



UNIVERSIDADE DE BRASÍLIA – UnB
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
DOUTORADO EM CIÊNCIA DE INFORMAÇÃO

THAYANE MORAIS DE ALENCAR

**Arquitetura de um repositório digital confiável para os acervos
no Tainacan: proposição de um modelo de aplicação**

Brasília

2026

Thayane Moraes de Alencar

**Arquitetura de um repositório digital confiável para os acervos
no Tainacan: proposição de um modelo de aplicação**

Defesa de tese apresentada ao
Curso de Doutorado em Ciência da
Informação do Programa de Pós-
graduação em Ciência da Informação
da Universidade de Brasília como
requisito para obtenção do grau de
Doutora em Ciência da Informação

Orientador: Prof. Dr. Dalton Lopes
Martins

Brasília

2026

Thayane Moraes de Alencar

Arquitetura de um repositório digital confiável para os acervos no Tainacan: proposição de um modelo de aplicação

Defesa de tese apresentada ao Curso de Doutorado em Ciência da Informação do Programa de Pós-graduação em Ciência da Informação da Universidade de Brasília como requisito para obtenção do grau de Doutora em Ciência da Informação.

Brasília, 15 de março de 2026.

BANCA EXAMINADORA

Cynthia Roncaglio
Universidade de Brasília UnB/FCI

Ivina Flores
Ministério da Gestão e Inovação em Serviços Públicos

Tiago Emmanuel Nunes Braga
Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Alencar, Thayane Moraes de

A368

Arquitetura de um repositório digital confiável para os acervos no Tainacan: proposição de um modelo de aplicação/ Thayane Moraes de Alencar; orientador Dalton Lopes Martins. -- Brasília, DF: Faculdade de Ciência da Informação, 2026.

133 p.

Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, 2026.

1. repositório digital confiável. 2. preservação digital 3. *open archival information system*. 4. *Trust worthy repositories audit and certification* 5. repositório arquivístico digital confiável. I. Martins, Dalton Lopes, orient. II. Título.

DEDICATÓRIA

Dedico esta tese a toda criança que sonha grande e busca, com todas as forças, mudar o seu destino.

Aos professores de escola pública, em especial aos que foram meus mestres, por terem plantado em meu coração o prazer de estudar e de transformar a minha vida por meio disso.

À minha família, sem a qual eu nada seria.

AGRADECIMENTOS

Agradeço à Universidade de Brasília e aos professores da Faculdade de Ciência da Informação pela dedicação ao ensino público. Sem eles, esta tese não seria possível.

Para a criança que eu fui.

À mulher que me tornei.

O mar revolto fez de nós boa marinheira.

RESUMO

Esta tese aborda o desafio preservação digital em instituições museológicas, propondo um modelo de aplicação arquitetônica que integra a plataforma Tainacan a um ecossistema de preservação digital, tendo em vista a crescente digitalização de acervos e a necessidade de garantir sua perenidade e autenticidade. O objetivo principal é formalizar um modelo de aplicação que articule o Tainacan, enquanto sistema de difusão, com requisitos de preservação digital, utilizando o Archivematica como *software* aplicador dos requisitos. A metodologia qualitativa adota parte da análise das funcionalidades do Tainacan e as correlaciona com os requisitos funcionais do modelo de referência *Open Archival Information System* (OAIS) e com os critérios de auditoria do *Trustworthy Repositories Audit & Certification* (TRAC). O resultado da pesquisa é a proposição de um modelo de aplicação de requisitos de preservação ao Tainacan. Atuando como a camada de acesso, responsável pela disponibilização do Pacote de Informação de Disseminação (DIP) e pela interface com o usuário. A camada de preservação, por sua vez, utiliza um sistema dedicado para a ingestão do Pacote de Informação de Arquivamento (AIP) produzido pelo Archivematica. Concluiu-se que a implementação do modelo proposto permite que acervos gerenciados no Tainacan alcancem um nível de confiabilidade arquivística compatível com padrões internacionais. A arquitetura soluciona uma lacuna funcional ao especializar as responsabilidades de acesso e preservação, alinhando-se aos princípios teóricos do OAIS, como a separação de funções para garantia da autenticidade e a gestão contínua de registros para a preservação do longo prazo.

Palavras-chave: repositório digital confiável, Tainacan, preservação digital, open archival information system, repositório arquivístico digital confiável.

ABSTRACT

This thesis addresses the challenge of digital preservation in museum institutions, proposing an architectural application model that integrates the Tainacan platform into a digital preservation ecosystem. This is in light of the increasing digitization of collections and the need to guarantee their permanence and authenticity. The main objective is to formalize an application model that articulates Tainacan, as a dissemination system, with digital preservation requirements, using Archivematica as the software to apply these requirements. The qualitative, applied methodology adopted begins with an analysis of Tainacan's functionalities and correlates them with the functional requirements of the Open Archival Information System (OAIS) reference model and the audit criteria of the Trustworthy Repositories Audit & Certification (TRAC). The research results in the proposition of a model for applying preservation requirements to Tainacan, acting as the access layer, responsible for making the Dissemination Information Package (DIP) available and for the user interface. The preservation layer, in turn, uses a dedicated system for ingesting the Archival Information Package (AIP) produced by Archivematica. It was concluded that the implementation of the proposed model allows collections managed in Tainacan to achieve a level of archival reliability compatible with international standards. The architecture solves a functional gap by specializing access and preservation responsibilities, aligning with OAIS theoretical principles, such as separation of duties to guarantee authenticity and continuous record management for long-term preservation.

Keywords: trusted digital repository, Tainacan, digital preservation, open archival information system, trustworthy repositories audit and certification , , repositórios digitais arquivísticos confiáveis.

LISTA DE QUADROS

QUADRO 1 - ELEMENTOS DA CADEIA DE CUSTÓDIA DIGITAL ARQUIVÍSTICA.	32
QUADRO 2 - DESCRIÇÃO DAS ENTIDADES FUNCIONAIS DO OAIS	64
QUADRO 3 - REQUISITOS ESSENCIAIS PARA QUE O TAINACAN SE TORNE UMA ENTIDADE DE ACESSO COM BASE NO OAIS.	65
QUADRO 4 - ASPECTOS DA OPERAÇÃO DE UM REPOSITÓRIO CONFIÁVEL.	67
QUADRO 5 - ESTRUTURA MACRO DO <i>CHECKLIST</i> TRAC	70
QUADRO 6 - ETAPAS DO FLUXO DE INTEGRAÇÃO ENTRE ARCHIVEMATICA E ATOM....	83
QUADRO 7 - ANÁLISE COMPARATIVA DOS MÉTODOS DE INTEGRAÇÃO.....	88
QUADRO 8 - ANÁLISE DE ADERÊNCIA DO ECOSISTEMA TAINACAN AOS REQUISITOS DA SEÇÃO A DO TRAC.	114
QUADRO 9 - SÍNTESE DE ADERÊNCIA DO ECOSISTEMA TAINACAN AOS REQUISITOS TRAC.	118
QUADRO 10 - METADADOS DESCRITIVOS (TAINACAN - PINACOTECA).....	126
QUADRO 11 - QUADRO DE <i>CROSSWALK</i> : TAINACAN (PINACOTECA) PARA PREMIS (ENTIDADE OBJETO).	127
QUADRO 12 - LÓGICA DO <i>SCRIPT</i> DE RECEBIMENTO E VERIFICAÇÃO DO DIP.....	134
QUADRO 13 - REPLICAÇÃO DOS REQUISITOS 3,4,5 E 6 DO QUADRO 12.	138
QUADRO 14 - QUADRO COMPARATIVO E RECOMENDAÇÃO	139

LISTA DE FIGURAS

FIGURA 1 - MODELO DE CADEIA DE CUSTÓDIA DE PRESERVAÇÃO CONTÍNUA (BASEADO NO OAIS)	35
FIGURA 2 - AMBIENTE OAIS E SUAS ENTIDADES.	62
FIGURA 3 - PACOTES DE INFORMAÇÃO NO MODELO OAIS	63
FIGURA 4 - ENTIDADES FUNCIONAIS DO MODELO OAIS	64
FIGURA 5 - PROCESSO DE AUDITORIA E CERTIFICAÇÃO TRAC.	68
FIGURA 6 - ESTRUTURA DE INTEGRAÇÃO ENTRE O ARCHIVEMATICA E O ATOM SIMPLIFICADA.	83
FIGURA 7 - DIAGRAMA DE ARQUITETURA DOS REQUISITOS TRAC APLICÁVEIS AO TAINACAN.	120
FIGURA 8 - ESQUEMA DE PRESERVAÇÃO DE PACOTES E SISTEMAS ABRANGENDO O TAINACAN	132

LISTA DE ABREVIATURAS E SIGLAS

ABNT – Associação Brasileira de Normas Técnicas

AIP – Archival Information Package

API – Application Programming Interface (Interface de Programação de Aplicações)

ASGI – Asynchronous Server Gateway Interface

AtoM – Access to Memory

CI/CD – Continuous Integration / Continuous Deployment (Integração Contínua / Entrega Contínua)

CIDOC CRM – CIDOC Conceptual Reference Model

CONARQ – Conselho Nacional de Arquivos

CORS – Cross-Origin Resource Sharing

CSS – Cascading Style Sheets

DIP – Dissemination Information Package

DOM – Document Object Model

DPI – Dots Per Inch

e-ARQ Brasil – Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos

ETL – Extract, Transform, Load (Extrair, Transformar, Carregar)

FOAF – Friend of a Friend

GED – Gerenciamento Eletrônico de Documentos

GUID – Globally Unique Identifier

HTML – HyperText Markup Language

HTTP – Hypertext Transfer Protocol

I/O – Input/Output (Entrada/Saída)

ICA – International Council on Archives

ICOM – International Council of Museums

IETF – Internet Engineering Task Force

ISO – International Organization for Standardization

JPEG – Joint Photographic Experts Group

JSON – JavaScript Object Notation

JWT – JSON Web Token

LAM – Libraries, Archives and Museums

LTO – Linear Tape-Open

ML – Machine Learning

METS – Metadata Encoding and Transmission Standard

MPEG – Moving Picture Experts Group

MVC – Model-View-Controller

MVP – Minimum Viable Product (Produto Mínimo Viável)

NLP – Natural Language Processing (Processamento de Linguagem Natural)

NOBRADE – Norma Brasileira de Descrição Arquivística

OAIS – Open Archival Information System

ORM – Object-Relational Mapper (Mapeador Objeto-Relacional)

OWASP – Open Web Application Security Project

PDF/A – Portable Document Format / Archival

PHP – PHP: Hypertext Preprocessor

PNG – Portable Network Graphics

PREMIS – Preservation Metadata: Implementation Strategies

RDC-Arq – Repositório Arquivístico Digital Confiável

RDF – Resource Description Framework

REST – Representational State Transfer

RFC – Request for Comments

SHA-1 – Secure Hash Algorithm 1

SIGAD – Sistema Informatizado de Gestão Arquivística de Documentos

SIP – Submission Information Package

SKOS – Simple Knowledge Organization System

SQL – Structured Query Language

TDD – Test-Driven Development (Desenvolvimento Orientado a Testes)

TI – Tecnologia da Informação

TRAC – Trustworthy Repositories Audit & Certification

UnB – Universidade de Brasília

URI – Uniform Resource Identifier

URL – Uniform Resource Locator

URN – Uniform Resource Name

UUID – Universally Unique Identifier

W3C – World Wide Web Consortium

WSGI – Web Server Gateway Interface

XML – eXtensible Markup Language

XSS – Cross-Site Scripting

SUMÁRIO

1 INTRODUÇÃO	16
1.1 Contextualização e cenário da preservação digital	18
1.2 Problema e justificativa da pesquisa	24
1.3 Relevância do estudo	26
1.4 Objetivo geral da tese	28
1.5 Objetivos específicos	28
1.6 Pergunta de pesquisa	28
2. REFERENCIAL TEÓRICO: PRESERVAÇÃO DIGITAL E DIFUSÃO	29
2.1 Preservação digital: conceitos fundamentais e desafios	30
2.2 O estado da arte da preservação digital em museus: uma revisão sistemática da literatura.....	35
2.2.3 Conclusão dos trabalhos correlatos	59
2.3 O Modelo OAIS (ISO 14721) e seus omponentes	61
2.4 O Modelo Trustworthy Repository Audit and Certification (TRAC).....	66
2.5 <i>Softwares</i> de aplicação do OAIS: análise comparativa de casos	72
2.6 Archivemática como <i>software</i> de aplicação do modelo OAIS: Análise da arquitetura de integração com o <i>software</i> de difusão (AtoM)	79
2.7 Aplicação do OAIS a <i>softwares</i> de difusão: Análise Comparativa de métodos de integração.	86
2 METODOLOGIA.....	89
3.1 Natureza e abordagem da pesquisa	90
3.2 Etapas de execução da pesquisa	92
3.3 Instrumentos de coleta e análise de dados	98
3.4 Considerações éticas e uso de ferramentas de Inteligência Artificial (IA)	100
4. O TAINACAN COMO OBJETO DE ESTUDO: ESTRUTURA E LIMITES	
101	
4.1 Descrição técnica do Tainacan	109

4.2	Análise da estrutura sociotécnica e aplicação dos requisitos TRAC/OAIS ao Tainacan.....	112
4.3	Modelagem de metadados e <i>crosswalk</i>	124
5.	PROPOSIÇÃO DE REQUISITOS DE PRESERVAÇÃO DIGITAL APLICÁVEIS AO TAINACAN.....	130
5.1	Arquitetura conceitual do modelo proposto.....	131
5.2	Detalhamento dos componentes e fluxos de preservação	133
5.3	Implicações e recomendações para o Ibram	142
6	CONCLUSÃO	144
	REFERÊNCIAS.....	148

1 INTRODUÇÃO

A preservação digital emergiu, na década de 1990, como uma preocupação central no campo da Ciência da Informação, à medida que pesquisadores passaram a reconhecer a vulnerabilidade dos suportes digitais e a rápida obsolescência de formatos e tecnologias. Hedstrom (2000) foi uma das primeiras autoras a demonstrar que a perda de objetos digitais constituía um risco estrutural para preservação de informações a longo prazo, iniciando debates que posteriormente se tornariam fundamentais para a consolidação do tema. A discussão ultrapassou a esfera técnica, ganhando contornos sociotécnicos à medida que autores como Conway (2010) argumentaram que a preservação digital envolve não apenas estratégias de migração e manutenção de sistemas, mas também práticas institucionais, políticas organizacionais e decisões de gestão que condicionam a continuidade e a autenticidade da informação.

