeias de bits (bitstream)¹, mas o conjunto de decisões que definiram a formação básica do objeto informacional como um objeto único. Isso envolve a descrição de documentos que podem ser representados de diferentes formas, mas o seu conteúdo será sempre interpretado do mesmo modo, no momento da sua recuperação, ou seja, será a mesma cadeia de bits com seu significado original.

De acordo com Santos e Flores (2015, a adoção de formatos de arquivo de código aberto e a implementação de políticas de preservação digital são essenciais para manter a acessibilidade e a usabilidade dos documentos digitais no futuro. As ações de preservação devem ser consistentes e confiáveis, garantindo a autenticidade e a integridade dos documentos, mesmo diante das rápidas transformações tecnológicas e da obsolescência.

Ainda para Márdero Arellano (2004). medidas proativas são indispensáveis para minimizar esses riscos, incluindo a conversão de documentos para formatos modernos e aceitos, o armazenamento em repositórios seguros e o uso de metadados adequados que documentem informações detalhadas sobre os arquivos. Profissionais da área de informação devem se manter atualizados sobre inovações tecnológicas e revisar suas estratégias de preservação digital, conforme necessário.

Apesar da complexidade da conservação de documentos arquivísticos digitais, a combinação de estratégias de preservação consiste em várias etapas que devem ser seguidas com responsabilidade. A migração de dados, que envolve a transferência de informações de um sistema de armazenamento para outro, geralmente exige a conversão de formatos de arquivo para garantir a compatibilidade com novas plataformas tecnológicas. Essa prática pode

impactar as propriedades significativas dos documentos, que asseguram sua autenticidade e relevância.

Ferreira (2006) descreve a migração de dados como o processo de mover informações de um sistema de armazenamento para outro, frequentemente exigindo a conversão de formatos para garantir compatibilidade com novas tecnologias. Esse processo pode impactar as características fundamentais dos documentos, que garantem sua autenticidade e pertinência.

Santos e flores (2015), sustenta que a migração digital é uma estratégia para garantir a preservação de documentos e arquivos ao longo do tempo. Ela envolve não apenas a transferência de dados, mas também a proteção da integridade estrutural e do conteúdo informativo dos materiais originais, assegurando que futuras tecnologias possam acessar e utilizar essas informações sem perda de dados ou contexto.

A conversão de formatos é relevante em um contexto de constante mudança tecnológica. Padronizar formatos promove a interoperabilidade entre diversos sistemas e plataformas, permitindo fácil acesso e uso dos dados por uma variedade de aplicativos e usuários. A conversão de formatos é uma maneira eficaz de mitigar a obsolescência tecnológica, garantindo que as informações permaneçam acessíveis mesmo quando as tecnologias originais se tornam obsoletas.

Para enfrentar os desafios da preservação de documentos digitais, é essencial implementar estratégias como a conversão de formatos e o armazenamento em repositórios confiáveis, garantindo a preservação em um ambiente seguro e monitorado continuamente. Segundo o glossário CTDE (2020), a conversão de formatos se refere à alteração de um formato para outro, com o principal objetivo de normalizar os formatos e evitar a obsolescência tecnológica.

Neste sentido, Ferreira (2006), menciona que a atualização de suporte, ou refrescamento, é uma técnica de migração que transfere dados de um suporte para outro sem alterar sua codificação, minimizando a perda de dados decorrente da degradação do suporte e acompanhando as inovações tecnológicas. A emulação cria um ambiente de software que replica o hardware antigo, possibilitando o uso de documentos como se estivessem na plataforma original. O encapsulamento envolve a adição de metadados e informações contextuais aos documentos digitais, garantindo que possam ser corretamente interpretados no futuro.

A utilização de metadados adequados é crucial para registrar informações abrangentes sobre os documentos, incluindo formatos, *softwares* e contextos de criação. Márdero Arellano

(2008) ressalta a importância de estratégias de preservação digital que englobem todos os aspectos associados ao problema tecnológico, como migração, emulação, arqueologia digital, criptografia, metadados, formatos padrão e *softwares* livres.

Essas estratégias são fundamentais para garantir que os documentos permaneçam acessíveis a longo prazo, protegendo-os da obsolescência tecnológica, perda de dados e deterioração, assegurando que continuem legíveis e utilizáveis para futuras gerações. É essencial que as abordagens de preservação levem em conta não apenas a conservação dos dados, mas também a manutenção do contexto e da interpretação, garantindo que o conteúdo permaneça acessível e relevante.

Conforme mencionado por Rocco (2021, p. 82), “[...] o procedimento de ‘migração’ atua como um conceito abrangente, englobando tanto a atualização de suportes (refrescamento) quanto a conversão de formatos”.

A migração é a estratégia de preservação digital mais utilizada em maior escala; é também a que possui um conjunto de procedimentos mais bem organizados e consolidados dentre as estratégias correntemente praticadas [...] O objetivo primordial da migração é preservar a integridade dos objetos digitais enquanto mantém – prioritariamente – a capacidade do usuário de recuperá-los (Sayão, 2005, p. 133).

A referência feita por Sayão (2005) destaca que a migração se posiciona como a estratégia de preservação digital mais comum e eficaz entre as diversas opções disponíveis. Essa técnica consiste na transição de dados de um formato digital que se tornou obsoleto para um novo formato que seja mais atual ou sustentável. O objetivo primordial dessa abordagem é assegurar que os objetos digitais preservem sua integridade ao longo do tempo. Além disso, busca-se garantir que os usuários possam acessar e recuperar esses objetos sempre que necessário.

A combinação de uma ou mais estratégias assegura que os documentos digitais permaneçam íntegros, autênticos e acessíveis ao longo do tempo. São fundamentais para uma preservação digital eficaz. Essas estratégias são indispensáveis para assegurar a continuidade e a segurança dos documentos digitais, devendo ser implementadas de forma integrada para evitar a perda de documentos arquivísticos digitais devido a falhas nos sistemas de armazenamento ou incompatibilidades tecnológicas.

A atualização dos arquivos digitais é essencial para garantir a integridade e acessibilidade dos documentos ao longo do tempo. Esse processo envolve a atualização periódica dos arquivos para prevenir obsolescência e degradação, mantendo os dados disponíveis e utilizáveis em um ambiente digital em constante mudança. Devido às mudanças tecnológicas, os arquivos digitais podem sofrer corrupção, perda de metadados ou danos.

Portanto, a atualização assegura que os arquivos permaneçam acessíveis e legíveis.

A prática de atualizar os suportes é essencial para evitar obsolescência e degradação, garantindo a acessibilidade e integridade dos dados. Isso envolve transferir os bits de um suporte físico para outro, reduzindo o risco de perda de dados devido à deterioração do suporte original. Essa medida é fundamental para a preservação a longo prazo, especialmente em um cenário tecnológico em constante evolução. Como defende Ferreira:

A migração tem como objectivo manter os objectos digitais compatíveis com tecnologias actuais de modo a que um utilizador comum seja capaz de os interpretar sem necessidade de recorrer a artefactos menos convencionais, como por exemplo, emuladores. Ferreira, p36, 2006

A migração de dados, que inclui a conversão de formatos e a atualização de suporte, é uma estratégia essencial na preservação de informações digitais. Essa abordagem capacita as organizações a manterem a continuidade em seus processos de negócios e a acessibilidade dos dados, enfrentando os desafios trazidos pelas rápidas mudanças tecnológicas e pelas necessidades em constante transformação dos sistemas de informação.

Optar por essa estratégia não significa que outras abordagens não sejam relevantes para esta pesquisa; no entanto, para manter um foco claro no problema, a análise dessa estratégia está alinhada com os objetivos do estudo. Conforme Márdero Arellano (2008, p. 22) destaca, a estratégia de migração envolve complexidades que merecem atenção, como será discutido a seguir.

A maioria dos bibliotecários e arquivistas aceita que a preservação digital depende, de modo geral, da migração de material digital que permite que ele seja compatível com novas tecnologias. Mas a migração é muito mais complexa do que apenas transferir o bitstream de uma mídia para outra. A estrutura interna e o conteúdo do material devem ser preservados e transferidos igualmente, para que, dessa forma, o “novo” objeto seja uma representação fiel do original.

Durante o processo de migração, é crucial preservar as características principais dos documentos, incluindo sua identidade e integridade. Mesmo com a mudança de formato, os documentos devem refletir com precisão seu conteúdo original e estar livres de alterações inadequadas. A conversão de formatos exige a atualização dos metadados associados, garantindo que todas as informações necessárias para sua interpretação e uso permaneçam acessíveis e corretas após a migração.

Esse procedimento é vital para manter a acessibilidade dos documentos arquivísticos digitais ao longo do tempo, mesmo com as transformações tecnológicas. A responsabilidade pela preservação desses documentos cabe às instituições que os guardam, que devem realizar a conversão de formatos de forma segura e eficiente. Implementar políticas e práticas de

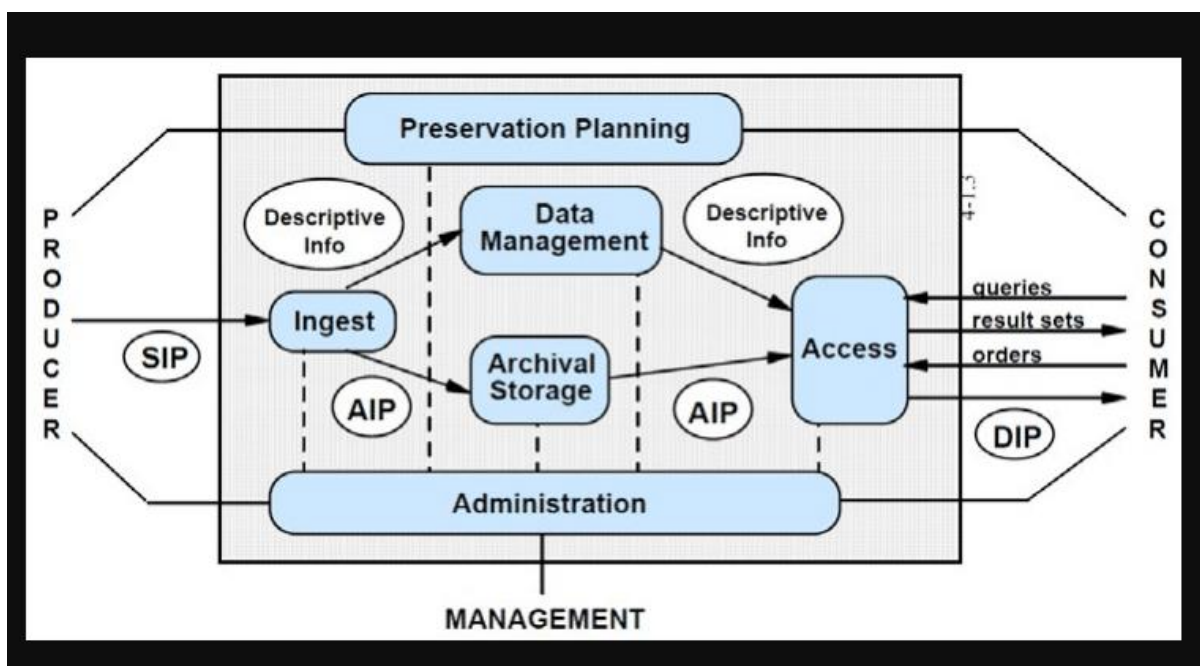
preservação é essencial para minimizar os riscos de perda de dados ou comprometimento da autenticidade dos documentos.

O Modelo de Referência *Open Archival Information System* (OAIS) é uma estrutura conceitual amplamente reconhecida para a preservação de informações digitais. Ele estabelece diretrizes para o gerenciamento e a preservação dos arquivos digitais, garantindo sua durabilidade e acessibilidade ao longo do tempo. O OAIS envolve seis funções essenciais: ingestão, armazenamento, gerenciamento de dados, administração, acesso e preservação, proporcionando um framework sistemático que orienta as práticas de preservação digital.

Desenvolvido pelo *Consultative Committee for Space Data Systems* (CCSDS), o Modelo OAIS se consolidou como um padrão ISO (ISO 14721:2012)². Ele descreve as funções e responsabilidades necessárias para garantir que os objetos digitais sejam preservados de maneira autêntica, íntegra e acessível ao longo do tempo. Além disso, oferece uma base conceitual que orienta a implementação de um ambiente para a preservação digital, definindo funcionalidades, um modelo de dados e os papéis e deveres para assegurar a adequada preservação dos objetos digitais.

² O Modelo OAIS (*Open Archival Information System*), desenvolvido pelo *Consultative Committee for Space Data Systems* (CCSDS), é um padrão internacional (ISO 14721:2012) que define um framework para a preservação e o acesso contínuo de informações digitais a longo prazo. Seus principais objetivos são garantir a integridade e acessibilidade das informações digitais ao longo do tempo, fornece um modelo comum para diferentes organizações desenvolverem sistemas de preservação digital, e assegurar que as informações permaneçam acessíveis, mesmo com mudanças tecnológicas.

Figura 2 - Modelo Funcional OAIS



Fonte: Consultative Committee for Space Data Systems (2002).