A consolidação conceitual da preservação digital levou, ainda no final da década de 1990, ao desenvolvimento de esforços internacionais para padronizar terminologias, funções e responsabilidades envolvidas na manutenção de objetos digitais. Nesse contexto, a NASA, por meio do *Consultative Committee for Space Data Systems* (CCSDS), iniciou a formulação do *Open Archival Information System* (OAIS), modelo concebido originalmente para atender às necessidades da comunidade espacial, que enfrentava dificuldades crescentes na preservação de dados científicos produzidos em larga escala. Sua primeira versão completa, publicada em 2002 e posteriormente normatizada como International Organization for Standardization (ISO) nº 14721 de 2012, estabeleceu um marco estrutural ao definir um conjunto de funções essenciais como: ingestão, armazenamento arquivístico, gestão de dados, administração, planejamento da preservação e acesso, bem como, ao propor um vocabulário comum para orientar repositórios digitais de longo prazo (CCSDS, 2012). A criação do OAIS responde, portanto, à necessidade de estabilizar concepções e práticas em um campo que se expandia rapidamente, oferecendo um referencial aplicável a diferentes cenários, mas originado de um ambiente altamente especializado.

Embora o OAIS tenha sido adotado como modelo global, diferentes autores apontam que sua construção teórica e sua interpretação prática têm sido fortemente influenciadas pelo campo arquivístico. Harvey e Weatherburn (2019) destacam que a ênfase do modelo em processos de autenticidade, proveniência e cadeia de custódia encontra maior aderência às tradições da arquivologia do que às práticas de bibliotecas e museus, cujas necessidades de preservação e interpretação nem sempre se alinham aos pressupostos do OAIS. Corrado e Moulaison Sandy (2017) reforçam essa crítica ao observar que, apesar de ser amplamente utilizado como referência central, o OAIS permanece mais robusto para ambientes arquivísticos estruturados do que para ecossistemas museológicos, que lidam com objetos multimídia, coleções heterogêneas, vínculos interpretativos e metadados curatoriais específicos. Esse desalinhamento gerou, ao longo dos anos, uma assimetria: a literatura e os padrões de certificação evoluíram de modo mais acelerado nos arquivos, enquanto bibliotecas e museus buscam adaptações próprias ou estratégias híbridas para suprir lacunas funcionais, ou não discutem sobre o tema.

O amadurecimento teórico do tema evidenciou a complexidade crescente dos ecossistemas digitais, que frequentemente envolvem múltiplas plataformas, fluxos distribuídos de informação e interdependências tecnológicas. Zierau e Schultz (2013) destacam que os sistemas contemporâneos raramente são estruturas monolíticas: sua confiabilidade depende de arquiteturas integradas, de rotinas formais de preservação e de mecanismos contínuos de monitoramento e avaliação. Becker *et al.* (2009) reforçam esse argumento ao demonstrar que abordagens de preservação precisam considerar camadas de descrição, armazenamento, acesso e governança, articuladas de forma coerente para evitar perdas e garantir autenticidade ao longo do tempo.

Padrões como o *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC) (CRL; OCLC, 2007) e a ISO 16363:2012 aprofundaram as exigências organizacionais, tecnológicas e de governança baseadas diretamente no OAIS, mas não expandiram de forma equivalente os requisitos específicos para museus e bibliotecas, mesmo que esses ambientes lidem com

objetos digitais que podem possuir valor de arquivo¹. Nesse sentido, Zierau e Schultz (2013) argumentam que a aplicação do OAIS em ecossistemas distribuídos, realidade comum em instituições museais, onde coexistem sistemas de gestão, plataformas de difusão, evidencia o desafio de adaptar o modelo a contextos cuja lógica não é arquivística, mas híbrida. Logo, a ausência de diretrizes mais amplas, capazes de contemplar os diferentes tipos de acervos e práticas curatoriais, contribuiu para que museus e bibliotecas não estivessem em local de igualdade nas discussões sobre preservação digital de longo prazo.

1.1 Contextualização e cenário da preservação digital

Nas últimas décadas, o avanço acelerado da transformação digital impôs às instituições informacionais, arquivos, bibliotecas e museus uma profunda reconfiguração de suas práticas de gestão, acesso e preservação da informação. Como observa Conway (2010), a digitalização dos processos institucionais e a crescente produção de objetos digitais transformaram o ambiente informacional em um ecossistema altamente dependente de tecnologias, no qual a vulnerabilidade dos suportes e a obsolescência acelerada constituem riscos estruturais para a memória social. Gilliland (2014) reforça que essa transformação não é apenas tecnológica: ela afeta modelos de governança, cadeias de custódia² e as próprias concepções de autenticidade e confiabilidade da informação preservada. Nesse cenário, Corrado e Moulaison Sandy (2017) destacam que a preservação digital deixou de ser um problema técnico isolado

¹ O conceito de *valor de arquivo* remete ao princípio de que determinados documentos, após cumprirem suas funções administrativas, legais ou fiscais, podem adquirir valores secundários - notadamente valor probatório, informativo ou histórico - que justificam sua preservação permanente. Bellotto explica que esses valores são identificados por meio da avaliação documental, processo que determina o destino dos documentos e fundamenta as práticas de seleção, recolhimento e preservação nos arquivos permanentes (Bellotto, 2002).

² O termo “cadeia de custódia” refere-se ao conjunto de procedimentos formais, registros e controles que documentam a trajetória de um objeto digital desde sua criação ou ingresso na instituição até seu uso presente, assegurando sua autenticidade, integridade e confiabilidade. No contexto da preservação digital, envolve documentação contínua, *logs* de eventos, controle de acesso e monitoramento sistemático, conforme discutido por Corrado e Moulaison Sandy ao tratarem da gestão e governança necessárias para garantir a confiabilidade dos objetos preservados (Corrado; Moulaison Sandy, 2017).

para se tornar um componente estratégico da gestão institucional, demandando políticas robustas, infraestrutura especializada e competências multidisciplinares.

No campo arquivístico, a necessidade de assegurar a autenticidade e o valor probatório dos documentos é central. Bellotto (2002) destaca que os documentos de arquivo são produzidos para registrar ações e servir como prova jurídica, mantendo ao longo de seu ciclo de vida valores que ultrapassam o mero registro informacional, influenciando decisões administrativas, legais e históricas. Essa natureza comprobatória implica a adoção de requisitos rigorosos de preservação que assegurem a integridade e autenticidade dos registros, sobretudo considerando a permanência de seu valor probatório mesmo quando passam a integrar acervos históricos e museais (Flores, 2015).

A preservação digital, nesse contexto, envolve um conjunto de procedimentos estruturais e operacionais que abrangem desde o estabelecimento de infraestrutura tecnológica, financiamento e padrões normativos até a implementação contínua de ações como migração, emulação e encapsulamento (Interpares, 2007). A combinação desses elementos exige políticas institucionais robustas, equipes capacitadas e modelos de gestão capazes de sustentar a continuidade da preservação ao longo do tempo.

A preservação digital no Brasil não pode ser compreendida sem referência ao arcabouço jurídico-normativo que, nas últimas décadas, passou a disciplinar tanto a gestão de documentos digitais quanto os requisitos mínimos para digitalização e manutenção de acervos de longo prazo. A Lei nº 8.159/1991, marco fundador da Política nacional de arquivos, estabeleceu princípios estruturantes da gestão documental e da proteção especial aos documentos arquivísticos públicos, incluindo aqueles produzidos em meio digital. Ao reconhecer a preservação como dimensão indissociável da gestão documental, a legislação brasileira inaugurou uma diretriz que, posteriormente, seria aprofundada por orientações técnicas do Conselho Nacional de Arquivos (Conarq). Nesse sentido, a Resolução Conarq nº 43/2015 tornou-se um marco ao instituir diretrizes para a preservação de documentos arquivísticos digitais, enfatizando requisitos de autenticidade, integridade, confiabilidade e acessibilidade, princípios coerentes com o Modelo OAIS e com as demandas

contemporâneas de governança da informação. A resolução introduziu a noção de repositórios digitais confiáveis no contexto nacional, apontando a necessidade de políticas institucionais, mecanismos de controle e infraestrutura tecnológica compatível com práticas internacionais de preservação.

Outro instrumento central é o Decreto nº 10.278/2020, que estabeleceu os padrões técnicos mínimos para a digitalização de documentos no Brasil, definindo parâmetros de resolução, integridade, autenticidade, metadados e rastreabilidade. O decreto representou um avanço concreto ao vincular juridicamente a qualidade técnica da digitalização à validade legal dos documentos digitalizados, configurando um marco regulatório que orienta tanto órgãos públicos quanto iniciativas particulares. A esse conjunto soma-se a resolução Conarq nº 51/2023, que atualiza princípios e requisitos para a implementação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) no Brasil. Inspirada diretamente na ISO 14721 (OAIS) e na ISO 16363 (TRAC), a resolução define parâmetros mínimos para ingestão, armazenamento, governança de metadados, segurança da informação, gestão de riscos e planejamento de preservação, reforçando a necessidade de arquiteturas institucionais que articulem sistemas de difusão, como o Tainacan, a sistemas especializados de preservação, como o Archivematica.

Esse conjunto normativo evidencia que, no Brasil, digitalização e preservação digital deixaram de ser meras atividades de gestão administrativas para se tornarem obrigações legais, estruturantes da governança documental e patrimonial. Quando analisado sob a perspectiva museológica, esse cenário torna-se ainda mais desafiador, pois a maioria das legislações foi originalmente formulada a partir de uma lógica arquivística, orientada a documentos textuais e registros administrativos. Como observam Alencar e Martins (2025), a predominância desse enfoque técnico-normativo contribui para a marginalização de objetos digitais museais — audiovisuais, modelagens 3D, mídias interativas, *datasets*³ — que não se ajustam facilmente aos requisitos estabelecidos pela legislação e pelos modelos consolidados de preservação. Tal assimetria

³ Um *dataset* pode ser entendido como uma coleção de dados específica para um determinado projeto ou análise, funcionando como uma amostra ou um subconjunto de um *database* (banco de dados) mais amplo (Date, 2004).

evidencia uma lacuna entre as normativas brasileiras e a realidade dos acervos museológicos digitais, cujo tratamento exige protocolos diferenciados de captura, descrição, migração e armazenamento, para os quais as legislações atuais nem sempre oferecem diretrizes técnicas suficientes.

Flores (2015) observa que replicar os requisitos de preservação digital de documentos convencionais para objetos não textuais é particularmente desafiador, sobretudo diante da diversidade de formatos, da ausência de padrões nos sistemas de difusão, das restrições de acessibilidade e interoperabilidade e das questões relacionadas a direitos autorais e propriedade intelectual.

A ampliação da diversidade e da complexidade dos acervos digitais tem intensificado os desafios enfrentados por museus e instituições com acervos informacionais, sobretudo porque a maior parte das estratégias consolidadas de preservação digital foi construída para documentos textuais. Alencar e Martins (2025) abordam que há uma predominância de abordagens voltadas a formatos como PDF/A e registros administrativos, deixando marginalizados objetos digitais não textuais, como audiovisuais, mídias interativas, modelagens tridimensionais e dados científicos, que exigem estratégias específicas de manutenção, migração e acesso.

Ademais, os autores evidenciam, por meio de revisão sistemática, que essa assimetria representa uma lacuna na literatura, uma vez que formatos museológicos demandam soluções diferenciadas para garantir autenticidade, integridade e preservação, extrapolando a capacidade das práticas tradicionalmente orientadas ao universo arquivístico. A partir dessa análise, os autores demonstram que a maturidade das soluções tecnológicas disponíveis não acompanha a complexidade dos objetos museais digitais contemporâneos, revelando um descompasso estrutural entre as necessidades reais das instituições patrimoniais e as diretrizes técnico-metodológicas atualmente predominantes.

O cenário torna-se mais crítico quando se examina a vulnerabilidade estrutural dos museus brasileiros frente aos riscos de perda informacional. Embora o documento canadense *Digital Preservation Recommendations for*

Small Museums, de 2024, reconheça que instituições de pequeno e médio porte enfrentam limitações significativas, como infraestrutura reduzida, equipes enxutas e ausência de políticas formais de preservação, tais desafios manifestam-se de forma mais acentuada no Brasil, onde a instabilidade tecnológica, recursos financeiros escassos, a ausência de políticas públicas consolidadas e a inexistência de estruturas de preservação de longo prazo, comprometem a sustentabilidade dos acervos digitais (Santos; Loureiro, 2018). Essas fragilidades revelam não apenas a dificuldade histórica das instituições museais brasileiras em implementar estratégias sistemáticas de continuidade digital, mas também a dependência de soluções emergenciais e desarticuladas, que não respondem adequadamente aos riscos associados à obsolescência tecnológica e à degradação dos suportes ao longo do tempo.