No modelo funcional acima, está representado o sistema de preservação composto por:

- a) Produtor: a entidade responsável por submeter informações ao repositório. Pode ser uma pessoa, uma organização ou um sistema.
- b) Consumidor: a entidade que acessa e utiliza as informações armazenadas no repositório. Pode ser um usuário final, uma aplicação ou outro sistema.
- c) Gestão: entidade responsável por planejar e desenvolver políticas.
- d) *Submission Information Package* (SIP): pacote de informação com os documentos e as informações disponíveis para recolhimento.
- e) *Archival Information Package* (AIP): pacote de informação com todas as informações relevantes para armazenamento e preservação de objeto digital.
- f) *Dissemination Information Package* (DIP): pacote de informação com os dados disponíveis para acesso.
- g) Ingestão: processos que aceitam materiais digitais dos produtores e os preparam para a inclusão no Armazenamento de Arquivos.
- h) Armazenamento de Arquivos: armazenamento e manutenção de longo prazo de materiais digitais, garantindo que os materiais digitais arquivados estejam na forma correta

de armazenamento e que os fluxos de bits permaneçam completos e renderizáveis ao longo do tempo.

i) Gerenciamento de Dados: manutenção de bancos de dados de metadados descritivos que identificam e descrevem os materiais digitais arquivados, além de gerenciar dados administrativos que dão suporte à operação do sistema OAIS, como dados de desempenho ou estatísticas de acesso.

j) Acesso: processos e serviços que os consumidores usam para localizar, solicitar e receber informações de materiais digitais na Loja de Arquivos.

k) Administração: gestão diária das operações e coordenação das outras entidades funcionais. Também é o hub central de comunicações com o ambiente externo: gestão, produtores e consumidores.

l) Planejamento de Preservação: monitoramento do ambiente externo para mudanças ou ameaças aos materiais digitais no Armazenamento de Arquivos. Mapeia a estratégia de preservação e recomenda revisões apropriadas em linha com as condições de mudança, garantindo a salvaguarda contra um ambiente de usuário e tecnológico em constante evolução.

A partir do que foi debatido acerca da preservação, os documentos que são o foco desta pesquisa são encaminhados para o repositório do Arquivo Nacional, tema que vamos abordar a seguir.

4 ARQUIVO NACIONAL: EM FOCO SEU REPOSITÓRIO

Para conceituar a relevância da pesquisa, apresentamos um breve histórico da entidade custodiadora responsável pela custódia do acervo. O Arquivo Nacional foi criado em 1838 e hoje é um órgão central do Sistema de Gestão de Documentos e Arquivos (Siga) da administração pública federal, do Ministério da Gestão e da Inovação em Serviços Públicos. Sua missão é implementar e acompanhar a política nacional de arquivos, definida pelo CONARQ, por meio da gestão, recolhimento, tratamento técnico, preservação e divulgação do patrimônio documental do país. Isso garante acesso à informação, apoia decisões governamentais, auxilia o cidadão na defesa de seus direitos e incentiva a produção de conhecimento científico e cultural.

O Arquivo Nacional desempenha um papel fundamental na manutenção da autenticidade dos documentos, garantindo que sejam preservados, catalogados e acessíveis. Ele assegura a preservação física e digital dos documentos, mantendo suas características essenciais de procedência e integridade. Com a missão de garantir à sociedade o direito à informação e à memória do país, o Arquivo Nacional contribui para a inovação e o fortalecimento do Estado.

Desde sua criação, o Arquivo Nacional tem desempenhado um papel crucial na preservação e na divulgação do patrimônio documental brasileiro. Ele coleta, trata tecnicamente e preserva documentos de valor histórico, garantindo que a memória do Brasil seja acessível para futuras gerações.

O Arquivo Nacional, hoje assim denominado, teve sua origem em 1838. Previsto na Constituição de 1824 e sob a denominação de Arquivo Público do Império, estabeleceu-se na Secretaria dos Negócios do Império pelo Regulamento nº 2, de 2 de janeiro de 1838, com a finalidade de guardar os documentos públicos. Naquele momento, os documentos sob sua guarda eram relativos à família imperial, às relações exteriores, documentos eclesiásticos e diplomas legais dos poderes legislativo, executivo e judiciário.

O Arquivo Nacional tem como objetivo preservar os documentos de valor administrativo ou histórico oriundos dos órgãos da União e entidades de direito privado por ela instituídas, bem como documentos históricos provenientes de entidades públicas e particulares. Com o propósito de atingir sua missão, o Arquivo Nacional instituiu em 2010 o Programa Permanente de Preservação e Acesso a Documentos Arquivísticos Digitais (AN Digital). A partir desse programa, o grupo de trabalho sugeriu a criação da Divisão de Preservação de

Documentos Digitais no Arquivo Nacional. Uma vez instituída, essa divisão tem como objetivo propor, planejar e executar ações relacionadas à preservação e ao acesso a objetos digitais.

No Arquivo Nacional, após testar algumas possibilidades disponíveis à época, optou-se por utilizar o Archivematica como principal software para apoiar a implantação do repositório digital.

O Archivematica é um sistema de preservação digital de código aberto desenvolvido pela Artefactual Systems, que apoia a preservação e o acesso de longo prazo a conteúdos digitais de forma confiável, autêntica e segura. O sistema utiliza um conjunto integrado de ferramentas de software para processar objetos digitais desde a ingestão até o acesso, em conformidade com o modelo funcional OAIS. Usuários podem monitorar e controlar a ingestão e a preservação através de um painel.

O Archivematica foi desenvolvido com base nos padrões e normas internacionalmente reconhecidos, como METS, PREMIS, Dublin Core e a especificação BagIt, da *Library of Congress*, para gerar Pacotes de Informações de Arquivamento (AIPs³) confiáveis e independentes do sistema. O Archivematica implementa funcionalidades previstas no modelo funções do OAIS, como a criação, manutenção, disseminação de pacotes de informação e gerenciamento da preservação, bem como permite o acesso aos pacotes por meio de integração com o AtoM⁴.

O Modelo OAIS apresenta três tipos de pacotes de informação: *Submission Information Package* (SIP), pacote de informação para submissão; *Archival Information Package* (AIP), pacote de informação para arquivamento; e *Dissemination Information Package* (DIP), pacote de informação para disseminação. Após a preparação do pacote SIP pelo produtor, este é submetido ao repositório através da função “Admissão”. Em seguida, é produzido o pacote AIP que contém os metadados descritivos, os metadados de controle, os objetos digitais submetidos ao repositório e as cópias desses objetos, normalizadas para o formato de preservação, identificadas por um identificador permanente (UUID).

Os pacotes AIP são armazenados em servidor e gerenciados pelas funcionalidades de “Armazenamento”. Dentro do pacote AIP não podem ser manipulados nem inseridos novos documentos ou metadados, este é padronizado conforme o plano de preservação do Arquivo, que prevê para qual extensão o formato será migrado para um formato de preservação. As funcionalidades de acesso produzem o DIP, que é apresentado ao consumidor com documentos

³ *Archival Information Package*.

⁴ O AtoM (*Access to Memory*) é uma ferramenta de software de código aberto, projetada para facilitar a descrição, gestão e acesso a arquivos e coleções arquivísticas.

digitais no formato definido para acesso. Conforme documentação do Archivematica Artefactual Systems Inc.

Nesse sentido, tal divisão logrou seu primeiro recolhimento nato-digital em 2015 o recolhimento do acervo da CNV ao Arquivo Nacional, ação fundamental para a preservação da memória histórica, promoção da verdade e proteção dos direitos humanos no Brasil. Quanto aos documentos oriundos da Comissão Nacional da Verdade, apresentaremos a seguir um breve resumo do propósito dessa comissão.

A Comissão Nacional da Verdade foi criada pela Lei nº 12.528/2011 como um órgão temporário, com o objetivo de examinar e esclarecer as graves violações de direitos humanos praticadas no Brasil entre setembro de 1946 e outubro de 1988. Todo o acervo documental resultante do trabalho da CNC foi encaminhado ao Arquivo Nacional e suas atividades foram finalizadas com a entrega do Relatório Final em dezembro de 2014.

Em dezembro de 2014, foi criada a Coordenação-Geral da Estrutura Temporária para Organização do Acervo da CNV, via realocação de cargos no âmbito da Casa Civil da Presidência da República (Decreto nº 8.378). Essa coordenação foi responsável por organizar e transferir o acervo ao Arquivo Nacional para sua integração ao Centro de Referência das Lutas Políticas no Brasil (1964-1985) – Memórias Reveladas, conforme disposto no parágrafo único do art. 11 da Lei nº 12.528/2011.

Desde então, o Arquivo Nacional mantém todo o acervo acessível para consultas locais e remotas, conforme requisitado pela CNV. O acervo é composto por documentos textuais e audiovisuais, além de dados e informações recebidos de diversas fontes, como instituições, universidades e pessoas físicas.

O recolhimento do acervo ocorreu em três etapas:

1. Primeira etapa: recolhimento de 2496 itens registrados e arquivados em 187 caixas-arquivo, acompanhados de memorial descritivo, 2 HDs externos com cópia dos documentos e 221 documentos digitais, totalizando 2717 itens.
2. Segunda etapa: recolhimento de 131 caixas com 1874 itens registrados, incluindo cópias de ofícios e documentos expedidos pela CNV, que podem conter dados pessoais e registros de audiências públicas e depoimentos.
3. Terceira etapa: entrega de 2 HDs externos contendo cópias digitais dos itens recolhidos nas 131 caixas e mais 42 itens digitais, sem correspondentes físicos. A tabela dos documentos recolhidos na segunda e na terceira etapas foi incluída na documentação de recolhimento.

A equipe do Arquivo Nacional está concluindo o arranjo provisório do fundo, cadastrando-o no Diretório Brasil Arquivos (Dibrarq) e descrevendo-o no Sistema de Informações do Arquivo Nacional (Sian).

5 PROJETO INSPECT: A DEFINIÇÃO DE PROPRIEDADES SIGNIFICATIVAS

O *Investigating the Significant Properties of Electronic Content over Time* (Projeto Inspect) foi uma iniciativa financiada pelo *Joint Information Systems Committee* (JISC) do Reino Unido, que se desenvolveu entre outubro de 2006 e dezembro de 2009, e sua atuação, atualmente está encerrada. Sob a liderança do *King's College London*, o projeto teve a colaboração do *The National Archives*. A meta central do Projeto Inspect era examinar as propriedades significativas de objetos digitais ao longo do tempo, assegurando a preservação dessas características para manter o conteúdo acessível e relevante. A pesquisa abarcou diversas categorias de conteúdo digital, incluindo áudio, e-mails, imagens matriciais e textos estruturados. O projeto visava investigar e determinar quais propriedades dos objetos digitais precisam ser resguardadas ao longo do tempo. O projeto Inspect define propriedades significativas como aqueles aspectos de um registro digital que devem ser preservados ao longo do tempo para que o objeto de informação permaneça autêntico.

Entender o objeto digital é fundamental para apoiar a garantia da autenticidade dos documentos na era digital, que pode ser validada por meio de *softwares* que empregam critérios estabelecidos pelo Projeto Inspect. Esse projeto delineou o conceito de propriedades significativas, o qual é essencial para a presunção de autenticidade dos documentos digitais.

Em suma, a consideração das propriedades significativas permite avaliar a autenticidade dos documentos digitais de forma mais precisa e confiável. Além disso, a utilização de critérios estabelecidos pelo Projeto Inspect garante uma abordagem sistemática e rigorosa para a validação da autenticidade.

A conceituação de propriedades significativas é apresentada com base no Projeto Inspect (2009), que caracteriza tais propriedades como elementos essenciais dos objetos digitais que devem ser preservados ao longo do tempo para garantir sua acessibilidade, usabilidade e significado, além de serem reconhecidos como evidência fidedigna do que pretendem registrar.

Conforme salientado por Hedstrom e Lee (2002), as propriedades relevantes de um objeto digital são aquelas que influenciam sua qualidade, usabilidade, visualização e comportamento. Ferreira (2006) aponta que as propriedades significativas se comportam como essência fundamental de um objeto digital, enfatizando que essa compreensão é essencial, pois determina o conjunto de características que devem ser preservadas de forma integral para que o objeto seja considerado autêntico, ou seja, continue sendo uma representação precisa e confiável da informação original ao longo do tempo.

Ferreira (2006) salienta, ainda, que esse conjunto de propriedades significativas não é rígido nem permanente, devendo levar em conta a natureza da organização responsável pela preservação, as particularidades da coleção e, sobretudo, as necessidades e exigências da comunidade envolvida.

Dessa forma, conforme Ferreira (2006) a essência de um objeto digital determina a maneira que o material deve ser preservado, assim como as propriedades significativas influenciam diretamente as estratégias de preservação.

Conforme o número de propriedades relevantes se torna mais claro, aumentam também as demandas por uma infraestrutura tecnológica adequada que assegure a preservação. Embora seja importante, a definição de um conjunto particular de propriedades significativas para cada item digital de um objeto digital não deve ser uma tarefa desvinculada das decisões da entidade custodiadora, uma vez que esta precisa se manter atualizada em termos tecnológicos.