A precariedade institucional e a ausência de políticas duradouras de salvaguarda configuram, entretanto, um fenômeno estrutural de longa duração. Rubim (2007) assinala que o Brasil convive com um problema crônico de desvalorização da cultura e das instituições responsáveis por sua preservação. Esse quadro tornou-se evidente no incêndio de grandes proporções ocorrido no Museu Nacional, no Rio de Janeiro, em 2 de setembro de 2018, que destruiu parcela importante do acervo histórico e científico acumulado ao longo de dois séculos, estimado em cerca de vinte milhões de itens catalogados (Brasil, 2018; Santos; Loureiro, 2018).

Esse episódio trouxe à luz, de forma contundente, que parte dos objetos não possuía documentação robusta ou representação digital, resultando não apenas na perda física irreversível das coleções, mas também na supressão de referências essenciais à memória nacional. A inexistência de acervos digitais ou de políticas sistemáticas de digitalização comprometeu, portanto, não somente a integridade material dos bens museológicos, mas a permanência de seu valor histórico e científico.

Para Nogueira e Amorim (2019), a imagem digital constitui um recurso estratégico para a continuidade da memória, a medida que viabiliza a reconstrução simbólica e a difusão ampliada dos acervos, desempenhando uma função mediadora entre a materialidade perdida e a permanência da informação inscrita nesses objetos. No caso do Museu Nacional, a ausência de processos

consolidados de digitalização reforçou a vulnerabilidade material das coleções e a insuficiência de políticas públicas voltadas à proteção do patrimônio científico e histórico do país.

Nesse cenário de crescente complexidade informacional e de necessidade de digitalização dos objetos museais, segundo o estudo de Martins *et al.* (2017), o Tainacan desponta como uma das principais plataformas brasileiras de difusão de acervos digitais em instituições museológicas. Desenvolvido como *software* livre e orientado por padrões internacionais de descrição e interoperabilidade, o Tainacan oferece uma infraestrutura flexível para organização, representação e disponibilização de coleções digitais, permitindo que instituições de diferentes portes incorporem práticas de gestão mais estruturadas, com significativa redução de barreiras técnicas e financeiras. Sua adoção crescente por museus, centros culturais, universidades e órgãos públicos evidência não apenas sua capacidade de atender às demandas de acessibilidade e democratização da informação, mas também sua potencialidade como solução nacional alinhada às diretrizes de transparência, difusão científica e valorização do patrimônio cultural (Martins *et al.* 2017).

Entretanto, embora o Tainacan represente um avanço significativo no campo da gestão e da mediação digital de acervos, sua arquitetura é fundamentalmente voltada para a organização, a descrição e o acesso, e não para a preservação de longo prazo. A plataforma dispõe de funcionalidades robustas para interoperabilidade, uso de vocabulários controlados, visualização avançada e integração com outras ferramentas digitais, mas carece de requisitos que assegurem a autenticidade, integridade, rastreabilidade e permanência dos objetos digitais em conformidade com modelos consolidados de preservação, como o OAIS (International Organization for Standardization, 2012a). Assim, à medida que instituições brasileiras passam a utilizar o Tainacan como solução central para difusão de acervos digitais, surgem desafios relacionados à continuidade, à sustentabilidade tecnológica e à necessidade de integração com sistemas de preservação capazes de assegurar a longevidade dos objetos que circulam nesse ambiente.

Diante desse contexto, torna-se necessária a revisão dos modelos tradicionais de gestão da informação e do patrimônio museal, incorporando

princípios de interdisciplinaridade entre Arquivologia, Biblioteconomia e Museologia. Dodebei (2011) argumenta que a integração entre essas áreas é fundamental para a constituição de instituições capazes de articular, simultaneamente, objetos, documentos e práticas curatoriais, superando barreiras epistemológicas⁴ que, historicamente, compartimentaram seus domínios de atuação. Essa convergência visa assegurar não apenas a preservação física ou digital dos bens museais, mas também a continuidade de seus valores simbólicos, sociais e históricos.

Assim, compreender a convergência entre práticas de preservação digital, modelos internacionais de referência e normativas brasileiras não é apenas um exercício teórico, mas uma necessidade concreta para o desenvolvimento de soluções aplicáveis, eficazes e juridicamente robustas. É nesse ponto que esta tese se insere: propor uma arquitetura de repositório confiável capaz de preencher a lacuna entre sistemas de difusão e requisitos de preservação, alinhando o Tainacan às exigências tecnológicas, científicas e normativas que moldam o campo da preservação digital contemporânea.

1.2 Problema e justificativa da pesquisa

A lacuna que fundamenta esta pesquisa emerge de um cenário em que a difusão de acervos digitais, por meio de plataformas amplamente adotadas, avança de maneira dissociada da implementação de estratégias formais de preservação de longo prazo. Embora tais sistemas respondam adequadamente às demandas de organização, descrição e acesso, a literatura demonstra que ferramentas orientadas à gestão e à mediação não garantem, por si só, a autenticidade, a integridade e a continuidade informacional necessárias para assegurar a permanência dos objetos digitais (Flores, 2015). Flores (2015) destaca que, sem mecanismos de preservação, cadeias de custódia confiáveis ou processos estruturados de gestão de risco, acervos digitais tornam-se

⁴ A epistemologia investiga os fundamentos, limites e validade do conhecimento científico (Tesser, 1994).

particularmente vulneráveis à obsolescência tecnológica, à degradação lógica e a falhas institucionais. Tal vulnerabilidade é agravada no contexto brasileiro, marcado por fragilidades estruturais, ausência de políticas integradas de preservação digital e descontinuidades históricas na salvaguarda do patrimônio cultural (Rubim, 2007; Santos; Loureiro, 2018).

Harvey e Weatherburn (2019) fornecem diretrizes práticas sobre como lidar com as mudanças tecnológicas e identificam lacunas que resultam na falta de padronização na preservação digital. Uma dessas lacunas é a inconsistência nos esquemas de metadados usados pelas organizações, o que leva a desafios institucionais na recuperação de dados e documentos. Além disso, ferramentas de preservação diversificadas, com diferentes tecnologias e abordagens, criam barreiras para a interoperabilidade entre sistemas. A falta de políticas e estratégias claras para a preservação digital, com cada organização desenvolvendo suas próprias práticas, resulta em abordagens descoordenadas. Para além das ações internas das instituições, há uma colaboração limitada entre as organizações, procedendo na duplicação de esforços e na falta de compartilhamento de soluções inovadoras. Essas lacunas destacam a necessidade de padrões e diretrizes que possam ser amplamente adotados para assegurar uma preservação digital mais eficaz e sustentável (Harvey; Weatherburn, 2019).

Nesse contexto de preservação, integridade e difusão das informações, o Brasil possui iniciativas para digitalizar o seu valioso acervo documental e histórico. O Tainacan, solução brasileira, foi criado como parte de um contexto de transformação digital e da necessidade de democratizar o acesso aos acervos culturais no Brasil. Desenvolvido em parceria entre o Ministério da Cultura, a Universidade Federal de Goiás (UFG) e o Instituto Brasileiro de Museus (Ibram), o projeto teve início como uma resposta à crescente demanda por soluções que permitissem a digitalização e a gestão eficiente de acervos, especialmente em instituições com recursos limitados (Tainacan, 2024). O Tainacan tem como objetivo oferecer uma plataforma que facilita a organização e a publicação de coleções *online*, utilizando tecnologias de *software* livre que podem ser adotadas por qualquer instituição cultural. A ferramenta promove a acessibilidade ao patrimônio cultural, tornando-o disponível ao público de forma inclusiva e conectando museus, bibliotecas e arquivos a um ecossistema digital mais amplo.

Com essa solução, os acervos são catalogados, organizados e compartilhados utilizando padrões modernos de metadados e princípios de interoperabilidade. A integração com o *WordPress* permite que instituições de diferentes portes tenham acesso a uma ferramenta flexível e eficiente para publicar seus acervos (Martins, 2017). No entanto, ainda que atue como um repositório de difusão, o Tainacan não é estruturado para assegurar a preservação digital a longo prazo, pois não incorpora nativamente mecanismos como auditoria de integridade, controle de versões, metadados de preservação ou estratégias baseadas em modelos como o OAIS.

A ausência de uma arquitetura de preservação integrada a essa ferramenta, amplamente adotada por instituições museais brasileiras, expõe esse patrimônio digital ao risco de perda. Diante disso, manifesta-se o seguinte problema de pesquisa: como a adoção de padrões baseados no modelo OAIS pode contribuir para a preservação digital a longo prazo dos acervos museais difundidos por meio do Tainacan?

1.3 Relevância do estudo

A relevância deste estudo fundamenta-se em uma dupla contribuição. Do ponto de vista científico, a pesquisa avança no campo da Ciência da Informação ao verticalizar a discussão teórica sobre modelos de preservação digital, como o OAIS, aplicados a uma ferramenta (Tainacan), que constitui uma política pública no Brasil e é de ampla utilização, investigando os desafios de sua aplicação em um *software* que não foi originalmente concebido para fins de preservação. A investigação das divergências entre os princípios do modelo OAIS e as funcionalidades atualmente oferecidas pelo Tainacan permite evidenciar os aspectos críticos que impedem a preservação de seus acervos a longo prazo, revelando lacunas e oportunidades de melhoria na estrutura e no funcionamento do Tainacan.

No âmbito institucional e social, a tese busca oferecer uma contribuição prática ao propor um modelo de aplicação que possa orientar desenvolvedores,

gestores de acervos e instituições museais, notadamente as vinculadas ao Ibram, na implementação de estratégias de preservação mais robustas. A aplicação dos princípios de preservação digital do modelo OAIS, reconhecido internacionalmente, pode assegurar que esses materiais sejam arquivados com segurança e que o seu acesso seja mantido a longo prazo. Além disso, ao alinhar o Tainacan com o padrão OAIS, busca-se garantir a interoperabilidade e a aceitação da plataforma entre instituições museais e arquivísticas. O modelo OAIS é amplamente aceito como referência na área de preservação digital, e aderir a esse padrão reforça a credibilidade e a confiança nas práticas adotadas pela ferramenta, tornando-a uma solução mais robusta e confiável.

A análise da viabilidade de aplicar os princípios OAIS ao Tainacan oferece oportunidades para identificar requisitos técnicos e práticas recomendadas que podem aprimorar a arquitetura técnica do sistema. Com isso, espera-se uma melhoria nos processos de gestão de metadados, armazenamento seguro, autenticação e acesso. Tais requisitos fornecem as diretrizes necessárias para aprimorar a confiabilidade da plataforma. Garantir a sustentabilidade financeira e organizacional do Tainacan é outra vantagem evidente desta pesquisa. Sugerir soluções que garantam a manutenção e a acessibilidade dos recursos digitais a longo prazo é crucial para as instituições museais que necessitam de uma solução que permaneça relevante e eficaz ao longo do tempo.

Ao avaliar os desafios técnicos, conceituais e organizacionais enfrentados pelas instituições que utilizam o Tainacan, será possível compreender os entraves que dificultam a integração de padrões de preservação. Por fim, o estudo visa propor soluções práticas e sustentáveis que possam ser incorporadas à ferramenta ou às práticas institucionais associadas a ela, com vistas a ampliar sua confiabilidade e efetividade na gestão do patrimônio museal digital. Ao corrigir essas lacunas e implementar melhorias, o Tainacan poderá ser fortalecido como uma solução pública robusta, beneficiando inúmeras instituições que confiam nele para difundir e preservar seus acervos.

Esse desafio é particularmente relevante no contexto do Ibram, autarquia federal responsável pela Política Nacional de Museus, que orienta e fomenta políticas para mais de 4.000 (quatro mil) museus em todo o território nacional, além de gerir diretamente cerca de 30 museus federais. Muitas dessas instituições utilizam a plataforma Tainacan para a organização e difusão online

de seus acervos. Diante da ausência de políticas de preservação integradas a essa ferramenta, a presente pesquisa se propõe a investigar como a adoção de padrões baseados no modelo OAIS pode contribuir para garantir a preservação a longo prazo dos acervos museais difundidos por meio do Tainacan.

1.4 Objetivo geral da tese

O objetivo geral desta tese é sugerir um modelo de aplicação de requisitos do OAIS/TRAC-ISO 16363 para orientar a implementação de um repositório digital confiável ao Tainacan.

1.5 Objetivos específicos

- a. Compreender o estado da arte sobre a situação da preservação digital no campo dos museus.
- b. Investigar uma arquitetura conceitual de maneira a orientar a constituição de um repositório digital confiável no Tainacan.
- c. Examinar a estrutura sociotécnica do Tainacan, identificando seus limites e capacidades para a incorporação dos requisitos OAIS.
- d. Propor modelo de aplicação ao Tainacan para interoperá-lo a um RDC-ARQ.