É, portanto, essencial a elaboração de políticas de preservação que definam, para cada categoria de objetos digitais, quais propriedades significativas serão garantidas pelo repositório. Essas propriedades significativas são características fundamentais dos dados que possuem relevância direta para o seu contexto ou uso. Elas oferecem informações essenciais sobre os dados, sendo indispensáveis para o funcionamento do sistema, uma vez que possibilitam aos usuários a pesquisa, a filtragem e a tomada de decisões com base nessas informações. Tais propriedades são imprescindíveis para a eficiência dos sistemas e das aplicações que dependem de informações específicas.

Dada a sua relevância, as características essenciais de um objeto digital que compõem o documento arquivístico devem ser preservadas para assegurar uma alta presunção de autenticidade, nutrindo tal documento dos elementos fundamentais para que ele possa servir de prova e atestar a ação que ele registra e/ou apoia.

Realizado entre 2006 e 2009, o projeto Inspect foi uma experiência pioneira na preservação de documentos arquivísticos digitais, criando uma metodologia para analisar a manutenção das propriedades significativas ao longo do tempo. A presente pesquisa pretende oferecer uma compreensão do conceito de propriedades significativas, fundamentada no Projeto Inspect, e contribuir para que as instituições responsáveis pela preservação de acervos digitais compreendam seu papel na tomada de decisões sobre os parâmetros aceitáveis na variação dessas propriedades, em vista da manutenção a longo prazo dos objetos digitais.

Espera-se que com o tempo surjam distintos “perfis” com margem de variação aceitável de propriedades significativas, que diferentes instituições preservadoras possam adotar ou

adaptar conforme julguem necessário. Nesse panorama, o Inspect desempenha um papel crucial, fornecendo uma estrutura para a identificação e a preservação das propriedades significativas dos documentos digitais. Essa estratégia não apenas reforça a teoria arquivística, mas também propõe uma aplicação prática que pode transformar as práticas de preservação digital.

O Projeto Inspect definiu, para cada um dos quatro tipos de registros digitais selecionados, um conjunto de características que precisam ser mantidas ao longo do tempo para que esses registros possam ser considerados autênticos. Assim, quando um registro digital passa por uma conversão de formato para fins de preservação digital, suas características podem ser analisadas para se verificar a manutenção das propriedades significativas. Essa questão é especialmente pertinente para o acervo da Comissão Nacional da Verdade, que contém registros digitais essenciais para a história e a memória nacional. O Projeto adota essa metodologia que identificou as propriedades significativas e da definição de perfis com margens aceitáveis de variação, as instituições responsáveis pela custódia dos documentos conseguem assegurar que as características fundamentais sejam preservadas de acordo com as necessidades do seu público-alvo, mesmo durante a transição para novos sistemas de armazenamento ou formatos de arquivo. Isso se torna fundamental em um contexto tecnológico em constante mudança, no qual a obsolescência de *software* e *hardware* pode representar riscos consideráveis à preservação a longo prazo dos documentos digitais.

A estrutura dos documentos arquivísticos digitais é analisada através de seus atributos. O Projeto Inspect⁵ partiu do princípio de que esses atributos são considerados propriedades significativas se impactarem a qualidade, usabilidade, acesso e comportamento, podendo ser organizados em cinco categorias: conteúdo, contexto, aparência, estrutura e comportamento. Essas propriedades são essenciais para a autenticidade dos documentos arquivísticos digitais e devem ser preservadas.

O projeto resultou na produção de quatro relatórios parciais para cada tipo de objeto digitais analisados, nos quais foram analisadas as propriedades significativas de quatro tipos de objetos: gravações de áudio, imagens matriciais, textos estruturados e *e-mails*, contribuindo para o desenvolvimento de estratégias de caracterização e manutenção dessas propriedades ao longo do tempo.

⁵ O site do projeto está disponível em <https://significantproperties.kdl.kcl.ac.uk>. Neste endereço, é possível encontrar mais informações sobre o projeto, além dos relatórios produzidos.

O principal objetivo do Projeto Inspect era aprofundar o conceito de “propriedades significativas” e estabelecer um conjunto de propriedades significativas para determinadas categorias de objetos digitais.

A abrangente metodologia do projeto compreendia diversas etapas, sumarizadas a seguir:

- a) Análise Conceitual: identificação e sistematização de conceitos relacionados a propriedades significativas de conteúdo eletrônico.
- b) Estudos de Caso: análise de diferentes categorias de conteúdo digital, como gravações de áudio, e-mails, imagens raster e textos estruturados.
- c) Desenvolvimento de Diretrizes: criação de orientações e metodologias para garantir a preservação digital eficiente.
- d) Testes e Validação: avaliação das diretrizes por meio de testes práticos e validação com diversos tipos de conteúdo digital.

O projeto definiu um conjunto de propriedades significativas para distintas categorias de objetos digitais e avaliou os métodos de mensuração dessas propriedades por meio de amostragens. A proposta foi testar, mapear, investigar e comparar as propriedades entre diferentes formatos.

Inicialmente, o projeto Inspect apresentou o seguinte conceito de propriedades significativas:

As características dos objetos digitais que devem ser preservadas ao longo do tempo, a fim de garantir a acessibilidade, usabilidade e significado contínuos dos objetos, e sua capacidade de serem aceitos como evidência do que pretendem registrar (Knight, 2009, p 4).

Conforme o Projeto Inspect avançou, a definição de propriedades significativas foi reformulada, conectando-se de maneira mais alinhada ao modelo *Open Archival Information System* (OAIS), o que exigiu uma análise aprofundada de suas diretrizes e definições.

A redefinição proposta pelo grupo do Inspect visou esclarecer e detalhar a relação entre propriedades significativas e a informação de representação.

As características dos objetos digitais que devem ser preservadas ao longo do tempo, a fim de garantir a acessibilidade, usabilidade e significado contínuos dos objetos, e sua capacidade de serem aceitos como evidência do que pretendem registrar e as delineou a partir das informações de representação (Knight, 2009, p. 4).

O Pacote de Informação do OAIS) e a Informação de Representação são a base do modelo de dados apresentados no modelo OAIS. O pacote de informação é uma estrutura conceitual que representa a unidade básica de informação dentro do OAIS, que contém todos os componentes necessários para preservar, gerenciar e acessar uma informação digital. O

pacote inclui dois componentes, sendo: i) Informação de conteúdo; e ii) Informação de preservação de documentos.

A Informação de Representação é um componente específico do Pacote de Informação do OAIS, que diz respeito ao conteúdo efetivamente preservado e gerenciado.

O conceito de propriedades significativas está vinculado ao de autenticidade. Os critérios necessários para assegurar a autenticidade de um objeto digital orientam a identificação das propriedades que podem ser consideradas significativas. A etapa inicial do projeto consistiu em uma revisão da literatura sobre propriedades significativas, analisando suas versões, perspectivas e abordagens.

Wilson (2007) fez um resumo sobre o uso e o entendimento do conceito de propriedades significativas em projetos precedentes, como o projeto *Curl Exemplar in Digital Archives* (Cedars) e o *Creative Archiving at Michigan and Leeds* (CAMiLEON), bem como em programas de preservação digital de várias instituições.

Durante o primeiro ano, a equipe fez uma extensa análise da literatura para identificar estruturas de estudo. Vários estudos abordaram métodos formais e semiformais para reconhecer propriedades significativas de objetos digitais. Rothenberg e Bikson elaboraram um modelo de análise de necessidades que consiste em quatro etapas principais. O Projeto Inspect incorporou elementos de diversos projetos, sendo a abordagem da RAND Europe particularmente útil para identificar as exigências das partes interessadas em contextos específicos. Assim, moldado com base em componentes e metodologias de outras iniciativas de preservação digital, o projeto considerado como elementos basilares:

Diretrizes de Preservação Digital: métodos e orientações desenvolvidos em iniciativas anteriores, como o OAIS, que oferecem uma estrutura para a preservação de conteúdo digital.

Estudos de Caso: investigações sobre diferentes tipos de conteúdo digital realizadas em projetos passados, como a Preservação de Conteúdo Digital da Biblioteca Britânica.

Ferramentas de Avaliação: instrumentos e metodologias para a avaliação de propriedades significativas que foram desenvolvidos em projetos anteriores, como o Projeto *Keeping Emulation Environments Portable* (KEEP).

Colaboração Internacional: cooperação com outros projetos internacionais de preservação digital, como o Projeto *Preservation and Long-term Access through Networked Services* (PLANETS).

Os componentes foram integrados ao Projeto Inspect, envolvendo uma abordagem robusta e eficaz para a preservação do conteúdo eletrônico ao longo do tempo, a qual conforme

apresentado por Knight (2009), reconhece que diversos *stakeholders* podem estar relacionados a um mesmo objeto digital e que os tipos de agentes, tais como entidades custodiadoras e usuários, podem mudar com o tempo. Cada agente possui diferentes níveis de conhecimento e necessidades, o que pode impactar as escolhas referentes à preservação digital. Assim, a variação aceitável das propriedades significativas após uma conversão pode ser diferente dependendo das exigências dos usuários.

A metodologia organiza o processo de avaliação em etapas e passos distintivos, criando um fluxo de trabalho estruturado e eficiente. A etapa inicial, denominada “análise de requisitos do objeto”, exige uma compreensão detalhada do objeto em questão, mapeando suas características principais e requisitos funcionais. Em seguida, a “análise de requisitos do agente” examina as necessidades e capacidades do agente que interage com o objeto. A fase de “reformulação” proporciona uma revisão e aprimoramento das análises anteriores, assegurando que as avaliações sejam precisas e pertinentes. O “avaliador” desempenha uma função fundamental, pois lidera o processo analítico, garantindo que cada subtarefa seja realizada com rigor e atenção aos detalhes.

O projeto parte do princípio de que existe uma margem de variação aceitável após a conversão de formato. Muitas partes interessadas podem aceitar um objeto que omita conteúdo ou funcionalidade específica, considerando o propósito do objeto e a parte interessada que o utiliza. Um curador pode identificar as funções exigidas pelo produtor nos estágios iniciais do ciclo de vida do objeto e avaliar se elas continuam necessárias quando usadas por uma comunidade diferente de usuários. Isso pode levar o curador a adotar uma estratégia de preservação mais rápida, simples e menos dispendiosa do que estratégias que mantêm todos os elementos do objeto.

Na fase do estudo, denominada análise do objeto, o avaliador analisou uma amostra representativa de cada tipo de objeto digital, identificando um conjunto de funções e comportamentos. Em cada caso, foi produzida uma lista para inspecionar as propriedades. Cada propriedade foi analisada quanto à sua função e determinada como significativa (ou não) se proporcionasse acessibilidade, usabilidade ou significado ao objeto.

As propriedades significativas elencadas pelo Projeto Inspect para cada tipo de objeto digital estão expressas nos Quadros, apresentados a seguir.

No que corresponde os quadros 4 e 5, o relatório de Teste de Propriedades Significativas para Gravações de Áudio Digital, Knight (2009), analisou as propriedades significativas do áudio digital, identificando quais deveriam ser mantidas e examinando seus elementos e

funções. Também explorou estratégias para preservação em o acesso a longo prazo aos ativos de áudio, descreveu experimentos realizados para identificar e avaliar ferramentas capazes de converter propriedades significativas entre diferentes formatos e revisou diversas especificações que estão em uso para o armazenamento e descrição de áudio digital.

Um objeto de áudio pode conter diversos formatos de metadados, dependendo do formato de representação utilizado e das escolhas do criador. Alguns formatos de metadados são desenvolvidos para finalidades específicas, como os metadados ID3 em arquivos MP3, enquanto outros, como o bloco Bext em arquivos Broadcast Wave, podem armazenar uma mistura de metadados técnicos e descritivos. A partir dessa análise chegou-se a listagem representada nos quadros abaixo:

Quadro 4 -Propriedades significativas de registro sonoro

Quantidade	Propriedades significativas
01	Duração
02	Profundidade de bits
03	Taxa de amostragem
04	Número de canais
05	Campo sonoro
06	Localização do mapa de som para cada canal

FONTE: Adaptado de Projeto Inspect (2009).

Se a gravação de áudio contiver metadados no formato *Broadcast Exchange Format* (BEXT), as informações elencadas no quadro a seguir também deverão ser retidas.

Quadro 5 - Propriedades significativas de registro sonoro no formato BEXT

Quantidade	Propriedades significativas
07	Descrição
08	Originador
09	Referência do Originador
10	Data de Origem
11	Tempo de Origem
12	Histórico de codificação
13	Relatório de qualidade
14	Folha de dicas

FONTE: Adaptado de Projeto Inspect (2009).

Já no quadro 6, está representado as propriedades significativas definidas a partir do relatório que examinou as propriedades significativas referente as mensagens de correio eletrônico, uma forma comum de comunicação digital. O objetivo foi identificar as propriedades essenciais do e-mail que devem ser preservadas, analisando cada um de seus elementos constituintes e suas funções designadas. O relatório explorou estratégias para manter

o acesso a ativos de e-mail a longo prazo. Por fim, descreve um conjunto de propriedades significativas. Knight (2009).

Quadro 6 - Propriedades significativas de mensagens de correio eletrônico⁶

Quantidade	Propriedades significativas
01	Local
02	Domínio
03	Relacionamento
04	Assunto
05	Campo de rastreamento
06	Corpo da mensagem sem marcação
07	Anexos

FONTE: Adaptado de Projeto Inspect (2009).

No que corresponde ao quadro 7, e aos dados apresentados no relatório que analisou as propriedades significativas de imagens matriciais, ou bitmaps, identificando quais devem ser mantidas e examinando seus elementos constituintes e funções. Para determinar essas propriedades, foi necessário um método formal e consistente. O Arquivo Nacional da Austrália (2002) desenvolveu um "Modelo de Desempenho", adotado pelo Projeto Inspect, que estabelece o conceito de "essência" de um registro digital, contendo características que devem ser preservadas para manter seu significado ao longo do tempo.

Quadro 7 -Propriedades significativas de imagem matricial

Quantidade	Propriedades significativas
01	Largura da imagem
02	Altura da imagem
03	Frequência de amostragem X
04	Frequência de amostragem Y
05	Bits por amostra
06	Amostra por pixel
07	Amostras extras

FONTE: Adaptado de Projeto Inspect (2009).

No que corresponde ao quadro 8, o relatório de Teste de Propriedades Significativas para Texto Estruturado de acordo com Montague (2009) analisou as propriedades significativas de documentos de texto estruturado, identificando quais devem ser mantidas e examinando seus elementos e funções. Também explorou estratégias para preservar o acesso a longo prazo aos ativos de texto estruturado. Além disso, descreveu experimentos realizados para identificar e avaliar ferramentas capazes de converter propriedades significativas entre diferentes formatos.

⁶ Definição do termo "Correio Eletrônico" pelo Glossário da Câmara Técnica de Documentos: sistema usado para criar, transmitir e receber mensagem eletrônica e outros documentos digitais por meio de rede de computadores..

Texto estruturado foi definido como uma ampla gama de conteúdos codificados por diversos métodos, contendo caracteres alfabéticos, numéricos e de pontuação, acompanhados por informações sobre sua descrição ou aparência. A principal característica que distingue texto estruturado de não estruturado é a presença de marcação, originada na indústria editorial. Exemplos de texto estruturado incluem HTML, XML, código-fonte e e-mails.

Para determinar as propriedades significativas de um registro digital, é necessário um método formal e consistente. O Arquivo Nacional da Austrália desenvolveu um "Modelo de Desempenho", adotado pelo Projeto Inspect, que estabelece o conceito de "essência" de um registro digital, contendo características que devem ser preservadas para manter seu significado ao longo do tempo. A partir desse cenário foi possível alcançar o resultado apresentado no quadro abaixo:

Quadro 8 - Propriedades significativas de texto estruturado

Quantidade	Propriedades significativas
01	Título
02	Criador
03	Data
04	Palavra-chave
05	Direitos
06	Divisão
07	Período
08	Idioma
09	Parágrafo
10	Quebra de Linha
11	Cabeçalho
12	Ênfase
13	Negrito
14	Itálico
15	Sublinhado
16	Forte Ênfase
17	Tachado
18	Regra Horizontal
19	Texto inserido
20	Texto excluído
21	Amostra
22	Citar
23	Termos definidos
24	Código
25	Abreviaturas
26	Sigla
27	Citações
28	Subscrito/Sobrescrito
29	Endereço
30	Botão

31	Elementos da lista
32	Elementos da tabela
33	Imagem
34	Ligação
35	Miniaplicativo
36	Quadro
37	Conjunto de Quadros

FONTE: Adaptado de Projeto Inspect (2009).

O projeto também considera que as entidades responsáveis pela custódia podem optar por aceitar uma certa perda de funcionalidade, para facilitar a preservação, ou seguir uma abordagem mais conservadora, visando manter todas as propriedades do objeto informacional.

A primeira versão do quadro de avaliação traçou um caminho que os avaliadores poderiam seguir e apresentou um modelo para registrar os valores atribuídos. O avaliador iniciaria a análise examinando o objeto informacional como um todo e, em seguida, exploraria seus subcomponentes até identificar as propriedades técnicas necessárias para a sua reprodução. Depois, colaboraria com as partes interessadas para avaliar os limites aceitáveis necessários para alcançar os objetivos propostos. Essa estrutura serviu como base para os testes, com o intuito de identificar e mensurar as propriedades significativas antes e após as migrações dos formatos dos arquivos.

Na fase de reformulação, o avaliador implementou ações de preservação dos arquivos, realizando a migração de formatos e avaliando os resultados em comparação com os objetos originais. As propriedades significativas foram extraídas dos arquivos amostrados e analisadas após a migração para formatos alternativos. Quatro relatórios pormenorizados documentaram as investigações para identificar as propriedades significativas em registros sonoros, e-mails, imagens matriciais e textos estruturados. Os avaliadores conseguiram definir um conjunto mínimo de propriedades significativas para cada categoria de objeto.

A metodologia e a estrutura voltadas para a avaliação das propriedades significativas foram eficazes em todos os quatro tipos de objetos analisados. Notou-se que os arquivos de registro sonoro apresentavam propriedades adicionais relevantes quando os metadados BEXT estavam presentes e que propriedades adicionais eram necessárias para manter a integridade de um encadeamento de e-mail.

Os estudos de caso sobre imagens matriciais contribuíram para delinear propriedades significativas a partir das Informações de Representação (Repinfo). Embora o espaço de cores seja fundamental para a reprodução da imagem, ele não é considerado significativo se o objeto digital for migrado de RGB para CMYK, por exemplo.

Através de múltiplas transformações, foi possível comparar diversos formatos de arquivo para determinar qual preservava da melhor forma as propriedades significativas. Essa análise pode indicar formatos de arquivo preferidos para o planejamento de preservação.

A ferramenta Comparador, desenvolvida no Projeto Planets, mede as diferenças entre os objetos digitais antes e depois do tratamento, e seria útil para comparar os resultados obtidos em ambos os projetos.

É importante ressaltar que a quantidade limitada de ferramentas de caracterização nem sempre conseguiu identificar todas as propriedades significativas. Portanto, há uma necessidade de desenvolver mais ferramentas nessa área. A análise funcional permitiu ao avaliador considerar o propósito de cada propriedade tanto em termos de características técnicas quanto em relação a uma comunidade específica, e essa análise pode ser revista na fase de reformulação.

Os relatórios de testes examinam a noção de propriedades significativas dos quatro tipos de objetos digitais escolhidos pelo projeto. Procura identificar as propriedades significativas para os quatro tipos de objetos digitais que devem ser mantidas, examinando cada um dos elementos constituintes e analisando a sua função designada, bem como examina as estratégias que podem ser utilizadas para manter o objeto digital acessível no longo prazo. Ainda, descreve um conjunto de experimentos que foram realizados para identificar e avaliar as ferramentas que podem ser utilizadas para converter em propriedades significativas.

Esses resultados contribuem para a compreensão e a preservação dos objetos digitais, garantindo que suas características essenciais sejam mantidas ao longo do tempo.

O Projeto Inspect gerou vários produtos, conforme destacado no relatório final:

- Um relatório sobre o conceito de propriedades significativas.
- Uma estrutura para avaliar as propriedades significativas.
- Um conjunto de propriedades significativas para cada tipo de objeto.
- Um dicionário de dados.
- Quatro relatórios de testes descritos na seção de metodologia.

Um destaque importante é o relatório de propriedades significativas de Knight (2009 *apud* Wilson, 2007), que ofereceu uma visão abrangente da discussão na literatura até o momento. Ele fornece uma definição explícita de propriedades significativas: “As características dos objetos digitais que devem ser preservadas ao longo do tempo, a fim de garantir a acessibilidade, usabilidade e significado contínuos dos objetos”.

No relatório final, o projeto apresenta cinco considerações que merecem ser destacadas.

As propriedades significativas oferecem uma maneira de analisar a manutenção das características dos objetos digitais ao longo de sua vida, mesmo com possíveis mudanças nos objetos ou no público.

As Comunidades Designadas são fundamentais e o Inspect garante que as necessidades dos usuários identificados sejam consideradas na determinação das propriedades significativas.

A identificação inicial das propriedades significativas pode ser realizada por um curador digital ao assumir o cuidado dos objetos, buscando representar a intenção do produtor. Avaliações adicionais podem ser necessárias à medida que os requisitos da comunidade designada mudam ou uma nova comunidade se torna o foco. Assim, as propriedades significativas podem atuar como uma ponte ao longo do tempo, garantindo a persistência do que é importante nos objetos digitais através de quaisquer transformações necessárias.

As informações de representação podem documentar o conhecimento necessário a uma comunidade designada para interpretar o objeto de dados codificados, enquanto as propriedades significativas podem documentar os estados anteriores e posteriores de quaisquer transformações, garantindo aos usuários que o significado do objeto de informação seja preservado.

As propriedades significativas precisam ser incorporadas aos fluxos de trabalho rotineiros dos curadores digitais. Os curadores precisam de acesso a ferramentas apropriadas, como o JHOVE, para identificar propriedades de um determinado tipo e formato de objeto, bem como metadados para registrar essas informações no objeto.

Tais conclusões ressaltam a importância das propriedades significativas na preservação digital e a necessidade de incorporar essas práticas aos fluxos de trabalho dos curadores digitais, garantindo a preservação e o acesso contínuo aos objetos digitais ao longo do tempo.

O projeto, propõe, ainda, algumas recomendações importantes, conforme expostas a seguir.

Realizar mais experimentações utilizando outros formatos de representações e categorias de objeto digital.

Aplicar a versão 2 da estrutura Inspect aos quatro outros projetos financiados pelo JISC, que abrangem imagens matriciais, imagens em movimento, software e objetos de e-learning, além dos quatro tipos de objetos realizados pelos avaliadores do Inspect, para permitir uma análise mais granular.

Realizar mais experimentações com outras migrações e caracterizações.

Fazer aperfeiçoamentos específicos em JHOVE e outras ferramentas de caracterização para permitir a captura de mais propriedades significativas.

Utilizar conjuntos maiores de amostras para validar este trabalho e os subsequentes. Nesse sentido, o Projeto Planets pode ser utilizado como um espaço para realizar experiências, incluindo formatos apropriados e compartilhamento dos resultados com o Projeto Planets.

Desenvolver casos de uso para propriedades significativas em Comunidades Designadas específicas. O Centro de Pesquisa Eletrônica conduzirá tal tarefa no Projeto Propriedades Significativas no Laboratório (SPIL), recentemente financiado pelo JISC como um demonstrador de propriedades significativas.

Realizar consulta formal a órgãos apropriados sobre a utilização do projeto de dados Dicionário com PREMIS, UDFR (sucessor planejado do PRONOM) e JHOVE.

As recomendações visam aprimorar a preservação e a análise de objetos digitais, garantindo a captura e a manutenção de propriedades significativas de maneira mais eficaz e abrangente.

A preservação de documentos arquivísticos digitais é uma tarefa coletiva que exige comprometimento contínuo, investimento em novas tecnologias e infraestrutura, além de um trabalho conjunto entre profissionais de diferentes áreas, como arquivistas, tecnólogos, legisladores e a sociedade em geral. Somente dessa forma conseguiremos assegurar que o legado digital da humanidade seja mantido para o benefício de todos.

O processo vai além de uma mera questão técnica; ele abarca também uma preocupação cultural e histórica. Os documentos digitais são parte essencial do nosso patrimônio cultural e histórico, e garantir sua preservação é fundamental para que as futuras gerações tenham acesso a esse legado. É crucial que instituições, pesquisadores e especialistas colaborem para criar soluções inovadoras que assegurem a durabilidade e a integridade dos recursos digitais, o que pode ser considerado um desafio nos dias de hoje, especialmente no que diz respeito a entidades responsáveis, como o Arquivo Nacional do Brasil. A autenticidade dos documentos digitais precisa ser cuidadosamente preservada ao longo do tempo. Trata-se de uma questão complexa e multifacetada, que envolve não apenas a manutenção da integridade física dos dados, mas também a preservação de seu significado e acesso ao longo dos anos.

A migração, como estratégia de preservação, implica não apenas transferência de dados, mas também conservação de suas propriedades significativas, as quais são fundamentais para a autenticidade e a usabilidade dos documentos.

O projeto Inspect sublinha a relevância de reconhecer e salvaguardar essas propriedades, de modo a garantir que os documentos digitais continuem acessíveis e compreensíveis no futuro. Nesse contexto, os metadados têm um papel fundamental, funcionando como um registro detalhado das características do documento, possibilitando verificar sua autenticidade e monitorar quaisquer mudanças ao longo do tempo.

Para verificar se a autenticidade dos documentos digitais arquivados foi mantida após a migração de formatos, é fundamental identificar e conservar suas propriedades significativas.

6 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa, de natureza exploratória e qualitativa, baseia-se no estudo de caso do acervo digital recolhido pela CNV, com objetivo de verificar se as propriedades significativas dos documentos foram mantidas durante a conversão dos formatos no processo de ingestão no Archivematica. O estudo fundamenta-se na presunção da autenticidade de documentos arquivísticos digitais e na necessidade de identificação e preservação de suas propriedades significativas.

Com o objetivo de compreender as propriedades significativas para garantir a autenticidade digital, foi analisada a metodologia adotada pelo Projeto Inspect⁷. Além disso, documentos do Arquivo Nacional foram examinados, incluindo os termos de recolhimento e as listagens recebidas pela CNV no momento da transferência dos arquivos.