1.6 Pergunta de pesquisa

Diante da obsolescência tecnológica, da fragilidade de formatos digitais e da ausência de políticas institucionais voltadas à preservação digital de objetos com valor híbrido (arquivo-museu), torna-se essencial refletir sobre

modelos e padrões que orientem práticas sustentáveis nesse campo. Nesse contexto, a presente pesquisa investigou “quais componentes do Modelo OAIS permitem a consideração do Tainacan como uma ferramenta de difusão capaz de interligar-se a um repositório digital confiável?”

2. REFERENCIAL TEÓRICO: PRESERVAÇÃO DIGITAL E DIFUSÃO

A crescente produção de documentos em formato digital impõe desafios significativos às instituições no que tange à sua preservação e ao acesso a longo prazo. A fragilidade inerente aos suportes digitais, a rápida obsolescência tecnológica e a facilidade de adulteração ameaçam a autenticidade e a disponibilidade de informações que servem como evidência de atividades, garantia de direitos e fonte para pesquisa (Conselho Nacional de Arquivos, 2023). Nesse contexto, a constituição de RDC-Arq emerge como uma necessidade para assegurar a custódia e o acesso perene ao patrimônio documental digital ao longo do tempo.

Paralelamente, surgem plataformas dedicadas à gestão e à difusão de acervos, provenientes das necessidades de instituições que custodiam acervos informacionais que demandam transparência ativa ⁵ e divulgação à sociedade, entre as quais se destaca o Tainacan. Desenvolvido no Brasil, o Tainacan é uma solução de política pública brasileira, de código aberto, que se integra ao WordPress para criar ambientes dinâmicos e flexíveis para publicação de coleções digitais (Tainacan, 2025). Sua natureza pública e facilidade de implementação o tornam uma opção atrativa para instituições que desejam dar visibilidade a seus acervos com pouco recurso aplicado. Contudo, sua arquitetura é primariamente voltada para a difusão, não contemplando nativamente todas as complexas funcionalidades exigidas para um repositório de preservação de longo prazo.

⁵ A transparência ativa é o princípio que determina a divulgação de dados e informações de interesse público por iniciativa do próprio setor público, independentemente de solicitações. Este dever está previsto na Lei de Acesso à Informação (Brasil, 2011).

Tendo isso em vista, este capítulo tem como objetivo elaborar uma arquitetura conceitual que oriente a constituição de um RDC-Arq, integrado à plataforma Tainacan. Para tal, propõe-se um modelo que articula o Tainacan como uma camada de acesso e difusão com um *software* e procedimentos que atendam aos requisitos de preservação a longo prazo definidos por normativas nacionais, como as do Conselho Nacional de Arquivos (Conarq), e internacionais, como o modelo de referência OAIS (ISO 14721:2012) e a TRAC.

2.1 Preservação digital: conceitos fundamentais e desafios

A preservação digital é definida pelo Conselho Nacional de Arquivos como o “conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo o acesso e a interpretação dos documentos digitais pelo tempo que for necessário” (Conselho Nacional de Arquivos, 2023, p. 9). Nesse escopo, o conceito de repositório digital confiável é central.

A transição para o ambiente digital trouxe consigo um paradoxo central para a gestão da informação, como observam Santos e Flores (2015), o ritmo acelerado da evolução tecnológica criou uma situação contraditória: ao mesmo tempo em que há grande facilidade para criar documentos digitais, os métodos de preservação ainda são escassos e de difícil aplicação. A informação registrada em bits⁶, embora aparentemente imaterial, está intrinsecamente ligada a um aparato tecnológico de *hardware* e *software* em constante mutação, tornando os documentos digitais “condicionados à tecnologia” (Santos; Flores, 2015, p. 201). Como alertam os autores, essa dependência gera vulnerabilidades significativas dos suportes, que ameaçam a capacidade das instituições de garantir o acesso contínuo e a fidedignidade⁷ de seus acervos.

⁶ Contração de *binary digit* (dígito binário). É a menor e mais fundamental unidade de informação em sistemas digitais, podendo assumir apenas um de dois valores, geralmente representados como 0 ou 1 (Tanenbaum; Wetherall, 2011).

⁷ Qualidade de um documento ser aquilo que diz ser e de estar livre de adulteração ou corrupção. Um documento fidedigno é confiável em sua representação dos fatos e atos (Duranti, 1995).

A preservação digital, nesse contexto, não se resume a uma simples ação de salvaguarda, mas a um “conjunto de ações gerenciais e técnicas exigidas para superar as mudanças tecnológicas e a fragilidade dos suportes, garantindo o acesso e a interpretação dos documentos digitais pelo tempo que for necessário” (Conselho Nacional de Arquivos, 2023, p. 9). Trata-se de um processo contínuo e proativo que visa manter a autenticidade, a integridade, a confiabilidade e a acessibilidade dos documentos arquivísticos digitais ao longo do tempo.

Para alcançar o objetivo de manter os requisitos de autenticidade e preservação a longo prazo, aderiu-se a outra área a um conceito que se tornou central nesse tema e que foi reformulado para o ambiente digital: o de cadeia de custódia. Ele teve sua origem no campo do Direito, particularmente na ciência forense, referindo-se à “documentação cronológica ou histórica que registra a sequência de custódia, controle, transferência, análise e descarte de evidências físicas ou eletrônicas de provas periciais” (Lopes; Bareta, 2014, p. 51). Seu propósito é garantir a integridade da prova, assegurando que a evidência apresentada em um processo judicial é a mesma que foi coletada, livre de adulterações ou contaminações. No Brasil, o Código de Processo Penal (Brasil, 2019) formaliza esse entendimento, definindo-a como o conjunto de procedimentos para manter e documentar a história cronológica do vestígio.

Na Arquivologia tradicional, a ideia de cadeia de custódia sempre esteve presente, ainda que implicitamente, ligada ao princípio da proveniência. A manutenção de uma cadeia de custódia ininterrupta, que documenta a origem e a posse dos documentos, é fundamental para a presunção de sua autenticidade (Hedstrom, 2000). A autenticidade de um documento analógico derivava, em grande parte, de sua guarda contínua e do controle físico por uma entidade arquivística confiável. Contudo, como aponta Flores (2015), essa abordagem, baseada na posse física e na inseparabilidade entre informação e suporte, tornou-se insuficiente para o ambiente digital.

Em seu artigo⁸ “Cadeia de custódia para documentos arquivísticos digitais”, Santos e Flores (2016) argumentam que, no ambiente digital, a cadeia

⁸ Este estudo resgata a abordagem da cadeia de custódia para documentos arquivísticos analógicos e propõe sua reformulação com o objetivo de contemplar documentos arquivísticos armazenados em ambientes digitais. Para isso, discute-se a importância da manutenção de uma

de custódia não pode mais se basear na posse física, mas, sim, em uma trilha de auditoria digital ininterrupta e confiável. Eles propõem uma Cadeia de Custódia Digital Arquivística (CCDA), que consiste na documentação de todas as ações e transferências que um documento sofre desde sua criação em um Sistema Informatizado de Gestão Arquivística de Documentos (SIGAD) até sua preservação em um RDC-Arq. A CCDA, portanto, é o mecanismo que permite comprovar a autenticidade de um documento digital. Ela não é um sistema único, mas um conjunto de procedimentos e registros distribuídos ao longo de diferentes sistemas, que, juntos, contam a história do documento. Essa cadeia é mantida por meio de tecnologias como assinaturas digitais, carimbos de tempo⁹, protocolos de transferência seguros e, fundamentalmente, a geração de metadados de preservação, como os do padrão *Preservation Metadata: Implementation Strategies* (PREMIS)¹⁰, que registra cada evento na vida do objeto digital (Santos; Flores, 2016).

A manutenção de uma CCDA robusta depende da implementação de uma série de elementos técnicos e procedimentais ao longo do ciclo de vida do documento. Com base nos trabalhos de Santos e Flores (2016), nas diretrizes do Conarq (2023) e no modelo OAIS, pode-se sistematizar os principais elementos, conforme o quadro abaixo.

Quadro 1 - Elementos da cadeia de custódia digital arquivística.

Elemento	Descrição	Mecanismo de Implementação
Identificação Única e Persistente	Atribuição de um identificador que não se altera, mesmo que a localização do objeto mude.	Identificadores como DOI, ARK ou Handle.

linha de custódia ininterrupta na implementação de sistemas informatizados para gestão, preservação e acesso aos documentos arquivísticos (Flores, 2015, p. 117).

⁹ Mecanismo criptográfico que associa um documento ou dado a um instante específico no tempo, provido por uma terceira parte confiável (Autoridade de Carimbo de Tempo - ACT). Um carimbo de tempo serve como uma evidência digital de que aquele dado já existia em uma determinada data e hora e não foi alterado desde então (Stallings; Brown, 2017).

¹⁰ Sigla para *Preservation Metadata: Implementation Strategies*. É o padrão internacional de fato para metadados de preservação, definindo um dicionário de dados com as informações que um repositório precisa para suportar a preservação de longo prazo de objetos digitais (Premis Editorial Committee, 2015).

Elemento	Descrição	Mecanismo de Implementação
Registro de Eventos (Auditoria)	Gravação de todas as ações significativas realizadas sobre o objeto digital.	Metadados de preservação (PREMIS), logs de sistema.
Verificação de Integridade	Mecanismos para garantir que o objeto digital não foi corrompido ou alterado indevidamente.	Algoritmos de <i>hash</i> (<i>checksums</i>) como SHA-256, verificados em cada transferência.
Controle de Versões	Gerenciamento das diferentes versões de um documento, preservando o histórico de alterações.	Sistemas de controle de versão, metadados que relacionam as diferentes versões.
Segurança e Controle de Acesso	Garantia de que apenas pessoal autorizado possa realizar ações sobre os documentos.	Autenticação de usuários, perfis de acesso, assinaturas digitais.
Documentação da Transferência	Registro formal da passagem de custódia entre diferentes sistemas ou responsáveis.	Geração de manifestos de transferência, pacotes de informação (SIP, AIP, DIP) com metadados.
Contexto de Produção	Preservação dos metadados que descrevem o contexto de criação, o produtor e a função do documento.	Metadados descritivos e de gestão, encapsulados junto ao objeto digital.

Fonte: Elaborado pela autora, com base em Santos e Flores (2016) e Conarq (2023).

O fluxo proposto por Santos e Flores (2016) para a manutenção da cadeia de custódia articula a passagem do documento do ambiente de gestão (SIGAD) para o ambiente de preservação (RDC-Arq), alinhando-se diretamente com o modelo de referência OAIS, normatizado pela ISO 14721¹¹. Assim, a transferência não é uma simples cópia de arquivos, mas um evento formal, mediado por um Pacote de Informação de Submissão (SIP), que encapsula o documento e seus metadados iniciais. Ao ser admitido no RDC-Arq, o SIP é validado e transformado em um Pacote de Informação de Arquivamento (AIP), a estrutura central para a preservação de longo prazo (Santos; Flores, 2016; International Organization for Standardization, 2012a).

É na gestão do AIP que a CCDA e a preservação de longo prazo se unem. Cada ação de preservação – como uma verificação de integridade, migração de formato para combater a obsolescência tecnológica ou mudança de suporte de

¹¹ ISO 14721:2012.

armazenamento – é registrada como um evento nos metadados de preservação do AIP, notadamente utilizando o padrão PREMIS. Esses metadados funcionam como um diário de bordo do objeto digital, documentando toda a sua trajetória e garantindo que a cadeia de custódia seja estendida e mantida indefinidamente. A implementação desses requisitos assegura que, mesmo após décadas e múltiplas transformações tecnológicas, seja possível rastrear e comprovar a autenticidade e a integridade do documento desde sua origem (International Organization for Standardization, 2012a)

Portanto, a CCDA, quando implementada por meio da arquitetura do OAIS, fornece o alicerce teórico e prático para a construção de sistemas de preservação confiáveis. Ela desloca o foco da posse física para a documentação processual, transformando a autenticidade de uma presunção para uma qualidade demonstrável e auditável, essencial para a chancela dos arquivos no meio digital (Santos; Flores, 2016; International Organization for Standardization, 2012a).

Este capítulo estabeleceu os fundamentos teóricos para a construção de um repositório digital confiável, partindo do paradoxo fundamental da preservação digital: a facilidade de criação *versus* a complexidade da perenidade. Demonstrou-se que a garantia de autenticidade, integridade e acesso a longo prazo transcende a simples salvaguarda de arquivos, exigindo uma abordagem estruturada e metodologicamente sólida. Essa discussão se conecta ao problema de pesquisa através da preocupação em garantir a longevidade dos acervos difundidos no Tainacan, pois são cerca de mil instalações ativas, com dezenas de milhares de documentos e informações culturais, museais e históricas de valor inestimável (Tainacan, 2025).

A análise revelou que a CCDA, conforme descrita por Santos e Flores (2016), constitui a resposta teórica a esse desafio. O conceito desloca o paradigma da custódia física para uma trilha de auditoria processual, onde a autenticidade não é presumida, mas continuamente demonstrada por meio de metadados e registros de eventos. Essa abordagem transforma a confiança em um atributo verificável, essencial para a credibilidade dos acervos no meio digital. A aplicação do modelo OAIS à realidade de acervos difundidos no Tainacan, vai possibilitar que instituições que o utilizam como única plataforma de

armazenamento e difusão, consigam implementar requisitos que assegurem a longevidade das informações. Acervos históricos estão contidos no universo do Tainacan, sem o arcabouço de segurança, passíveis de significativa perda (Tainacan, 2025).