A CNV desempenhou um papel importante na apuração de graves violações de direitos humanos ocorridas entre 1946 e 1988, incluindo os 21 anos da ditadura militar no Brasil. O trabalho resultou em um relatório final de três volumes, documentando as conclusões e recomendações sobre essas violações.

6.1 ETAPAS DA PESQUISA

A pesquisa teve seu desenvolvimento estruturado em três etapas principais: i) identificação e seleção dos documentos; ii) processamento e submissão ao sistema de preservação, e iii) análise comparativa dos resultados.

6.1.1 Identificação e seleção dos documentos

Devido ao quantitativo variado de formatos e extensões recolhido nesse acervo, iniciaram-se os métodos de preservação atribuídos a esse acervo. Verificou-se que esses documentos, recolhidos em 2015, ainda não estavam submetidos à plataforma de preservação, mas os diversos fatores relacionados a isso não compõem o escopo de investigação desta pesquisa. Porém, com base no conhecimento da metodologia do Projeto Inspect, identificou-se a necessidade de verificar as propriedades significativas desses objetos digitais e analisar seu

⁷ Disponível em: <https://significantproperties.kdl.kcl.ac.uk/index.html>. Acesso em: 19 mar. 2025.

comportamento na plataforma de preservação. Optamos, então, por elaborar o checklist a seguir para sintetizar as etapas da metodologia quando as estratégias de preservação fossem aplicadas:

Identificação dos documentos digitais recolhidos do fundo estudado;

Levantamento das quantidades de formatos recolhidos;

Seleção dos formatos para análise;

Identificar as extensões dos arquivos;

Definir dos arquivos a serem analisados;

Escolha da ferramenta de extração de metadados (EXIFtool);

Definição da metodologia de análise;

Escolha da plataforma de preservação (Archivematica).

Depois desses passos iniciais, fizemos uma pesquisa de quais ferramentas poderiam auxiliar nessa análise para alcançarmos o propósito desta pesquisa.

6.1.2 Processamento e submissão ao Sistema de Preservação

A próxima etapa foi realizada a partir dos seguintes pontos:

- a) Análise dos arquivos digitais recolhidos;
- b) Criação da estrutura de submissão (SIP);
- c) Extração dos metadados antes da submissão via EXIFtool;
- d) Submissão do SIP ao Archivematica;
- e) Acompanhamento dos microserviços até a formação do AIP;
- f) Download do AIP e extração dos metadados;
- g) Organização dos dados em quadros comparativos (metadados do SIP versus metadados AIP);
- h) Comparação das propriedades significativas do objeto antes e depois da preservação.

6.1.3 Análise comparativa dos resultados

Para apresentação dos resultados, realizaram-se na etapa:

- a) Verificação de possíveis alterações nas propriedades significativas;
- b) Avaliação da permanência das propriedades significativas, conforme os conceitos do Projeto Inspect;
- c) Discussão dos resultados e sua relação com a preservação digital a longo prazo.

6.1.4 Ferramentas utilizadas

Para auxiliar na visualização das ferramentas utilizadas indicam-se:

- a) EXIFtool: ferramenta de extração de metadados antes e depois da submissão ao Archivematica;
- b) Archivematica: plataforma de preservação digital utilizada pela entidade custodiadora;
- c) JHOVE: software de validação e caracterização de formatos de arquivos digitais;
- d) MediaInfo e Metadata Online: ferramentas utilizadas para tentativa de extração inicial de metadados.

7 RESULTADOS

7.1 DOCUMENTOS DA COMISSÃO NACIONAL DA VERDADE RECOLHIDOS PARA O ARQUIVO NACIONAL;

O acervo da Comissão Nacional da Verdade, recolhido ao Arquivo Nacional em 2015, foi integrado ao Centro de Referência das Lutas Políticas no Brasil (1964-1985) – Memórias Reveladas. Esse centro oferece acesso aos documentos digitais e não digitais, facilitando a pesquisa e a consulta por historiadores, estudantes e o público em geral.

O acervo inclui documentos textuais, audiovisuais, iconográficos, cartográficos e sonoros, todos organizados para facilitar o acesso e a pesquisa, contribuindo para a transparência e a memória histórica do Brasil.

Primeiramente, identificamos os formatos dos arquivos recolhidos no acervo, resultando em uma grande variedade de formatos. Após essa identificação, decidimos estudar especificamente os arquivos de áudio presentes no acervo da CNV, contabilizando a variação das extensões desses arquivos.

Utilizando um software, elaboramos uma planilha para demonstrar os tipos de extensões desses formatos de arquivos no fundo da CNV.

A análise visa fornecer uma visão detalhada dos tipos de arquivos e suas extensões, facilitando a compreensão e o manejo dos documentos no processo de preservação digital.

Figura 3 -Quantitativo de extensão em formato áudio

Audio Files	251.03GB	251.03GB	5.35%	1503 Files containing music, sounds or playlists
.wav	161.00GB	161.00GB	3.00%	294 Arquivo WAV
.mp3	89.00GB	89.00GB	2.00%	1156 Arquivo MP3
.wma	663.00MB	663.00MB	0.00%	22 Arquivo WMA
.aac	545.00MB	545.00MB	0.00%	14 Arquivo AAC
.m4a	301.00MB	301.00MB	0.00%	11 Arquivo M4A
.ogg	15.00KB	16.00KB	0.00%	4 Arquivo OGG
.wpl	567.00Bytes	8.00KB	0.00%	2 Arquivo WPL

Fonte: A autora (2024).

Figura 4 -Quantitativo de extensão em formato imagem matricial

Graphic Files	1.24TB	1.24TB	27.07%	1036501 Files containing pictures, images or mouse cursors
.dng	13.00MB	13.00MB	0.00%	1 Arquivo DNG
.gif	6.00MB	6.00MB	0.00%	182 Arquivo GIF
.tif	52.00GB	52.00GB	1.00%	42017 Arquivo TIF
.indd	1.00GB	1.00GB	0.00%	617 Arquivo INDD
.dwg	95.00MB	96.00MB	0.00%	254 Arquivo DWG
.png	425.00MB	427.00MB	0.00%	943 Arquivo PNG
.cr2	5.00GB	5.00GB	0.00%	286 Arquivo CR2
.psd	340.00MB	340.00MB	0.00%	11 IrfanView PSD File
.bmp	77.00MB	78.00MB	0.00%	8 Arquivo BMP
.jpeg	15.00MB	15.00MB	0.00%	40 Arquivo JPEG
.tiff	20.00MB	20.00MB	0.00%	16 Arquivo TIFF
.ai	12.00KB	24.00KB	0.00%	3 Arquivo AI
.nef	13.00GB	13.00GB	0.00%	711 Arquivo NEF
.arw	38.00GB	38.00GB	1.00%	2729 Arquivo ARW
.jpg	1.00TB	1.00TB	25.00%	988683 Arquivo JPG

Fonte: A autora (2024).

7.2 ANÁLISE DAS PROPRIEDADES SIGNIFICATIVAS DOS DOCUMENTOS DIGITAIS COM BASE NA METODOLOGIA DO PROJETO INSPECT

Em seguida, os arquivos foram organizados em uma estrutura de diretórios. Para cada extensão de arquivo foi criada uma pasta específica; dentro de cada uma, uma subpasta nomeada com a extensão do arquivo em ordem sequencial. Por exemplo: se o arquivo tinha a extensão MP3, a pasta foi nomeada da seguinte maneira: 0001_mp3. Dentro de cada subpasta, foram criados dois diretórios: “*metadata*” e “*objetic*”. O arquivo de áudio correspondente à extensão foi colocado na pasta “*objetic*” e foi criado um arquivo de metadados utilizando uma planilha do LibreOffice, armazenado na pasta “*metadata*”. Esse procedimento foi repetido para todas as outras extensões de áudio.

Após organizar os arquivos de áudio, utilizamos a *ferramenta Metadata Online* para extrair os metadados. No entanto, não foi possível obter os metadados correspondentes às propriedades significativas previstas no Projeto Inspect para documentos de áudio. Em seguida, foi utilizada a ferramenta *MediaInfo*, que também é acessível via web, eliminando a necessidade de baixar programas executáveis para verificar os metadados. Concluímos que a *MediaInfo* era a melhor opção para alcançar os resultados desejados.

MediaInfo era a melhor opção para alcançar os resultados desejados.

No entanto, nem a *MediaInfo* nem a *Metadata Online* retornaram os metadados buscados. Nenhuma das duas ferramentas trouxe a correspondência esperada. Consultando o Projeto Inspect, foi identificado que a extração dos metadados havia sido realizada com a utilização da ferramenta JHOVE (JSTOR/Harvard Object Validation Environment), que é uma ferramenta de identificação, validação e caracterização desenvolvida pela JSTOR e pela Harvard University Library. Portanto, foi solicitada à Seção de Tecnologia, da Divisão de

Tecnologia do Arquivo Nacional, a disponibilização dessa ferramenta para a extração dos dados necessários.

Inicialmente, consideramos verificar quinze arquivos no formato MP3. No entanto, constatamos que analisar apenas um tipo de formato não seria vantajoso. Decidimos, então, seguir a política de preservação digital do arquivo nacional, que estabelece quais são os formatos e as extensões oficiais para o recolhimento. Além do MP3, foram identificados arquivos nos formatos WMA, WAVE, AIF, AAC e OGG. Acreditamos que essa variação de formatos no acervo da CNV seria suficiente para a análise.

Recorrendo à política de preservação digital do arquivo nacional, identificamos que o método de preservação exige que o AIP seja mantido em extensão WAVE e que os arquivos sejam convertidos para MP3 como formato de acesso (DIP), independentemente do formato original. Isso garante que o objeto seja preservado e acessível.

Para testar o comportamento da ferramenta, os SIPs foram submetidos. Em seguida, abrimos a pasta “*transfer*” e localizamos o objeto a ser submetido ao software Archivematica. O objeto foi selecionado e adicionado, assim, a transferência foi iniciada dentro do software, que analisou seu comportamento ao longo do processo. Após a conclusão dos microserviços, foi feito o download do AIP para extração dos metadados.

O Projeto Inspect fez uso de ferramentas diferente das que optamos usar nesta pesquisa, pois levamos em consideração o tempo da pesquisa realizada pelo projeto e acreditamos que no correr desta pesquisa teríamos ferramentas mais robustas que pudessem retornar melhor os metadados, já que sabemos que a tecnologia sofre obsolescência rápida.

Assim como no relatório de texto do Inspect, sugere-se a eleição de outras ferramentas. No relatório de teste do Projeto, recomenda-se que mais experimentações sejam realizadas com outras ferramentas de migração e caracterização, a fim de comparar resultados e desenvolver ferramentas conforme necessário.

7.3 ESTUDO DAS VARIAÇÕES NAS PROPRIEDADES SIGNIFICATIVAS ENTRE OS DOCUMENTOS DIGITAIS RECOLHIDOS E PRESERVADOS NO REPOSITÓRIO DIGITAL CONFIÁVEL DO ARQUIVO NACIONAL

Nesta pesquisa, optou-se por utilizar a mesma ferramenta de extração de metadados que a plataforma de preservação adota, o EXIFtool. Isso garantiu a continuidade do uso das ferramentas e plataformas já implementadas pela entidade custodiadora, o Arquivo Nacional,

que integra o campo empírico dessa pesquisa. A escolha do EXIFtool para extração de metadados e do Archivematica para conversão das extensões para formatos de preservação seguiu Política de Preservação Digital do Arquivo Nacional

Ao adotar a infraestrutura da entidade custodiadora, sabíamos que precisaríamos alinharmos ao tempo e às regras da instituição. Durante a execução da pesquisa, enfrentamos algumas dificuldades, que foram prontamente resolvidas com o apoio dos colaboradores da TI e da chefe da Divisão de Preservação do Acervo Digital. A pesquisa foi conduzida utilizando os recursos do Arquivo Nacional, com o objetivo de contribuir para a relevância das atividades desenvolvidas pela Divisão de Preservação do Acervo Digital, que é responsável pela manutenção e preservação do acervo nato-digital e dos representantes digitais do Executivo Federal, no caráter permanente.

Trabalhar em uma organização governamental traz desafios adicionais, pois todo o trabalho precisou ser realizado dentro dos procedimentos operacionais padrão relacionados ao uso da internet e do *software*, o que nos levou a depender da Divisão de Tecnologia. Este foi um risco que assumimos ao longo do processo. Após superar tais questões, realizou-se o levantamento quantitativo dos dados analisados.

Optamos por analisar os seguintes formatos de arquivo:

- Áudio:
 - 4 (quatro) Áudios em WMA;
 - 4 (quatro) Áudios em MP3.
- Imagem Matricial:
 - 4 (quatro) imagens em BMP;
 - 2 (quatro) Imagens em PSD;
 - 4 (quatro) imagens em TIFF;
 - 4 (quatro) Imagens em JPEG.

A escolha desses formatos pode parecer curiosa, dado a variedade de extensões apresentadas nas Figuras 5 e 6. Essa decisão foi motivada pela resposta limitada do Archivematica, que, durante o processo de admissão dos arquivos, gerou apenas esses formatos de saída (AIP), o que nos restringiu às extensões escolhidas.