O objetivo deste estudo é diagnosticar os requisitos essenciais para a integração da plataforma Tainacan a um *software* de preservação digital. Tal integração visa a estabelecer um fluxo processual preconizado pelo modelo de referência OAIS, cuja sequência, preservar para depois difundir, é destacada por Santos e Flores (2016). A proposta busca viabilizar a preservação de acervos, inaugurando, por conseguinte, uma metodologia que permita a futuras coleções gerenciar o ciclo completo de preservação e difusão de forma integrada e contínua no âmbito da própria plataforma Tainacan, como ilustra a figura abaixo:

Figura 1 - Modelo de cadeia de custódia de preservação contínua (baseado no OAIS)



Fonte: Elaborado pela autora, com base no modelo OAIS.

2.2 O estado da arte da preservação digital em museus: uma revisão sistemática da literatura

O objetivo desta revisão de literatura é analisar como a adoção de padrões baseados no modelo OAIS (Open Archival Information System) pode

contribuir para a preservação digital a longo prazo dos **acervos museais** disponibilizados por meio do Tainacan. Esta revisão deu-se com o objetivo de fornecer uma fundamentação teórica detalhada que demonstre a relevância do modelo OAIS no contexto da preservação digital de acervos culturais e de formatos não convencionais de arquivos¹². Esta etapa busca esclarecer a relação entre os princípios do OAIS e as práticas de preservação adotadas por instituições que utilizam ferramentas tecnológicas para difundir os seus acervos, destacando os benefícios potenciais e os desafios encontrados.

O escopo desta revisão abrange estudos realizados com foco em iniciativas de preservação digital em museus e instituições detentoras de acervos museais ao redor do mundo. Foram analisados artigos acadêmicos, relatórios técnicos e estudos de caso que abordam a implementação de padrões de metadados, estratégias de preservação a longo prazo e a interoperabilidade entre sistemas, baseados no modelo OAIS.

Para garantir uma análise completa do universo da pesquisa, adotou-se um protocolo de revisão sistemática, seguindo diretrizes metodológicas PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (MOHER et al., 2009). A revisão sistemática permite sintetizar as evidências científicas existentes, identificando padrões, desafios e estratégias focados no tema da pesquisa, considerando a aplicabilidade do modelo OAIS (ISO 14721) como referência metodológica central.

A problematização central da revisão consistiu em diagnosticar o que a literatura trata sobre a aplicação do modelo OAIS na preservação digital de objetos culturais.

Para responder a essa questão, foram definidos critérios de inclusão e exclusão. Foram selecionados estudos que tratam da aplicação do OAIS na

¹² Formatos não convencionais são aqueles que exigem estratégias especializadas de preservação digital devido à sua dependência de software, hardware e metadados específicos. Entre os principais formatos estão arquivos audiovisuais, mídias interativas, objetos 3D, dados científicos e documentos estruturados.:CORRADO, Edward M.; MOULAISON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017.;Becker, Christoph et al. *Digital Preservation: Theory and Practice*. Berlin: Springer, 2009.

preservação de objetos culturais ou que não fossem tradicionalmente documentos de arquivo¹³ considerando abordagens aplicáveis a mídias interativas, audiovisuais, modelagem tridimensional e dados científicos (Arellano, 2012; Gilliland, 2014). Excluíram-se trabalhos que se restringem à preservação de documentos textuais convencionais, abordando formatos clássicos.

Os termos de busca foram estruturados para abranger diferentes abordagens da preservação digital conforme a pergunta de pesquisa: ("digital preservation" OR "preservação digital") AND ("cultural heritage" OR "objetos culturais") AND ("OAIS").

As bases de dados utilizadas para a busca de literatura incluem Scopus, Web of Science, Google Scholar e a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), conforme a recomendação da literatura de Becker (2009). Além dessas fontes, foi utilizada a inteligência artificial SCISPaCE, uma ferramenta especializada em pesquisa acadêmica assistida por IA, que permite analisar grandes volumes de publicações, identificar padrões temáticos e sugerir referências relevantes com base em modelagem semântica. A inclusão da SCISPaCE ampliou a abrangência da revisão, garantindo um levantamento mais eficiente e sistemático dos estudos (XU; LIN, 2021).

A Scopus e a Web of Science são bases de dados multidisciplinares e indexadas, amplamente utilizadas para revisões sistemáticas devido à sua rigorosa curadoria de publicações científicas. Ambas oferecem filtros avançados para refinamento da pesquisa, permitindo a seleção de artigos de alto impacto e relevância na área pesquisada (Becker, 2009). O Google Scholar foi incluído

¹³ Um documento de arquivo é aquele produzido ou recebido por uma instituição ou indivíduo no decorrer de suas atividades, preservado como evidência de informações, ações ou decisões. Diferente de documentos comuns, os documentos arquivísticos possuem valor probatório e administrativo, sendo geridos segundo princípios arquivísticos para garantir sua autenticidade, fidedignidade e acessibilidade ao longo do tempo. Esses documentos podem existir em diversos suportes, como papel, microfilme e formatos digitais. BRASIL. Conselho Nacional de Arquivos (CONARQ). *Glossário de Terminologia Arquivística*. Brasília: Arquivo Nacional, 2005.; SCHELLENBERG, Theodore R. *Arquivos modernos: princípios e técnicas*. Rio de Janeiro: FGV, 2004.

para complementar a busca, pois indexa uma variedade mais ampla de documentos, incluindo *preprints*, teses, dissertações e publicações institucionais que podem não estar disponíveis em bases de dados comerciais (HADDOW; GENONI, 2010; MARTÍNEZ-LÓPEZ et al., 2018). Essa base permite um panorama mais abrangente da produção científica, incluindo estudos emergentes e de acesso aberto. A BDTD (Biblioteca Digital Brasileira de Teses e Dissertações) foi selecionada para identificar pesquisas acadêmicas desenvolvidas no Brasil, garantindo a inclusão de estudos nacionais. A plataforma reúne teses e dissertações de diversas universidades brasileiras, muitas das quais trazem contribuições relevantes para o contexto (SANTOS et al., 2017). Por fim, a SCISPaCE foi utilizada para otimizar a busca por estudos relevantes. A ferramenta emprega inteligência artificial para mapear tendências, sugerir conexões entre artigos e identificar pesquisas correlatas. Essa abordagem permitiu reduzir vieses de seleção, detectar artigos altamente citados e explorar publicações menos acessíveis por métodos convencionais de busca (XU; LIN, 2021). Os termos de busca foram estruturados considerando a garantia de que a recuperação dos trabalhos seriam focados no tema.

Para a seleção dos artigos, adotou-se um processo de triagem em múltiplas etapas. Inicialmente, os trabalhos recuperados foram analisados por meio da leitura de títulos e resumos, verificando se abordavam a aplicação do modelo OAIS na preservação digital de objetos culturais não convencionais ou museológicos. Aqueles que não atenderam a esse critério foram excluídos. Os estudos que passaram por essa primeira fase foram submetidos à leitura integral, na qual foram avaliados quanto à relevância metodológica, aplicabilidade prática e alinhamento com os desafios específicos da preservação digital em contextos museológicos e patrimoniais. Os artigos incluídos foram analisados detalhadamente, extraindo-se informações sobre as estratégias de preservação utilizadas, os desafios identificados e as soluções propostas pelos autores. Os trabalhos, que incluíram artigos, livros e normas técnicas, foram analisados, permitindo uma síntese crítica das abordagens existentes e a identificação de lacunas na literatura, que serviram de base para a formulação de recomendações e diretrizes futuras para a preservação digital de acervos museológicos.

A busca foi realizada em cinco bases de dados e plataformas: Scopus (45 resultados), Web of Science (38), Google Scholar (109), Biblioteca Digital Brasileira de Teses e Dissertações – BDTD (5) e SciSpace (50), totalizando 247 publicações inicialmente recuperadas (Ver Tabela 1). Essa amostragem ampla se justifica pela diversidade de escopos e cobertura de cada base. A Scopus e a Web of Science forneceram artigos indexados e com rigor acadêmico, enquanto o Google Scholar e a SciSpace complementaram com literatura cinzenta e produções institucionais ou emergentes. A BDTD, por sua vez, permitiu incorporar a produção científica brasileira no recorte da pesquisa (Santos et al., 2017; Xu; Lin, 2021).

Tabela 1 - Busca nas bases de dados sobre preservação digital museal.

Base de dados	Número de Resultados
Scopus	45
Web of Science	38
Google Scholar	109
BDTD	5
SciSpace	50

Fonte: Elaborado pela autora, realizada no dia 24/02/2025.

A investigação nas bases de dados resultou em um volume significativo de publicações relacionadas ao tema “preservação digital”, sendo a maioria dos estudos voltada para a preservação de documentos digitais convencionais, como **arquivos textuais**. Embora a pesquisa tenha sido estruturada para recuperar trabalhos que abordassem a preservação digital de objetos culturais sob a ótica do modelo OAIS, verificou-se que a maioria dos resultados se concentra em estratégias de preservação documental e repositórios institucionais, sem um foco específico em objetos museológicos complexos, como modelos tridimensionais, mídias interativas e acervos audiovisuais (Apendice I). Essa lacuna evidencia a necessidade de um aprofundamento na literatura para compreender as abordagens metodológicas aplicadas a esse contexto, bem

como o desenvolvimento de pesquisas voltadas para a adequação do OAIS à preservação digital de acervos museológicos e patrimoniais.

Apesar do resultado não expressivo, do ponto de vista quantitativo, três trabalhos possuem temas semelhantes ao desta pesquisa. São eles: "Digital Preservation Recommendations for Small Museums" (CANADA, 2024); "Digital Preservation for Libraries, Archives, and Museums" (CORRADO; MOULAISON SANDY, 2015) ;e "Creating a Framework for Applying OAIS to Distributed Digital Preservation" (ZIERAU; SCHULTZ, 2013). A seguir serão descritos os principais aspectos de cada um deles.

O manual *Digital Preservation Recommendations for Small Museums*, desenvolvido pelo Canadian Heritage Information Network (CHIN), apresenta diretrizes essenciais para auxiliar pequenos e médios museus na implementação de práticas de preservação digital. Reconhecendo as limitações orçamentárias e técnicas dessas instituições, o documento propõe abordagens simplificadas, porém eficazes, para garantir a integridade e acessibilidade de acervos digitais ao longo do tempo. As recomendações se concentram na criação de um inventário digital organizado, na adoção de estratégias de armazenamento adaptáveis e na implementação de rotinas de manutenção e *backup*.

O primeiro passo destacado pelo manual é a utilização do Digital Preservation Toolkit do CHIN, um conjunto de ferramentas que auxilia na catalogação dos recursos digitais e no desenvolvimento de políticas de preservação. A elaboração de um inventário de ativos digitais é uma etapa crucial para garantir que todos os conteúdos relevantes sejam identificados e protegidos contra riscos de perda ou corrupção. Além disso, o documento incentiva a definição de políticas institucionais de preservação, estabelecendo diretrizes claras sobre as responsabilidades dos profissionais envolvidos e os procedimentos de manutenção dos arquivos digitais. Para fortalecer a governança da preservação digital, recomenda-se a criação de planos e procedimentos estruturados, garantindo que cada etapa do processo seja documentada e executada de forma sistemática.

Uma das soluções práticas sugeridas pelo manual é a simplificação do processo de inventário, aproveitando ferramentas automatizadas ou métodos

áveis para reduzir a carga de trabalho das equipes responsáveis pela preservação. A organização centralizada dos ativos digitais também é incentivada, pois facilita a localização e o gerenciamento dos arquivos armazenados. A adoção de estratégias alternativas ao modelo OAIS (Open Archival Information System) é proposta para instituições que não dispõem de infraestrutura avançada, permitindo a implementação gradual de medidas de preservação sem a necessidade de investimentos imediatos em sistemas complexos.

Dentre as práticas recomendadas, destaca-se a criação de metadados de preservação, garantindo que as informações essenciais sobre a origem e a estrutura dos arquivos digitais sejam registradas. Para evitar a obsolescência dos conteúdos armazenados, o manual recomenda a geração anual de cópias de preservação, acompanhada de processos regulares de manutenção e atualização de formatos. Além disso, a realização de backups frequentes é fundamental para mitigar riscos de falhas técnicas e perda de dados. A manutenção adequada dos dispositivos de armazenamento, aliada a uma política de nomeação consistente dos arquivos, contribui para a acessibilidade e recuperação eficiente das informações preservadas.

No entanto, o manual reconhece desafios inerentes às abordagens sugeridas, como a violação da regra 3-2-1 de backup, que estipula a existência de três cópias dos dados, armazenadas em dois tipos de mídia diferentes, com pelo menos uma cópia mantida off-site. A recomendação do CHIN, que prevê apenas duas cópias em locais distintos, pode não ser suficiente para garantir a redundância ideal dos arquivos. Além disso, a falta de um processo formal de arquivamento de cópias autoritativas pode comprometer a autenticidade e a confiabilidade dos dados preservados.

Dessa forma, o *Digital Preservation Recommendations for Small Museums* apresenta uma abordagem acessível e escalável para a preservação digital em pequenos museus, proporcionando diretrizes práticas para garantir a continuidade dos acervos digitais com recursos limitados. Embora as recomendações ofereçam soluções viáveis para instituições com infraestrutura restrita, a adoção de estratégias mais robustas, como a implementação do

modelo OAIS à medida que a capacidade institucional aumenta, é essencial para garantir a preservação digital a longo prazo. Assim, a combinação de planejamento estratégico, práticas de backup eficientes e uso de tecnologias apropriadas pode permitir que pequenos museus superem as barreiras da preservação digital, assegurando o acesso contínuo ao seu patrimônio cultural.