Em seguida, passou-se para a extração e análise dos metadados. A pesquisa seguiu por várias etapas, começando com o teste de diversos arquivos de uma mesma extensão e formato utilizando ferramentas web de extração de metadados. Não obtendo resultados satisfatórios, modificou-se a abordagem, testando extensões variadas dentro de um único formato. Mesmo

com a ampliação da variedade, observou-se que os resultados não eram adequados, o que levou à decisão de acrescentar um novo formato e ampliar ainda mais a gama de extensões a serem analisadas.

Esta pesquisa se justifica pelo fato de que, até o momento, não foi identificada nenhuma análise que examine as propriedades significativas dos objetos digitais antes e depois da submissão a uma plataforma de preservação e do momento da conversão.

O levantamento dos dados foi realizado e está reunido em quadros apresentados abaixo. Primeiramente, apresenta-se os quadros que reúnem os dados do formato Áudio e as extensões WMA e MP3.

Observando os Quadros 8 a 15, temos os dados coletados na etapa de pesquisa, na primeira coluna temos as propriedades significativas determinadas pelo projeto Inspect, na segunda coluna temos os metadados do pacote SIP do objeto digital a ser submetido na plataforma de preservação digital e na terceira coluna temos os metadados do AIP, objeto digital já convertido para formato de preservação.

Nos quadros de 9 a 13, a ferramenta de extração de metadados não conseguiu retornar o campo “Bit per sample” nos metadados referente ao SIP, já nos quadros de 14 a 16 a ferramenta de extração de metadados não retornou o campo “*audio channels e bit per sample*”, os demais campos a ferramenta retornou, e foi possível observar que em todos os casos o AIP, objeto convertido para formato de preservação, apresenta uma alteração no tamanho do objeto digital. Esta pesquisa, também optou por manter a data de criação desse objeto, para mostrar que esse documento foi criado em uma data distante da data de realização dessa pesquisa, como marco para que se um dia esse objeto nessa analisado novamente poderá comparar se houve alguma alteração.

Dado essa análise, para o formato áudio percebemos que não houve alterações nas propriedades significativas dos objetos digitais.

Quadro 9 - Áudio 0001: extensão .wma

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_WMA.WMA	0001_WMA-97367a2b-4bcc-41b9-aa2f-406d480038f9.wav
Directory	S:/Ambiente_Treinamento/SIP/Treinamento/CNV/AUDIO/WMA/0001_WMA/objects	S:/Ambiente_Treinamento/AIP/8b36/ba81/1c54/4825/b084/62af/ca72/637e
File Size	29 MB	322 MB
File Modification Date/Time	2015:07:20 11:00:05-03:00	2025:02:25 18:13:18-03:00

File Type	WMA	WAV
File Type Extension	WMA	WAV
Audio Channels	2	2
Audio Sample Rate	44100	44100
Duration	0:30:26	0:30:25
Bits Per Sample		16 bits

Fonte: A autora (2024).

Quadro 10 - Áudio 0002: extensão .wma

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_WMA.WMA	0002_WMA-245b5e7a-4a5b-44d1-8cd7-ba74e4be6ef5.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ WMA/0002_WMA/objects	S:/Ambiente_Treinamento/AIP/ 7ce3/26c5/8173/4154/ac7b/97 20/517d/61d8
File Size	28 MB	306 MB
File Modification Date/Time	2015:07:20 11:09:33-03:00	2025:02:25 18:13:30-03:00
File Access Date/Time	2015:07:20 11:09:33-03:00	2025:02:25 18:13:30-03:00
File Creation Date/Time	2025:01:19 15:46:25-03:00	2025:02:25 18:13:44-03:00
File Type	WMA	WAV
File Type Extension	WMA	WAV
Audio Channels	2	2
Audio Sample Rate	44100	44100
Duration	0:28:57	0:28:55
Bits Per Sample	-	16 bits

Fonte: A autora (2024).

Quadro 11 - Áudio 0003: extensão .wma

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0003_WMA.WMA	0003_WMA-1457299a-e9cc-49ec-8d8b-f3289f7ca8c3.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ WMA/0003_WMA/objects	S:/Ambiente_Treinamento/AIP/ 3284/f04d/d6fd/4b90/9c37/491 b/818b/2c5
File Size	28 MB	308 MB
File Modification Date/Time	2015:07:20 11:17:17-03:00	2025:02:27 19:25:45-03:00
File Access Date/Time	2015:07:20 11:17:17-03:00	2025:02:27 19:25:45-03:00
File Creation Date/Time	2025:01:19 15:46:25-03:00	2025:02:27 19:25:59-03:00
File Type	WMA	WAV

File Type Extension	WMA	WAV
Audio Channels	2	2
Audio Sample Rate	44100	44100
Duration	0:29:06	0:29:05
Bits Per Sample	-	16 bits

Fonte: A autora (2024).

Quadro 12 - Áudio 0004: extensão .wma

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0004_WMA.WMA	0004_WMA-dc5c3d58-f416-4bb4-bc57-eabea95cc59e.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ WMA/0004_WMA/objects	S:/Ambiente_Treinamento/AIP/ 6fc8/5c64/3d3b/4cdd/b3c0/27f a/3eb3/bf65
File Size	28 MB	309 MB
File Modification Date/Time	2015:07:20 11:27:06-03:00	2025:02:27 19:26:09-03:00
File Access Date/Time	2015:07:20 11:27:06-03:00	2025:02:27 19:26:09-03:00
File Creation Date/Time	2025:01:19 15:46:25-03:00	2025:02:27 19:26:24-03:00
File Type	WMA	WAV
File Type Extension	WMA	WAV
Audio Channels	2	2
Audio Sample Rate	44100	44100
Duration	0:29:11	0:29:10
Bits Per Sample	-	16 bits

Fonte: A autora (2024).

Quadro 13 - Áudio 0001: extensão .mp3

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_mp3.mp3	0001_MP3-41e579bf-8461-44a8-a98c-1e4925938d65.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ MP3/0001_MP3/objects	S:/Ambiente_Treinamento/AIP/ acab/bd79/1850/4181/aef6/f3c a/7a8c/4efd
File Size	380 MB	2.1 G
File Modification Date/Time	2013:12:16 11:38:54-03:00	2025:02:25 18:44:00-03:00
File Access Date/Time	2013:12:16 11:38:54-03:00	2025:02:25 18:44:00-03:00
File Creation Date/Time	2025:01:19 15:46:17-03:00	2025:02:25 18:44:34-03:00
File Type	MP3	WAV
File Type Extension	MP3	WAV

Audio Channels		2
Audio Sample Rate	44100	44100
Duration	3:17:47 (approx)	3:17:47
Bits Per Sample	-	16

Fonte: A autora (2024).

Quadro 14 - Áudio 0002: extensão .mp3

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_mp3.mp3	0002_MP3-a216ede1-7317-478f-8006-2f7c975ca1c6.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ MP3/0001_MP3/objects	S:/Ambiente_Treinamento/AIP/ 1221/5455/7d54/4daf/834f/a74 4/c494/b168
File Size	329 MB	1813 MB
File Modification Date/Time	2013:12:16 16:01:18-03:00	2025:02:25 18:43:43-03:00
File Access Date/Time	2013:12:16 16:01:18-03:00	2025:02:25 18:43:43-03:00
File Creation Date/Time	2025:01:19 15:46:19-03:00	2025:02:25 18:44:26-03:00
File Type	MP3	WAV
File Type Extension	MP3	WAV
Audio Channels		2
Audio Sample Rate	44100	44100
Duration	2:51:18 (approx)	2:51:18
Bits Per Sample	-	16

Fonte: A autora (2024).

Quadro 15 - Áudio 0003: extensão .mp3

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0003_mp3.mp3	0003_MP3-13a8ed50-6fa1-4164-8b59-1d830ea067b8.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ MP3/0003_MP3/objects	S:/Ambiente_Treinamento/AIP/ 9e30/b4df/f633/453b/9e90/242 9/b4bf/9789
File Size	156 MB	1143 MB
File Modification Date/Time	2011:02:10 21:10:34-03:00	2025:02:27 19:23:12-03:00
File Access Date/Time	2011:02:10 21:10:34-03:00	2025:02:27 19:23:12-03:00
File Creation Date/Time	2025:02:27 19:16:46-03:00	2025:02:27 19:23:38-03:00
File Type	MP3	WAV
File Type Extension	MP3	WAV
Audio Channels		2

Audio Sample Rate	44100	44100
Duration	1:48:02 (approx)	1:48:02
Bits Per Sample	-	16

Fonte: A autora (2024).

Quadro 16 - Áudio 0004: Extensão .mp3

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0004_mp3.mp3	0004_MP3-360bd20a-ab7f- 44b9-b290-23fdb9af8c10.wav
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/AUDIO/ MP3/0004_MP3/objects	S:/Ambiente_Treinamento/AIP/ ef3c/4fd9/6a5f/48d6/a5e6/1e0 3/223c/0765
File Size	254 MB	1121 MB
File Modification Date/Time	2015:07:16 12:13:15-03:00	2025:02:27 23:49:48-03:00
File Access Date/Time	2015:07:16 12:13:15-03:00	2025:02:27 23:49:48-03:00
File Creation Date/Time	2025:02:27 19:16:15-03:00	2025:02:27 23:50:04-03:00
File Type	MP3	WAV
File Type Extension	MP3	WAV
Audio Channels		2
Audio Sample Rate	44100	44100
Duration	1:45:50 (approx)	1:45:55
Bits Per Sample	-	16

Fonte: A autora (2024).

Assim como observado, no texto acima, os Quadros 17 a 30, também apresentam dados coletados nesta etapa de pesquisa, para este parágrafo vamos explorar os dados referentes a imagens matriciais, com variadas extensões, no quadro abaixo os dados foram organizados da seguinte maneira: na primeira coluna temos as propriedades significativas determinadas pelo projeto Inspect , na segunda coluna temos os metadados do pacote SIP do objeto digital à ser submetido na plataforma de preservação digital e na terceira coluna temos os metadados do AIP, objeto digital já convertido para formato de preservação.

Nos quadros 17 a 20, observa-se a análise da extensão BMP. A ferramenta de extração de metadados não conseguiu identificar os campos “*Bit per sample*”, “*Samples per pixel*” e “*Extras samples*” no SIP, enquanto, no mesmo, o único dado ausente foi “*Extra sample*”. Notou-se uma diferença nos campos “*X Sampling Frequency*” e “*Y Sampling Frequency*”: no SIP, os valores são inteiros, e no AIP, apresentam unidades decimais. Inicialmente, isso poderia sugerir uma alteração na propriedade significativa do objeto digital. Contudo, ao analisar os metadados do AIP, verifica-se que a resolução está em centímetros (cm), enquanto no SIP a

unidade é por metro quadrado. Portanto, a diferença é apenas na unidade de medida, não implicando alteração na propriedade significativa do objeto digital. Ainda nesses quadros foi possível identificar também que a ferramenta de extração de metadados do AIP, não retorno o dado “amostra extras”, esses dados não foram apresentados em nenhum momento pela ferramenta de extração de metadados.

Quadro 17 - Imagem matricial 0001: extensão .bmp

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_BMP.bmp	0001_BMP-a11ba53d-64b5-4e5a-816d-e5050bb426b7.tif
Directory	S:/Ambiente_Treinamento/SIP/Treinamento/CNV/IMAGE M_MATRICIAL/BMP/0001_BMP/objects	S:/Ambiente_Treinamento/AIP/1e27/5401/85d9/487d/9f83/c2a1/23c6/9896
File Size	12 MB	23 MB
File Modification Date/Time	2012:06:19 16:54:34-03:00	2025:02:25 00:50:06-03:00
File Access Date/Time	2012:06:19 16:54:34-03:00	2025:02:25 00:50:06-03:00
File Creation Date/Time	2025:02:24 18:23:29-03:00	2025:02:25 00:50:21-03:00
File Type	BMP	TIFF
File Type Extension	BMP	TIFF
Largura da imagem Image Width	1654	1654
Altura da imagem Image Height	2337	2337
Frequência de amostragem X X Sampling Frequency	7874	78.73999818
Frequência de amostragem Y Y Sampling Frequency	7874	78.73999818
Bits por amostra Bits per sample		8
Amostras por pixel Samples per pixel		3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 18 - Imagem matricial 0002: extensão .bmp

IMAGEM MATRICIAL 0002 – EXTENSÃO .BMP		
METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_BMP.bmp	0002_BMP-a027b76d-3d28-4195-992f-96a66960230e.tif
Directory	S:/Ambiente_Treinamento/SIP/Treinamento/CNV/IMAGE M_MATRICIAL/BMP/0002_BMP/objects	S:/Ambiente_Treinamento/AIP/5d94/1b44/b91c/4691/8df6/9f4f/f7eb/9502
File Size	39 kB	39 KB
File Modification Date/Time	2014:10:24 09:18:10-03:00	2025:02:25 00:50:16-03:00
File Access Date/Time	2013:12:16 16:01:18-03:00	2025:02:25 00:50:16-03:00
File Creation Date/Time	2025:01:19 15:46:19-03:00	2025:02:25 00:50:31-03:00
File Type	BMP	TIFF
File Type Extension	BMP	TIFF
Largura da imagem Image Width	98	98
Altura da imagem Image Height	130	130
Frequência de amostragem X X Sampling Frequency	2834	28.34000016
Frequência de amostragem Y Y Sampling Frequency	2834	28.34000016
Bits por amostra Bits per sample		8 8 8
Amostras por pixel Samples per pixel		3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 19 - Imagem matricial 0003: extensão .bmp

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0003_BMP.bmp	0003_BMP-00de09e2-1bef-46a1-acbe-bb3eeaaa1621.tif
Directory	S:/Ambiente_Treinamento/SIP/Treinamento/CNV/IMAGE M_MATRICIAL/BMP/0003_BMP/objects	S:/Ambiente_Treinamento/AIP/da69/496b/d86b/478f/aa17/2d76/a757/0c24
File Size	12 MB	243 MB

File Modification Date/Time	2012:06:19 16:54:40-03:00	2025:02:25 00:54:01-03:00
File Access Date/Time	2012:06:19 16:54:40-03:00	2025:02:25 00:54:01-03:00
File Creation Date/Time	2025:02:27 17:48:29-03:00	2025:02:25 00:54:48-03:00
File Type	BMP	TIFF
File Type Extension	BMP	TIFF
Largura da imagem Image Width	1654	1654
Altura da imagem Image Height	2337	2337
Frequência de amostragem X X Sampling Frequency	7874	78.73999818
Frequência de amostragem Y Y Sampling Frequency	7874	78.73999818
Bits por amostra Bits per sample		8 8 8
Amostras por pixel Samples per pixel		3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 20 - Imagem matricial 0004: extensão .bmp

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0004_BMP.bmp	0004_BMP-265a6a60-b4c8-490e-b0ba-a89627626b32.tif
Directory	S:/Ambiente_Treinamento/SIP/Tr e i n a m e n t o /C N V /i m a g e m _m a t r i c i a l /b m p /0004_bmp/o b j e c t s	S:/Ambiente_Treinamento/AIP/ 7547/9506/2e12/4172/8afc/c5c 8/432b/1494
File Size	12 MB	23 MB
File Modification Date/Time	2012:06:19 16:54:42-03:00	2025:02:27 20:15:51-03:00
File Access Date/Time	2012:06:19 16:54:42-03:00	2025:02:27 20:15:51-03:00
File Creation Date/Time	2012:06:19 16:54:42-03:00	2025:02:27 20:16:39-03:00
File Type	BMP	TIFF
File Type Extension	BMP	TIFF
Largura da imagem Image Width	1654	1654

Altura da imagem Image Height	2337	2337
Frequência de amostragem X X Sampling Frequency	7874	78.73999818
Frequência de amostragem Y Y Sampling Frequency	7874	78.73999818
Bits por amostra Bits per sample		8 8 8
Amostras por pixel Samples per pixel		3
Amostras extras Extras samples		

Fonte: A autora (2024).

Nos quadros de 21 a 30, analisam-se as extensões .psd, .tiff e .jpeg. A ferramenta de extração de metadados não conseguiu retornar o campo "Extras samples" nos metadados referentes ao SIP e AIP. Os demais campos apresentaram todas as propriedades significativas com valores iguais, levando esta pesquisa a concluir que, para essas extensões, as propriedades significativas foram todas mantidas.

Quadro 21 - Imagem matricial 0001: extensão .psd

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_PSD.psd	0001_PSD-d6d37db3-3eba-4d99-b570-3ee05bd4215e.tif
Directory	S:/Ambiente_Treinamento/SIP/Treinamento/CNV/IMAGE M_MATRICIAL/PSD/0001_PSD/objects	S:/Ambiente_Treinamento/AIP/ed71/3395/4cc4/45b2/a445/b8f1/da6c/3548
File Size	132 MB	167 MB
File Modification Date/Time	2014:05:29 16:53:00-03:00	2025:02:25 00:53:13-03:00
File Access Date/Time	2014:05:29 16:53:00-03:00	2025:02:25 00:53:13-03:00
File Creation Date/Time	2025:02:24 18:23:30-03:00	2025:02:25 00:53:37-03:00
File Type	PSD	TIFF
File Type Extension	PSD	TIFF
Largura da imagem Image Width	2656	4176
Altura da imagem Image Height	4176	2656

Frequência de amostragem X X Sampling Frequency	72	72
Frequência de amostragem Y Y Sampling Frequency	72	72
Bits por amostra Bits per sample	8 8 8	8 8 8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 22 - Imagem matricial 0002: extensão .psd

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_PSD.psd	0002_PSD-74b39478-2bda-4c81-b090-790034479ee6.tif
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/PSD/0002_P SD/objects	S:/Ambiente_Treinamento/AIP/ 80c6/2e5d/28a1/4798/ae62/53 e6/815a/bddc
File Size	109 MB	134 MB
File Modification Date/Time	2014:10:31 16:05:01-03:00	2025:02:25 00:53:22-03:00
File Access Date/Time	2014:10:31 16:05:01-03:00	2025:02:25 00:53:22-03:00
File Creation Date/Time	2025:02:24 18:23:31-03:00	2025:02:25 00:53:47-03:00
File Type	PSD	TIFF
File Type Extension	PSD	TIFF
Largura da imagem Image Width	4272	4272
Altura da imagem Image Height	2848	2848
Frequência de amostragem X X Sampling Frequency	72	72
Frequência de amostragem Y Y Sampling Frequency	72	72
Bits por amostra Bits per sample	8 8 8	8 8 8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 23 - Imagem matricial 0001: extensão .tiff

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_TIFF.tiff	0001_TIFF.tiff
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/TIFF/0001_TI FF/objects	S:/Ambiente_Treinamento/AIP/ 64bf/23fc/13eb/4c8f/a102/5879 /3657/6011
File Size	865 kB	865 kB
File Modification Date/Time	2014:05:20 12:35:32-03:00	2025:02:25 00:54:49-03:00
File Access Date/Time	2014:05:20 12:35:32-03:00	2025:02:25 00:54:49-03:00
File Creation Date/Time	2025:02:24 18:23:32-03:00	2025:02:25 00:55:09-03:00
File Type	TIFF	TIFF
File Type Extension	TIFF	TIFF
Largura da imagem Image Width	2720	2720
Altura da imagem Image Height	3624	3624
Frequência de amostragem X X Sampling Frequency	300	300
Frequência de amostragem Y Y Sampling Frequency	300	300
Bits por amostra Bits per sample	8 8 8	8 8 8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 24 - Imagem matricial 0002: extensão .tiff

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_TIFF.tiff	0002_TIFF-34e39351-f27a-4734- a326-d4a87f7ed8a9.tif
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/TIFF/0002_TI FF/objects	S:/Ambiente_Treinamento/AIP/ 305b/c0de/3e89/4269/a0ad/6f 56/e191/a033
File Size	1811 kB	59 MB
File Modification Date/Time	2014:05:20 15:08:12-03:00	2025:02:25 00:55:11-03:00

File Access Date/Time	2014:05:20 15:08:12-03:00	2025:02:25 00:55:11-03:00
File Creation Date/Time	2025:02:24 18:23:32-03:00	2025:02:25 00:55:29-03:00
File Type	TIFF	TIFF
File Type Extension	TIFF	TIFF
Largura da imagem Image Width	2720	2720
Altura da imagem Image Height	3624	3624
Frequência de amostragem X X Sampling Frequency	300	300
Frequência de amostragem Y Y Sampling Frequency	300	300
Bits por amostra Bits per sample	8 8 8	8 8 8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 25 - Imagem matricial 0003: extensão .tiff

IMAGEM MATRICIAL 0003 – EXTENSÃO .TIFF		
METADADOS	METADADOS SIP	METADADOS AIP
File Name	0003_tif.tif	0003_TIF.tif
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/imagem _matricial/tiff/0003_tif/objects	S:/Ambiente_Treinamento/AIP/ bc1c/0c8d/25b3/41fe/b1a8/c2c 6/cf22/42d8
File Size	46 kB	237 kB
File Modification Date/Time	2014:01:30 11:30:30-03:00	2025:02:27 19:26:45-03:00
File Access Date/Time	2014:01:30 11:30:30-03:00	2025:02:27 19:26:45-03:00
File Creation Date/Time	2025:02:24 18:23:31-03:00	2025:02:27 19:27:37-03:00
File Type	TIFF	TIFF
File Type Extension	TIFF	TIFF
Largura da imagem Image Width	1648	1648
Altura da imagem Image Height	2336	2336

Frequência de amostragem X X Sampling Frequency	200	200
Frequência de amostragem Y Y Sampling Frequency	200	200
Bits por amostra Bits per sample	1	1
Amostras por pixel Samples per pixel	1	1
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 26 - Imagem matricial 0004: extensão .tiff

IMAGEM MATRICIAL 0004 – EXTENSÃO .TIFF		
METADADOS	METADADOS SIP	METADADOS AIP
File Name	0004_tif.tif	0004_TIF.tif
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/imagem _matricial/tiff/0004_tif/objects	S:/Ambiente_Treinamento/AIP/ ecfc/582e/4f06/4f7e/9752/5e41 /74d1/d0c0
File Size	15 kB	205 kB
File Modification Date/Time	2014:01:23 13:52:18-03:00	2025:02:27 19:26:50-03:00
File Access Date/Time	2014:01:23 13:52:18-03:00	2025:02:27 19:26:50-03:00
File Creation Date/Time	2025:02:24 18:23:31-03:00	2025:02:27 19:27:38-03:00
File Type	TIFF	TIFF
File Type Extension	TIFF	TIFF
Largura da imagem Image Width	1648	1648
Altura da imagem Image Height	2336	2336
Frequência de amostragem X X Sampling Frequency	200	200
Frequência de amostragem Y Y Sampling Frequency	200	200
Bits por amostra Bits per sample	1	1
Amostras por pixel Samples per pixel	1	1

Amostras extras Extras samples		
-----------------------------------	--	--

Fonte: A autora (2024).

Quadro 27 - Imagem matricial 0001 – extensão .jpeg

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0001_JPEG.jpeg	0001_JPEG.jpeg
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/JPEG/0001_J PEG/objects	S:/Ambiente_Treinamento/AIP/ 77fc/af4f/0316/4e51/9dae/a8a 9/3f24/792f
File Size	27 kB	273 kB
File Modification Date/Time	2014:05:13 18:16:24-03:00	2025:02:25 00:51:22-03:00
File Access Date/Time	2014:05:13 18:16:24-03:00	2025:02:25 00:51:22-03:00
File Creation Date/Time	2025:02:24 18:23:30-03:00	2025:02:25 00:51:22-03:00
File Type	JPEG	JPEG
File Type Extension	JPEG	JPEG
Largura da imagem Image Width	370	370
Altura da imagem Image Height	270	270
Frequência de amostragem X X Sampling Frequency	72	72
Frequência de amostragem Y Y Sampling Frequency	72	72
Bits por amostra Bits per sample	8	8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 28 - Imagem matricial 0002: extensão .jpeg

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0002_JPEG.jpeg	0002_JPEG.jpeg
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/JPEG/0002_J PEG/objects	S:/Ambiente_Treinamento/AIP/ 1465/d15e/1b26/499a/a1fd/8a de/83f7/60d6
File Size	29 Kb	275 kB

File Modification Date/Time	2014:05:13 18:16:24-03:00	2025:02:25 00:51:34-03:00
File Access Date/Time	2014:05:13 18:16:24-03:00	2025:02:25 00:51:34-03:00
File Creation Date/Time	2025:02:24 18:23:30-03:00	2025:02:25 00:52:20-03:00
File Type	JPEG	JPEG
File Type Extension	JPEG	JPEG
Largura da imagem Image Width	240	240
Altura da imagem Image Height	415	415
Frequência de amostragem X X Sampling Frequency	100	100
Frequência de amostragem Y Y Sampling Frequency	100	100
Bits por amostra Bits per sample	8	8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 29 - Imagem matricial 0003: extensão .jpeg

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0003_JPEG.jpeg	0003_JPEG.jpg
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/IMAGE M_MATRICIAL/JPEG/0003_J PEG/objects	S:/Ambiente_Treinamento/AIP/ 5edd/344c/8446/4126/9c52/f9b 5/9486/1bf8
File Size	166 kB	413 kB
File Modification Date/Time	2014:03:27 18:19:00-03:00	2025:02:27 19:25:34-03:00
File Access Date/Time	2014:03:27 18:19:00-03:00	2025:02:27 19:25:34-03:00
File Creation Date/Time	2025:02:24 18:23:30-03:00	2025:02:27 19:26:44-03:00
File Type	JPEG	JPEG
File Type Extension	JPEG	JPEG
Largura da imagem Image Width	863	863

Altura da imagem Image Height	1200	1200
Frequência de amostragem X X Sampling Frequency	1	1
Frequência de amostragem Y Y Sampling Frequency	1	1
Bits por amostra Bits per sample	8	8
Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

Fonte: A autora (2024).