O manual reconhece que a implementação completa do modelo OAIS (Open Archival Information System) pode ser desafiadora para pequenos museus devido a limitações de recursos. Portanto, o manual sugere que essas instituições considerem a adoção de práticas de preservação digital mais simples e escaláveis inicialmente, evoluindo para a implementação completa do modelo OAIS conforme desenvolvem infraestrutura e recursos adequados

Para auxiliar nesse processo de transição, as seguintes etapas gerais podem ser consideradas:

1. Avaliação Inicial e Planejamento:
 - Inventário de Ativos Digitais: Catalogar todos os recursos digitais existentes para entender o escopo e as necessidades de preservação.
 - Análise de Recursos: Avaliar a infraestrutura atual, incluindo hardware, software e competências da equipe, para identificar lacunas em relação aos requisitos do OAIS.
 - Desenvolvimento de Políticas de Preservação: Estabelecer diretrizes e objetivos claros para a preservação digital, alinhados com as capacidades e metas da instituição.
2. Implementação Gradual de Componentes do OAIS:
 - Ingestão (Ingest): Desenvolver procedimentos para a submissão e aceitação de materiais digitais no repositório, garantindo a captura adequada de metadados.
 - Armazenamento Arquivístico (Archival Storage): Implementar soluções de armazenamento que assegurem a integridade e a segurança dos dados a longo prazo.
 - Gestão de Dados (Data Management): Estabelecer sistemas para organizar, manter e fornecer acesso aos metadados associados aos objetos digitais.
 - Planejamento de Preservação (Preservation Planning): Monitorar a evolução tecnológica e desenvolver estratégias para mitigar a obsolescência de formatos e suportes.

- Acesso (Access): Criar mecanismos que permitam aos usuários localizar e acessar os materiais preservados de forma eficiente.
3. Capacitação e Melhoria Contínua:
- Treinamento da Equipe: Investir na formação dos colaboradores para garantir a compreensão e a competência na aplicação dos princípios do OAIS.
 - Revisão e Atualização de Processos: Avaliar regularmente as práticas de preservação e realizar ajustes conforme necessário para acompanhar as mudanças tecnológicas e institucionais.

CANADA. Canadian Heritage Information Network. *Digital Preservation Recommendations for Small Museums*. Ottawa: Government of Canada, 2024. Disponível em: <https://www.canada.ca/en/heritage-information-network/services/digital-preservation/recommendations-small-museums.html>. Acesso em: 19 fev. 2025.

Embora o manual do CHIN não forneça um guia detalhado para a implementação do OAIS, ele enfatiza a importância de práticas de preservação digital adaptadas às realidades dos pequenos museus, com a possibilidade de evoluir para modelos mais robustos como o OAIS à medida que os recursos e as necessidades institucionais cresçam.

O livro *Digital Preservation for Libraries, Archives, and Museums*, de Edward M. Corrado e Heather Moulaison Sandy (2015), apresenta uma abordagem abrangente e prática sobre a preservação digital no contexto de bibliotecas, arquivos e museus (LAMs). A obra discute os desafios técnicos, administrativos e conceituais relacionados à preservação de objetos digitais, oferecendo um guia detalhado para profissionais que desejam implementar estratégias eficazes de preservação. O livro enfatiza a importância da preservação digital como uma questão de gestão, não apenas uma preocupação técnica.

A metodologia utilizada pelos autores é baseada no conceito da *Digital Preservation Triad*, que estrutura a preservação digital em três pilares fundamentais: gestão, tecnologia e conteúdo. A gestão envolve a definição de políticas, alocação de recursos e planejamento estratégico, garantindo que a preservação seja sustentável a longo prazo. A tecnologia diz respeito à

infraestrutura utilizada, incluindo repositórios digitais, formatos de arquivos e estratégias de emulação e migração. O pilar do conteúdo refere-se à curadoria dos materiais digitais, abrangendo critérios de seleção, metadados e acessibilidade. Para embasar essas diretrizes, os autores realizam uma revisão de literatura abrangente, analisam estudos de caso e exploram modelos teóricos, como o Open Archival Information System (OAIS), padrão internacional para repositórios confiáveis.

Os autores enfatizam que a preservação digital vai além da simples criação de backups ou da disponibilidade temporária de arquivos. Para garantir a acessibilidade de longo prazo, é necessário adotar estratégias que previnam a degradação dos arquivos digitais e a obsolescência de formatos. Entre os principais riscos estão a deterioração dos suportes físicos, como CDs e discos rígidos, e a descontinuação de softwares necessários para leitura de determinados formatos. A fim de mitigar esses problemas, recomenda-se o uso de formatos padronizados e amplamente aceitos, como PDF/A para documentos, TIFF para imagens e METS para estruturação de objetos digitais. Além disso, são fundamentais técnicas como migração de formatos – que consiste na conversão para padrões mais sustentáveis – e emulação, que permite simular o funcionamento de softwares antigos em novas plataformas.

Outro aspecto essencial abordado na obra é a gestão da preservação digital, que deve ser tratada como uma questão organizacional e não apenas técnica. Para que uma instituição possa garantir a integridade e acessibilidade de seu acervo digital, é necessário um planejamento estruturado, que inclua políticas institucionais bem definidas, mecanismos de financiamento sustentáveis e colaborações entre instituições. A falta de financiamento contínuo é apontada como um dos principais desafios para a preservação digital, pois sem investimentos recorrentes, os esforços de digitalização e preservação podem se tornar inviáveis. Dessa forma, os autores sugerem que instituições busquem financiamento por meio de modelos colaborativos, consórcios e parcerias interinstitucionais, reduzindo custos e aumentando a eficiência das ações de preservação.

A escolha da infraestrutura e dos softwares de preservação digital é um fator crítico para garantir a longevidade e a acessibilidade dos objetos. O livro apresenta uma análise detalhada das principais soluções disponíveis no

mercado, comparando softwares proprietários e de código aberto em termos de funcionalidades, custo, aplicabilidade e desafios operacionais. A análise busca fornecer subsídios para que instituições de diferentes portes possam tomar decisões informadas sobre qual solução se adapta melhor às suas necessidades e restrições financeiras.

Dentre as soluções proprietárias, destacou-se o ‘Preservica’ e o ‘Rosetta’ como dois dos sistemas mais robustos e amplamente utilizados. O Preservica é um software de preservação digital que opera como um sistema completo de gerenciamento de repositórios, permitindo desde a ingestão de arquivos até sua preservação a longo prazo, com suporte a auditorias, migração de formatos e controle automatizado da integridade dos objetos digitais. Sua maior vantagem, segundo os autores, reside na automação de processos de preservação, reduzindo a necessidade de intervenções manuais e minimizando o risco de corrupção de arquivos ao longo do tempo. No entanto, seu alto custo e a dependência da empresa fornecedora para atualizações e suporte técnico podem ser barreiras para pequenas instituições.

Por sua vez, o Rosetta, desenvolvido pela Ex Libris, é um sistema igualmente robusto, adotado principalmente por bibliotecas nacionais, arquivos públicos e grandes universidades. Ele incorpora princípios do modelo OAIS, oferecendo funcionalidades como gerenciamento de políticas de preservação, compatibilidade com metadados de preservação (como PREMIS e METS) e integração com sistemas de bibliotecas digitais. Uma das principais vantagens do Rosetta é sua capacidade de gerenciar grandes volumes de dados e sua integração com outros produtos da Ex Libris, como o Alma e o Primo. No entanto, assim como o Preservica, o Rosetta apresenta altos custos operacionais e exige contratos de manutenção e suporte contínuo, o que pode torná-lo inviável para instituições com orçamentos mais limitados.

Em contraposição às soluções comerciais, os autores exploram alternativas *open source*, que oferecem maior flexibilidade e independência tecnológica. Dentre elas, destacam-se o Archivematica e o DSpace, dois dos sistemas mais amplamente adotados por instituições de ensino e pesquisa. O Archivematica, desenvolvido pela Artefactual Systems, é um software de

preservação digital baseado no modelo OAIS, focado na normalização e na preservação de longo prazo de objetos digitais. Sua principal vantagem é a transparência no código-fonte, permitindo que instituições personalizem suas funcionalidades conforme suas necessidades. Além disso, o Archivematica oferece suporte a múltiplos formatos de arquivo e inclui ferramentas para validação de metadados e migração de formatos, garantindo que os conteúdos digitais permaneçam acessíveis em novas plataformas tecnológicas. Entretanto, sua complexidade técnica e a necessidade de infraestrutura local para implementação podem representar desafios para instituições que não possuem equipes especializadas em TI.

O DSpace, por outro lado, é uma solução mais voltada para o gerenciamento de repositórios institucionais, sendo amplamente utilizado por universidades para armazenamento e disseminação de documentos científicos, dissertações e teses. Embora o DSpace ofereça algumas funcionalidades de preservação digital, sua principal limitação é a falta de ferramentas nativas para auditoria e migração de formatos, tornando-se mais adequado para acesso e gestão de conteúdos digitais do que para preservação em longo prazo. Contudo, sua grande comunidade de usuários e desenvolvedores permite constantes aprimoramentos e integração com outras soluções de preservação digital.

Além da comparação entre softwares proprietários e open source, os autores analisam o impacto da computação em nuvem na preservação digital, um tema cada vez mais relevante para instituições que buscam reduzir custos operacionais. A adoção de serviços de nuvem oferece vantagens significativas, como a eliminação da necessidade de manutenção de servidores físicos locais e a escalabilidade do armazenamento conforme a demanda. Soluções como AWS Glacier (Amazon Web Services), Google Cloud Storage e Microsoft Azure possibilitam o armazenamento de grandes volumes de dados com custos relativamente baixos e alta redundância, garantindo a recuperação dos arquivos em caso de falha de hardware.

Os autores alertam que a preservação digital baseada em nuvem também apresenta desafios consideráveis. Um dos principais problemas é a dependência de fornecedores externos, o que pode comprometer a governança dos dados

preservados. Caso um provedor de nuvem altere suas políticas ou encerre seus serviços, as instituições podem enfrentar dificuldades na migração dos dados para outras plataformas. Além disso, há preocupações com a segurança e privacidade dos dados armazenados na nuvem, especialmente em arquivos governamentais e documentos sensíveis. Dessa forma, os autores recomendam que as instituições adotem uma estratégia híbrida, combinando armazenamento em nuvem com repositórios locais, a fim de garantir maior controle sobre os dados e minimizar riscos associados à terceirização da infraestrutura de preservação digital.

Os resultados apresentados no livro indicam que não existe uma solução única para a preservação digital, e que a escolha entre sistemas proprietários, *open source* e armazenamento em nuvem deve levar em conta fatores institucionais, orçamentários e técnicos. Enquanto sistemas comerciais oferecem mais suporte e funcionalidades prontas, soluções *open source* garantem maior flexibilidade e independência. Por sua vez, a computação em nuvem pode ser uma opção viável para redução de custos, mas exige planejamento cuidadoso para evitar riscos de perda de controle sobre os dados preservados. Assim, os autores reforçam que a decisão sobre qual infraestrutura adotar deve estar alinhada às necessidades institucionais e ao planejamento estratégico da preservação digital, garantindo que os conteúdos permaneçam acessíveis e autênticos a longo prazo.

Além da discussão central sobre softwares e armazenamento, os autores trazem ao debate os acervos que não são tradicionalmente manuseados em arquivos, tais como os de formatos complexos¹⁴. Para os autores, a preservação de objetos culturais em formatos complexos representa um dos maiores desafios na preservação digital, pois envolve materiais cuja estrutura, funcionalidade e autenticidade dependem de fatores técnicos específicos. Corrado e Moulaison

¹⁴ 'Formatos complexos' são aqueles que exigem estratégias especializadas de preservação digital devido à sua dependência de software, hardware e metadados específicos. Entre os principais formatos estão arquivos audiovisuais, mídias interativas, objetos 3D, dados científicos e documentos estruturados.:CORRADO, Edward M.; MOULAISSON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017.;BECKER, Christoph et al. *Digital Preservation: Theory and Practice*. Berlin: Springer, 2009.

Sandy (2017) destacam a importância de abordar a preservação de conteúdos como audiovisuais, mídias interativas e dados científicos por meio de estratégias especializadas, dada a sua alta vulnerabilidade à obsolescência tecnológica. A preservação desses objetos culturais exige uma abordagem integrada, que combine metadados detalhados, repositórios digitais especializados e políticas institucionais bem definidas para garantir a acessibilidade e autenticidade a longo prazo.

Os conteúdos audiovisuais são um dos formatos mais críticos no campo da preservação digital, uma vez que dependem de codecs e contêineres¹⁵ específicos para reprodução. Vídeos e gravações sonoras podem ser armazenados em formatos proprietários, que frequentemente se tornam obsoletos com a evolução da tecnologia. A obra enfatiza que a adoção de formatos padronizados e amplamente aceitos, como Broadcast WAVE (BWF) para áudio e Matroska (MKV) ou Motion JPEG2000 para vídeo, é uma estratégia essencial para mitigar os riscos de obsolescência. Além disso, a preservação audiovisual deve considerar a captura e armazenamento de metadados técnicos detalhados, garantindo que informações como taxa de bits¹⁶, taxa de amostragem e codificação¹⁷ sejam preservadas para permitir a futura reprodução fiel do conteúdo original. Como resultado das análises dos autores, conclui-se que instituições que não adotam padrões de preservação adequados enfrentam sérias dificuldades na migração e recuperação de conteúdos audiovisuais, podendo perder materiais de valor histórico e cultural irreparáveis.