Quadro 30 - Imagem matricial 0004: extensão .jpeg

METADADOS	METADADOS SIP	METADADOS AIP
File Name	0004_JPEG.jpeg	0004_JPEG.jpg
Directory	S:/Ambiente_Treinamento/SI P/Treinamento/CNV/imagem _matricial/JPEG/0004_JPEG/ objects	S:/Ambiente_Treinamento/AIP/ 1574/513e/2ed8/4155/baa0/dc 19/e273/ad51
File Size	183 kB	431 Kb
File Modification Date/Time	2014:03:27 18:19:38-03:00	2025:02:27 19:25:58-03:00
File Access Date/Time	2014:03:27 18:19:38-03:00	2025:02:27 19:25:58-03:00
File Creation Date/Time	2025:02:24 18:23:30-03:00	2025:02:27 19:26:47-03:00
File Type	JPEG	JPEG
File Type Extension	JPEG	JPEG
Largura da imagem Image Width	864	864
Altura da imagem Image Height	1200	1200
Frequência de amostragem X X Sampling Frequency	1	1
Frequência de amostragem Y Y Sampling Frequency	1	1
Bits por amostra Bits per sample	8	8

Amostras por pixel Samples per pixel	3	3
Amostras extras Extras samples		

FONTE: A autora (2024).

Esta pesquisa, também optou por manter a data de criação desse objeto, para mostrar que esse documento foi criado em uma data distante da data de realização dessa pesquisa, como marco para que se um dia esse objeto nessa analisado novamente poderá comparar se houve alguma alteração.

8 CONSIDERAÇÕES FINAIS

A preservação da autenticidade dos documentos arquivísticos digitais é uma questão essencial e contemporânea para a Arquivologia. A fragilidade dos suportes, as características dos objetos digitais e a obsolescência tecnológica apresentam desafios significativos para a conservação a longo prazo e para a manutenção de sua autenticidade.

Esses documentos, gerados por órgãos e entidades como resultado de suas atividades, precisam de tratamento técnico adequado desde sua criação até a destinação final. O foco desta pesquisa é o recolhimento desses documentos. Para aqueles que atingem a idade permanente e necessitam de preservação, os procedimentos devem seguir as diretrizes estabelecidas pelas autoridades arquivísticas competentes.

Portanto, os órgãos e entidades produtoras devem adotar procedimentos e critérios para a criação, uso e manutenção desses documentos em um ambiente seguro. Aspectos como os sistemas utilizados, a organização dos documentos, a gestão documental, a padronização dos elementos da forma documental nesses sistemas e o registro de metadados precisam ser considerados para garantir a máxima presunção de autenticidade desses documentos. Neste sentido, a pesquisa conseguiu alcançar a proposta de verificar as propriedades significativas dos objetos digitais recolhidos pelo Arquivo Nacional, com base na metodologia do projeto Inspect, e identificar que esses objetos têm suas propriedades significativas preservadas, contribuindo para a manutenção da presunção de autenticidade do objeto digital.

Entretanto, essa fase não foi fácil, e muitos altos e baixos interpelaram a pesquisa. Considerando essa condição, esta pesquisa deixa algumas recomendações que podem ser executadas para facilitar a análise das propriedades significativas de forma mais transparente. Pretende-se propor melhorias e direcionamentos para pesquisas futuras.

Essa pesquisa justifica-se pela escassez de estudos que abordem a comparação das propriedades significativas de objetos digitais antes e depois de sua submissão a uma plataforma de preservação e conversão para formatos de longo prazo.

Sugere-se a criação de perfis que auxiliem na decisão sobre quais propriedades significativas são importantes para cada tipo de documento ou grupo de documentos, com base na busca de informações pelos usuários. Outra possibilidade é a criação de grupos de trabalho que façam estudos das extensões existentes, baseados nos formatos oficiais de objetos digitais que serão recolhidos pelo Arquivo Nacional a partir da Administração Pública Federal, com o

intuito de identificar amplamente quais são e como se comportam os metadados de cada extensão.

Dessa forma, conclui-se esta dissertação com a esperança de ter contribuído para a preservação dos documentos arquivísticos digitais e incentivado a realização de trabalhos futuros sugeridos.

REFERÊNCIAS

ARQUIVO NACIONAL (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005. (Publicações Técnicas, 51).

_____. **Política de Preservação Digital**. 2. ed. Rio de Janeiro: Arquivo Nacional, 2016.

BRASIL. Lei nº 8159 de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Brasília, DF: Diário Oficial da União, 1991. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/18159.htm. Acesso em: 15 fev. 2025

CAMPOS, F. M. Informação digital: um novo patrimônio a preservar. **Cadernos BAD**, [S.l.] n. 2, 2002. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/62432>. Acesso em: 01 mar. 2024.

CONARQ. **Carta para a preservação do patrimônio arquivístico digital**. Rio de Janeiro, 2004. Disponível em: <https://shre.ink/eAcj>. Acesso em: 22 fev. 2018.

_____. Diretrizes para a Implementação de Repositórios Arquivísticos Digitais. Rio de Janeiro: Conarq, 2012. Disponível em: https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_presuncao_autenticidade_completa.pdf. Acesso em: 21 mar. 2024.

_____. **Glossário de Documentos Arquivísticos Digitais**. 8. ed. Rio de Janeiro: Conarq, 2020. Disponível em: https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf. Acesso em: 21 mar. 2024

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEM. **Reference model for an open archival information system (OAIS): recommended practice**. Washington: CCSDS Secretariat, 2012. Disponível em: <http://public.ccsds.org/publications/archive/650x0m2.pdf>. Acesso em: 14 mar. 2024.

DUCKER, P. Além da revolução da informação. **HSM Management**, [São Paulo], v. 4, n. 18, jan./fev. 2000. Disponível em: https://ulhoa.net/Leituras/Alem_da_revolucao_da_informacao.pdf. Acesso em: 14 mar. 2024.

FERREIRA, M. **Introdução à preservação digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade de Minho, 2006.

FORMENTON, D. *et al.* Os padrões de metadados como recursos tecnológicos para a garantia da preservação digital. **Biblios**, Pittsburgh, n. 68, p. 82-95, jul. 2017. Disponível em: <http://www.scielo.org.pe/pdf/biblios/n68/a06n68.pdf>. Acesso em: 13 maio 2025.

GAVA, T. B. S.; FLORES, D. O papel do Archivematica no RDC-Arq e possíveis cenários de uso. **Ágora: Arquivologia em debate**, Trindade, SC, v. 31, n. 63, p. 1–21, 2021. Disponível em: <https://agora.emnuvens.com.br/ra/article/view/1018>. Acesso em: 14 maio. 2025.

GRACE, S.; KNIGHT, G.; MONTAGUE, L. **Inspect**: Final Report v. 1.0. [S.l.: s.n.], 2009.

HOBSBAWM, E. **Da revolução industrial inglesa ao imperialismo**. 5. ed. Rio de Janeiro: Forense Universitária, 2000.

INDOLFO, A. C. **Dimensões político-arquivísticas da avaliação de documentos na Administração Pública Federal (2004-2012)**. 2013. 312 f. Tese (Doutorado em Ciência da Informação). – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.

INTERPARES TRUST. **Preservation as a Service for Trust (PaaST)**: functional and data requirements for digital preservation, versão 1.0. [S.l.: s.n.], 2017. Disponível em: http://interparestrust.org/assets/public/dissemination/PreservationasaServiceforTrust1_0.pdf . Acesso em: 12 maio 2024.

INTERPARES 2 Project. **Diretrizes do produtor**: a elaboração e a manutenção de materiais digitais: diretrizes para indivíduos. Rio de Janeiro: Arquivo Nacional, [2011a].

_____. **Diretrizes do preservador**: a preservação de documentos arquivísticos digitais: diretrizes para organizações. Rio de Janeiro: Arquivo Nacional, [2011b].

INVESTIGATING SIGNIFICANT PROPERTIES OF ELECTRONIC CONTENT OVER TIME (INSPECT). Londres: Kings College London, 2010. Disponível em: <https://significantproperties.kdl.kcl.ac.uk/index.html> . Acesso em: 20 mar. 2024.

KNIGHT, G. **Inspect Framework Report**: v. 2.1. [S.l.: s.n.], 2009. Disponível em: https://www.webarchive.org.uk/wayback/archive/20130423072330mp_/http://www.significantproperties.org.uk/Inspect-framework.pdf . Acesso em: 20 mar. 2024.

_____. **Significant properties testing report**: audio recordings: v1.1. [S.l.: s.n.], 2010. Disponível em: <https://significantproperties.kdl.kcl.ac.uk/audio-testingreport.pdf> . Acesso em: 04 jan. 2024.

_____. **Significant properties testing report**: electronic Mail. [S.l.: s.n.], 2009. Disponível em: <https://significantproperties.kdl.kcl.ac.uk/email-testingreport.html> . Acesso em: 04 jan. 2024.

LACOMBE, C.; SILVA, M. da. Padrões para Garantir a Preservação e o Acesso aos Documentos Digitais. **Acervo**, [S. l.], v. 20, n. 1-2, p. 113–124, 2007. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/76> . Acesso em: 20 fev. 2025.

MÁRDERO ARELLANO, M. A. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <https://shre.ink/eAa4> . Acesso em: 05 fev. 2025.

_____. **Critérios para a preservação digital da informação científica**. 2008. 356 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2008. Disponível em: <http://repositorio.unb.br/handle/10482/1518> . Acesso em: 23 dez. 2023.

MONTAGUE, L. **Significant properties testing report**: Raster Images. [S.l.: s.n.], 2009. Disponível em: <https://significantproperties.kdl.kcl.ac.uk/rasterimages-testingreport.html>. Acesso em: 04 jan. 2024.

ONLINE-METADATA. **Visualizador online de metadados e exif**. [S.l.: s.n.], c2025. Disponível em: <https://online-metadata.com/pt>. Acesso em: 27 dez. 2024.

PREMIS EDITORIAL COMMITTEE. **PREMIS Data Dictionary for Preservation Metadata**. 2012. Disponível em: <http://www.loc.gov/standards/premis/v2/premis-2-2.pdf>. Acesso em: 10 mar. 2024.

ROCCO, B. C. B. **A preservação de documentos em ambiente digital**: contribuições da teoria social na ampliação da abordagem técnica. 2021. 213 f. Tese (Doutorado em Ciência da Informação) – Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, 2021. Disponível em: <https://ridi.ibict.br/handle/123456789/1163>. Acesso em: 15 mar. 2024.

RONDINELLI, R. C. **O documento arquivístico ante a realidade digital**: uma revisão conceitual necessária. Rio de Janeiro: FGV, 2013.

SANTA CATARINA. Tribunal de Justiça. **Glossário de segurança da informação**. Florianópolis: PJSC, [202-?]. Disponível em: www.tjsc.jus.br/web/tecnologia-da-informacao/glossario-de-seguranca-da-informacao. Acesso em: 05 fev. 2025.

SAYÃO, L. F. **Preservação digital no contexto das bibliotecas digitais**: uma breve introdução. Bibliotecas digitais: saberes e práticas. Salvador: UFBA, 2005. Disponível em: [Preservao_digital_no_contexto_das_biblio20211014-18652-1i82zgo.pdf](#). Acesso em: 16 dez. 2024.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. Enc. **Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 15, n. 30, p.1-31, 2010. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2010v15n30p1>. Acesso em: 15 jan. 2024.

SCHLLENBERG, T. R. **Arquivos modernos**: princípios e técnicas. 4. ed. Rio de Janeiro: FGV, 2006.

SILVA, A. M. *et al.*(Orgs.) **Arquivística**: teoria e prática de uma Ciência da Informação. 3. ed. Porto: Edições Afrontamento, 2009. (Biblioteca das Ciências Sociais; 2).

RUIZ RODRIGUÉZ, A. A. **Manual de arquivística**. Madrid: Síntesis, 1995

TAKAHASHI, T. (Org.). **Sociedade da informação no Brasil**: livro verde. Brasília: Ministério da Ciência e Tecnologia, 2000. xxv, 195 p. Disponível em: https://repositorio.mcti.gov.br/bitstream/mctic/6501/1/2000_sociedade_da_informacao_no_brasil_livro_verde.pdf. Acesso em: 26 jan. 2024.

TAVARES, M. F. D. Preservação digital: entre a memória e a história. **Ciência da Informação**, Brasília, v. 41, n. 1, p. 9-21, 2012. Disponível em: <http://revista.ibict.br/ciinf/article/view/1348> . Acesso em: 24 jan. 2024.

THE INTERPARES 2 PROJECT. **Diretrizes do produtor:** a elaboração e a manutenção de materiais digitais. Diretrizes para Indivíduos. Rio de Janeiro: Arquivo Nacional, [2013].

THE INTERPARES 3 PROJECT. Team Canada. **InterPARES 2 Project Chain of Preservation (COP) Model Metadata.**[*S.l.: s.n.*], 2019.

SOUSA, R. T. B. Os princípios arquivísticos e o conceito de classificação. *In:* RODRIGUES, G. M.; ILZA, L. L (Orgs). **Organização e Representação do Conhecimento.** [*S.l.: s.n.*]: 2003. <https://core.ac.uk/download/pdf/33529652.pdf> Acesso em: 20 fev. 2025

SANTOS, H. M; FLORES, D. Las vulnerabilidades de los documentos digitales: Obsolescencia tecnológica y ausencia de políticas y prácticas de preservación digital. **Biblios Journal of Librarianship and Information Science**, [*S. l.*], n. 59, p. 45–54, 2015. Disponível em: <https://biblios.pitt.edu/ojs/biblios/article/view/215>. Acesso em: 22 maio 2025.