¹⁵ 'Codecs' são algoritmos responsáveis pela compressão e descompressão de arquivos multimídia, enquanto contêineres organizam e armazenam diferentes fluxos de dados dentro de um único arquivo: CORRADO, Edward M.; MOULAISON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017.; PEREIRA, Cristiano Amaral. *Preservação Digital de Conteúdos Audiovisuais: Desafios e Estratégias*. Revista Brasileira de Preservação Digital, v. 5, n. 1, p. 45-60, 2021.

¹⁶ A 'taxa de bits' refere-se à quantidade de dados por segundo em um arquivo digital, enquanto a taxa de amostragem determina a frequência com que um sinal analógico é convertido em digital. Já a codificação define o processo de compressão e armazenamento dos dados. CORRADO, Edward M.; MOULAISON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017. ; CASEY, Mike. *Why Media Preservation Can't Wait: The Gathering Storm*. IASA Journal, v. 44, p. 14-22, 2015.

¹⁷

As mídias interativas, que incluem jogos digitais, websites dinâmicos e aplicativos interativos, representam um desafio ainda maior, pois sua funcionalidade está atrelada a fatores como engines de software,¹⁸ compatibilidade de hardware e sistemas operacionais específicos. Os autores ressaltam que a preservação desse tipo de objeto cultural exige estratégias avançadas, como a emulação e a virtualização, que permitem recriar o ambiente original no qual o objeto foi desenvolvido. Ferramentas como Emularity, DOSBox e ScummVM são citadas como exemplos de soluções eficazes para preservar a experiência original de softwares e jogos interativos. No entanto, a pesquisa aponta que a emulação não é uma solução definitiva, pois depende da manutenção contínua dos softwares emuladores e do desenvolvimento de documentação detalhada sobre o funcionamento dos sistemas originais. Dessa forma, os autores recomendam que, além da emulação, seja realizada a extração e documentação de elementos fundamentais do software, incluindo código-fonte (quando disponível), metadados de interatividade e materiais de apoio, garantindo que futuras gerações possam compreender e reproduzir o funcionamento dessas mídias.

Outro formato de grande complexidade abordado é o dos dados científicos, que incluem conjuntos massivos de dados (*big data*), resultados de pesquisas experimentais e modelos computacionais. A preservação desses materiais apresenta desafios tanto em relação ao volume de armazenamento quanto à interoperabilidade entre diferentes plataformas. Os autores sugerem que instituições científicas adotem repositórios digitais específicos para preservação de dados, como o Dataverse, Zenodo e Dryad, que oferecem funcionalidades voltadas para a curadoria e disseminação de dados de pesquisa. Além disso, destacam a importância da adoção de metadados descritivos e técnicos rigorosos, utilizando padrões como Dublin Core, DataCite e METS, garantindo a reutilização e interpretação dos dados em longo prazo. O estudo

¹⁸ Engines de software são plataformas utilizadas para o desenvolvimento de aplicações interativas, fornecendo suporte para renderização gráfica, física, inteligência artificial e interação do usuário: CORRADO, Edward M.; MOULAISON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017.; LOWOOD, Henry. *Preserving Virtual Worlds: Challenges and Approaches*. American Journal of Play, v. 3, n. 2, p. 185-198, 2011.

conclui que instituições que não adotam padrões de preservação de dados científicos enfrentam dificuldades significativas na reprodutibilidade de pesquisas e na validação de informações, comprometendo a integridade da ciência aberta.

Diante dos desafios apresentados, os autores enfatizam a necessidade de políticas institucionais específicas para cada tipo de formato complexo, assegurando sua preservação e acessibilidade futura. Eles recomendam que instituições culturais e científicas elaborem diretrizes claras sobre captura, armazenamento, documentação e disseminação de conteúdos digitais. Essas diretrizes devem incluir protocolos para migração periódica de formatos, auditorias regulares de integridade dos arquivos e estratégias de replicação geográfica, garantindo que objetos culturais digitais permaneçam acessíveis mesmo diante de falhas técnicas ou desastres inesperados.

Como resultado da análise, concluiu-se que a preservação de formatos complexos exige um esforço coordenado entre diferentes áreas do conhecimento, combinando tecnologia, curadoria digital e governança institucional. A pesquisa demonstra que instituições que adotam práticas padronizadas e investem em tecnologias inovadoras têm maior sucesso na preservação e reutilização de objetos culturais digitais. No entanto, também se observa que muitas organizações ainda não possuem uma abordagem estruturada para a preservação digital desses materiais, o que pode levar à perda irreversível de conteúdos valiosos. Dessa forma, os autores reforçam que a preservação digital de formatos complexos deve ser tratada como uma prioridade estratégica, demandando investimentos em infraestrutura, capacitação profissional e colaboração interinstitucional para garantir a sustentabilidade e acessibilidade dos objetos culturais no futuro.

Com relação aos acervos museológicos, existem questões legais não definidas que desafiam a preservação digital de objetos. A complexidade desse tema reside no fato de que muitos objetos digitais preservados em bibliotecas, arquivos e museus são compostos por múltiplos elementos protegidos por diferentes regimes de propriedade intelectual. Isso inclui textos, imagens, vídeos, áudio, software, metadados e até interações dinâmicas que podem ter autores

distintos e licenças variadas. Assim, ao longo do livro, os autores analisam como essas camadas de complexidade impactam diretamente as estratégias de preservação, acesso e reutilização desses acervos.

Um dos principais desafios mencionados na obra é a dificuldade em obter permissões legais para preservar e disseminar objetos digitais protegidos por direitos autorais. No caso de documentos textuais e imagens estáticas, a legislação pode ser relativamente clara, permitindo o uso sob a doutrina do uso justo (fair use) ou exceções para bibliotecas e arquivos. No entanto, ao lidar com mídias interativas, videogames, simulações científicas ou realidade virtual, o problema se agrava, pois, esses materiais frequentemente envolvem múltiplos criadores e tecnologias proprietárias que limitam a preservação e o acesso futuro. Os autores destacam que, em muitos casos, as instituições que buscam preservar esses conteúdos precisam negociar diretamente com os detentores de direitos, o que pode ser um processo demorado e oneroso.

Outro aspecto abordado é a relação entre propriedade intelectual e software utilizado na criação e preservação de objetos digitais. Ferramentas como engines de software (*Unity*, *Unreal Engine*, *CryEngine*), codecs proprietários e plataformas de armazenamento em nuvem podem restringir a preservação a longo prazo. Se uma engine for descontinuada ou se os termos de uso do software mudarem, a instituição pode perder o direito de acessar ou reproduzir um objeto digital, mesmo que tenha uma cópia arquivada. Os autores sugerem que uma das formas de mitigar esse risco é priorizar o uso de formatos abertos e software de código aberto, garantindo que a preservação digital não fique dependente de empresas privadas que podem mudar suas políticas no futuro.

Além disso, os autores discutem o impacto da legislação sobre repositórios digitais e acesso aberto. Embora muitas instituições defendam a disponibilização pública de acervos digitais para fins de pesquisa e preservação cultural, restrições legais podem impedir a publicação de determinados materiais, especialmente no caso de obras protegidas por contratos comerciais. No contexto internacional, os autores citam a importância de iniciativas como a Europeia e o Internet Archive, que trabalham para disponibilizar conteúdos

digitais sob licenças abertas. No entanto, alertam que, mesmo nesses casos, diferenças na legislação de direitos autorais entre países podem dificultar a harmonização das políticas de acesso.

Outro problema identificado é a perda de contexto do objeto digitalizado. Enquanto um livro ou um documento digitalizado mantém sua estrutura informacional relativamente intacta, um artefato museológico pode perder significados se não for acompanhado de metadados detalhados sobre sua origem, uso, materiais e contexto histórico. Os autores destacam a necessidade de utilizar metadados específicos para museologia, como CIDOC CRM¹⁹, garantindo a documentação precisa dos objetos preservados digitalmente.

A preservação digital de objetos museológicos frequentemente lida com formatos complexos, como arquivos 3D, realidade aumentada (AR), realidade virtual (VR) e mídias interativas, que apresentam alta vulnerabilidade à obsolescência tecnológica. Os autores alertam que, sem estratégias de migração de formatos e suporte a padrões abertos, há o risco de que objetos digitalizados se tornem inacessíveis devido à descontinuação de softwares e hardwares necessários para sua visualização. Os formatos proprietários são outro ponto crítico. Muitas ferramentas utilizadas para digitalização e modelagem tridimensional salvam arquivos em formatos fechados, tornando-se uma barreira para a preservação a longo prazo. O estudo recomenda o uso de formatos abertos e amplamente adotados na comunidade de preservação digital, como OBJ, STL e GLTF para modelagem 3D, reduzindo o risco de dependência tecnológica.

Abordam-se a infraestrutura necessária para armazenar e manter repositórios digitais de objetos museológicos. Diferente de bibliotecas e arquivos, que podem contar com repositórios institucionais padronizados, os museus

¹⁹ * O CIDOC CRM (Conceptual Reference Model) é um modelo conceitual padronizado para a organização e intercâmbio de informações sobre o patrimônio cultural, garantindo a interoperabilidade entre diferentes sistemas de informação de museus, arquivos e bibliotecas.; DOERR, Martin. *The CIDOC Conceptual Reference Model: An Ontological Approach to Semantic Interoperability of Metadata*. *AI Magazine*, v. 24, n. 3, p. 75-92, 2003; CIDOC CRM. *Definition of the CIDOC Conceptual Reference Model (Version 7.1.1)*. Disponível em: <https://www.cidoc-crm.org/>. Acesso em: 19 fev. 2025.

frequentemente precisam de infraestruturas especializadas, capazes de lidar com grandes volumes de dados, metadados complexos e repositórios interativos. A preservação digital em museus envolve custos elevados, tanto para aquisição de equipamentos e software quanto para a manutenção de sistemas. Os autores enfatizam que, muitas vezes, os museus não possuem financiamento contínuo para preservar digitalmente seus acervos, o que leva à obsolescência de arquivos e à perda de informações. Como solução, sugerem parcerias interinstitucionais, financiamento colaborativo e o uso de plataformas open source, reduzindo os custos e garantindo maior acessibilidade aos conteúdos.

Como resultado de suas análises, os autores concluem que a preservação de acervos complexos exige uma abordagem multidisciplinar, na qual juristas, arquivistas, bibliotecários e especialistas em tecnologia devem trabalhar juntos para definir políticas institucionais que conciliam preservação digital e conformidade legal. Entre as recomendações para lidar com a complexidade dos direitos autorais, destacam-se:

- Mapear previamente os detentores de direitos e buscar licenças que permitam a preservação a longo prazo.
- Priorizar formatos e softwares abertos sempre que possível, reduzindo a dependência de plataformas proprietárias.
- Utilizar licenças flexíveis, como Creative Commons²⁰, para conteúdos produzidos por instituições públicas e acadêmicas.
- Desenvolver políticas institucionais de acesso restrito para materiais protegidos, garantindo que possam ser preservados mesmo que sua disponibilização pública seja limitada por questões legais.

Ressaltam ainda que a falta de clareza na legislação de direitos autorais pode comprometer seriamente a preservação digital. Se não houver políticas

²⁰ Creative Commons é um sistema de licenciamento aberto que permite aos criadores conceder permissões flexíveis sobre o uso de suas obras, equilibrando direitos autorais e acesso público ao conhecimento. CREATIVE COMMONS. *About the Licenses*. Disponível em: <https://creativecommons.org/licenses/>. Acesso em: 19 fev. 2025.; CORRADO, Edward M.; MOULAISON SANDY, Heather. *Digital Preservation for Libraries, Archives, and Museums*. 2nd ed. Lanham: Rowman & Littlefield, 2017.

bem definidas e parcerias estratégicas entre instituições culturais e detentores de direitos, muitos objetos digitais podem se perder por falta de permissão para serem preservados ou disseminados. Dessa forma, a obra enfatiza que a preservação de acervos complexos não é apenas uma questão técnica, mas também uma questão legal e ética, exigindo soluções que equilibrem a proteção dos direitos dos criadores com o acesso ao patrimônio digital.

Conclui-se que a importância do modelo OAIS (ISO 14721) como estrutura fundamental para garantir a integridade e autenticidade de repositórios digitais confiáveis reside em sua capacidade de prover diretrizes normativas e funcionais que asseguram a preservação de longo prazo, a interoperabilidade entre sistemas e a sustentabilidade das iniciativas de preservação digital em contextos institucionais diversos, incluindo bibliotecas, arquivos e museus (ISO, 2012; ZIERAU; SCHULTZ, 2013). No entanto, os autores alertam que a aplicação do OAIS deve ser adaptada conforme o contexto institucional, especialmente quando se trata da preservação distribuída, onde múltiplas instituições compartilham responsabilidades na preservação de conteúdos digitais. Além disso, reforçam a necessidade de metadados específicos, para preservação e estruturação de objetos digitais e garantindo que objetos culturais sejam descritos de forma contextualizada e interoperável. Um dos aspectos mais relevantes do livro é a reflexão sobre o papel da preservação digital não apenas como um processo de arquivamento, mas como uma ferramenta de democratização do conhecimento. Os autores destacam como tecnologias emergentes, como realidade virtual e exposições digitais, podem transformar a experiência do público com o patrimônio cultural, tornando coleções antes restritas a espaços físicos acessíveis globalmente. No entanto, alertam que o sucesso dessas iniciativas depende de políticas institucionais bem estruturadas, garantindo que a preservação digital não seja tratada como um projeto isolado, mas como um compromisso institucional de longo prazo.

Dessa forma, o *Digital Preservation for Libraries, Archives, and Museums* conclui que a preservação digital é uma responsabilidade coletiva, que exige colaboração entre profissionais de diferentes áreas, investimento contínuo em infraestrutura, atualização de padrões e diretrizes institucionais claras. Ao integrar os desafios técnicos, jurídicos e institucionais em uma abordagem

sistemática, a obra se torna uma referência essencial para profissionais que atuam na preservação digital, fornecendo ferramentas conceituais e práticas para garantir que os acervos digitais permaneçam acessíveis e autênticos para as futuras gerações.

O artigo "Creating a Framework for Applying OAIS to Distributed Digital Preservation" propõe a criação de um arcabouço metodológico para aplicar o Modelo de Referência para um Sistema Aberto de Informação Arquivística (OAIS) à Preservação Digital Distribuída (DDP). O estudo aborda a necessidade de um framework conceitual e terminológico para sistematizar e padronizar abordagens de preservação digital que envolvem múltiplas infraestruturas e organizações o que vai de encontro ao objetivo desta tese, pois algumas instituições avaliadas possuem acervos museológicos e culturais.

A pesquisa adota um conjunto de estratégias metodológicas para fundamentar o desenvolvimento de um framework que possibilite a aplicação do Modelo OAIS em contextos de Preservação Digital Distribuída (DDP). Essas estratégias visam garantir uma abordagem sistemática e científica para a criação de um modelo que seja funcional, adaptável e validado pela comunidade acadêmica e profissional da área de preservação digital. A revisão de literatura teve como objetivo mapear e sintetizar os estudos existentes sobre o modelo OAIS e sua aplicação à preservação digital distribuída, como identificar as lacunas conceituais e metodológicas entre a teoria e os desafios encontrados em sistemas distribuídos de preservação. A revisão revelou que, embora o OAIS seja amplamente adotado para sistemas de preservação centralizados, há poucas diretrizes sobre como aplicar suas funções em contextos distribuídos. Identificou-se que as lacunas mais significativas estão relacionadas à governança, responsabilidade compartilhada, e à interoperabilidade entre diferentes repositórios. Além disso, observou-se uma falta de consenso sobre a terminologia e a aplicação de metadados em redes de preservação distribuída, o que dificulta a colaboração e a integração de sistemas. Essas lacunas motivaram a necessidade de expandir o OAIS para acomodar as especificidades dos sistemas distribuídos (Zierau & Schultz, 2013).

A etapa que envolve a análise do Modelo OAIS e sua aplicabilidade a sistemas distribuídos, explorou sua estrutura conceitual avaliando o desempenho de suas funções em ambientes distribuídos de preservação digital.

O OAIS define um conjunto de entidades funcionais, como Ingestão (*Ingest*), Armazenamento Arquivístico (*Archival Storage*) e Acesso (*Access*), mas, segundo o estudo, não oferece uma abordagem clara sobre como essas funções podem ser distribuídas entre várias instituições que compartilham responsabilidades. O principal apontamento feito indicou que o modelo precisa ser expandido para incluir novos conceitos de governança colaborativa e responsabilidade compartilhada. A função de acesso, por exemplo, precisa considerar novas formas de garantir que os objetos digitais possam ser acessados de forma segura e consistente por várias entidades em uma rede distribuída. Além disso, os sistemas distribuídos demandam soluções específicas para interoperabilidade e certificação entre diferentes repositórios, aspectos que não são totalmente abordados pelo OAIS.

Para complementar a análise teórica, a pesquisa incorporou entrevistas com especialistas e estudos de caso de instituições que implementaram ou estão implementando estratégias de preservação digital distribuída. O objetivo das entrevistas foi identificar os desafios, melhores práticas e soluções que estão sendo adotadas no campo. Além disso, os estudos de caso fornecem exemplos concretos de como diferentes organizações abordam a preservação distribuída e como superam os desafios de governança, interoperabilidade e certificação. Os especialistas destacaram a importância de estabelecer acordos formais entre as instituições participantes, como contratos de serviços e responsabilidades. Os estudos de caso revelaram que repositórios distribuídos como o MetaArchive Cooperative e o LOCKSS (*Lots of Copies Keep Stuff Safe*) oferecem soluções robustas para a replicação de dados e a preservação de longo prazo, mas enfrentam dificuldades para garantir a integridade dos metadados em uma rede de várias partes. Esses estudos ajudaram a identificar padrões e soluções que devem ser incorporados ao framework.

A última fase metodológica envolveu o desenvolvimento de um *white paper*, um documento consolidado que apresenta o framework proposto com base nas pesquisas realizadas. O *white paper* foi submetido à avaliação da comunidade científica, permitindo revisões e sugestões de melhorias. O objetivo foi garantir que o framework estivesse alinhado às necessidades reais das instituições. A principal contribuição culminou na criação de um modelo híbrido de governança, que combina a autonomia das instituições participantes com

uma coordenação centralizada para garantir a interoperabilidade e a integridade dos dados ao longo do tempo.

Nesse contexto, o modelo híbrido de governança propõe um equilíbrio entre a autonomia das instituições participantes e uma coordenação centralizada, garantindo a integridade dos dados, a replicação eficiente e o alinhamento com diretrizes padronizadas. Esse modelo estrutura-se em três camadas principais: a Camada Central de Coordenação (CCC), os Nodos de Preservação (NP) e as Unidades Autônomas de Preservação (UAP). Cada uma dessas camadas desempenha papéis específicos, permitindo a descentralização dos processos de preservação sem comprometer a interoperabilidade e a confiabilidade do sistema (Becker et al., 2009; Lor; Reis, 2020).

A Camada Central de Coordenação (CCC) é responsável por definir diretrizes gerais para a preservação digital distribuída, estabelecendo padrões técnicos, políticas de auditoria e protocolos de interoperabilidade entre os diferentes repositórios. Segundo os resultados do estudo, essa camada atua como um órgão coordenador, garantindo que os repositórios mantenham a conformidade com as melhores práticas do setor e possam operar de forma colaborativa. Entre as principais funções, destacam-se a certificação de repositórios, a realização de auditorias periódicas e a coordenação dos processos de replicação de dados. Embora desempenhe um papel normativo, a CCC não interfere diretamente na gestão dos repositórios individuais, permitindo que cada instituição mantenha controle sobre suas próprias operações (Zierau e Schultz, 2013).

Os Nodos de Preservação (NPs) são entidades que executam as diretrizes estabelecidas pela CCC, operando como repositórios distribuídos que garantem a replicação e a sincronização dos objetos digitais. Dentro das sugestões de atualização do modelo, os nodos são responsáveis pela implementação das políticas de retenção e descarte, assegurando que os dados sejam armazenados com integridade e acessibilidade ao longo do tempo. Além disso, os NPs desempenham um papel essencial na realização de verificações periódicas de integridade, garantindo que os dados preservados não sofram corrupção ou perda. Embora estejam alinhados às diretrizes centrais, os NPs possuem autonomia operacional, permitindo que cada instituição adapte às

políticas de preservação de acordo com suas necessidades específicas, sem comprometer a interoperabilidade do sistema (Ribeiro; Lucchesi, 2017, adaptado por Zierau; Schultz, 2013, p. 18).

Ainda sobre os resultados do estudo de Zierau e Schultz, 2013, às Unidades Autônomas de Preservação (UAPs) representam instituições independentes que fazem parte da rede distribuída, mas operam com alto grau de autonomia. Essas unidades seriam responsáveis por definir e executar suas próprias políticas de preservação, podendo armazenar dados proprietários que não necessitem de replicação em outros nodos. Além disso, as UAPs deteriam o controle sobre estratégias de acesso e disseminação dos objetos digitais armazenados, permitindo maior flexibilidade na disponibilização dos conteúdos. Embora tenham total autonomia sobre suas operações, as UAPs devem garantir interoperabilidade com os NPs, assegurando que os dados compartilhados sigam os padrões estabelecidos pela CCC. Esse modelo descentralizado possibilita que diferentes instituições participem da rede de preservação sem renunciar a sua independência administrativa e técnica (Zierau E Schultz, 2013).

Dentre as principais vantagens do modelo híbrido de governança, destacam-se sua flexibilidade e escalabilidade, permitindo que instituições de diferentes tamanhos e com distintas capacidades técnicas se integrem à rede sem comprometer a coesão do sistema. Além disso, a distribuição das responsabilidades reduz os custos operacionais e melhora a sustentabilidade da preservação digital a longo prazo. A replicação de dados entre os NPs aumenta a resiliência contra perdas e corrupção de dados, enquanto a governança colaborativa assegura transparência e participação ativa das instituições envolvidas. O modelo também proporciona um equilíbrio entre padronização e autonomia, permitindo que cada organização implemente suas próprias estratégias de preservação sem perder a capacidade de colaborar com a rede (Zierau E Schultz, 2013).

Dessa forma, o modelo híbrido de governança para preservação digital distribuída oferece uma abordagem inovadora que combina a coordenação centralizada da CCC com a autonomia operacional dos NPs e das UAPs (Zierau e Schultz, 2013). Esse modelo permite que diferentes instituições colaborem na

preservação digital sem comprometer sua independência, garantindo a integridade, acessibilidade e interoperabilidade dos dados preservados. Ao superar as limitações do modelo OAIS tradicional, essa estrutura contribui significativamente para a construção de um ambiente de preservação digital distribuído mais robusto e sustentável (Ribeiro e Lucchesi, 2017)

2.2.3 Conclusão dos trabalhos correlatos

A revisão de literatura realizada permitiu estabelecer um panorama sobre o estado atual da preservação digital em instituições Museais, com foco na aplicação do modelo OAIS (Open Archival Information System) como estrutura de referência para assegurar a longevidade e a integridade dos objetos digitais, especialmente em contextos museológicos. O levantamento bibliográfico realizado, concluiu que há um número reduzido de estudos que atendem integralmente aos critérios metodológicos definidos para a presente investigação, sobretudo no que concerne à aplicação do modelo OAIS (ISO 14721) em contextos museológicos e à análise de formatos digitais complexos. Evidenciando uma lacuna significativa na temática, embora amplamente discutida em termos conceituais, ainda carece de pesquisas sistemáticas e aprofundadas que abordem de forma rigorosa e contextualizada as especificidades da preservação digital em instituições culturais. Essa carência de estudos demonstra a necessidade de avanços teóricos e empíricos que contribuam para o aprimoramento de práticas e modelos de preservação digital, favorecendo a consolidação de políticas, metodologias e instrumentos técnicos aplicáveis em diferentes realidades institucionais (Zierau e Schultz, 2013; Corrado e Moulaison Sandy, 2015).

A análise dos estudos selecionados revelou que os desafios da preservação digital vão além das questões técnicas, envolvendo também barreiras institucionais, legais e orçamentárias. A ausência de políticas públicas robustas e de diretrizes normativas unificadas para a preservação de acervos digitais culturais contribui para abordagens fragmentadas e pouco sustentáveis.

Além disso, constatou-se uma escassez de pesquisas voltadas especificamente à preservação digital em museus de pequeno e médio porte, que frequentemente enfrentam restrições significativas de infraestrutura e pessoal qualificado. Nesse sentido, experiências como a do manual do CHIN (2024) e os estudos de Corrado e Moulaison Sandy (2015) oferecem contribuições relevantes, ao propor soluções escaláveis e adaptadas à realidade dessas instituições.

O levantamento bibliográfico também destacou a importância crescente dos softwares de preservação digital — tanto proprietários quanto de código aberto — e a necessidade de selecionar tecnologias alinhadas às capacidades técnicas e financeiras das instituições. Soluções como Archivematica e DSpace foram apresentadas, com ênfase em sua compatibilidade com o modelo OAIS e seu potencial de integração com plataformas de difusão como o Tainacan. Observou-se que, apesar da ênfase na difusão digital dos acervos não há menções claras sobre a inclusão de requisitos voltados a preservação digital de longo prazo, o que reforça a pertinência da presente pesquisa ao investigar como os princípios do modelo OAIS podem ser incorporados à sua arquitetura ou por meio de integração com ferramentas complementares.

A literatura examinada confirma que a preservação digital deve ser concebida como uma responsabilidade institucional de longo prazo, exigindo planejamento estratégico, capacitação de equipes, definição de políticas específicas e a adoção de padrões reconhecidos. A proposta de frameworks híbridos e distribuídos, como os apresentados por Zierau e Schultz (2013), amplia as possibilidades de aplicação do OAIS em ambientes colaborativos, descentralizados e adaptáveis, sugerindo caminhos viáveis para instituições que operam em rede ou com infraestrutura limitada.

Portanto, a revisão evidencia a lacuna existente na aplicação prática do modelo OAIS no contexto dos museus brasileiros. Esta constatação justifica a escolha dos objetivos específicos desta pesquisa: identificar as lacunas entre o OAIS e o Tainacan; avaliar os desafios técnicos, organizacionais e conceituais dessa integração; e propor soluções para aprimorar a preservação digital a longo prazo nesse ambiente. Assim, a presente investigação se propõe a contribuir tanto para o campo teórico da preservação digital quanto para a prática

institucional das políticas públicas de memória, acesso e sustentabilidade dos acervos culturais no Brasil.

2.3 O Modelo OAIS (ISO 14721) e seus componentes

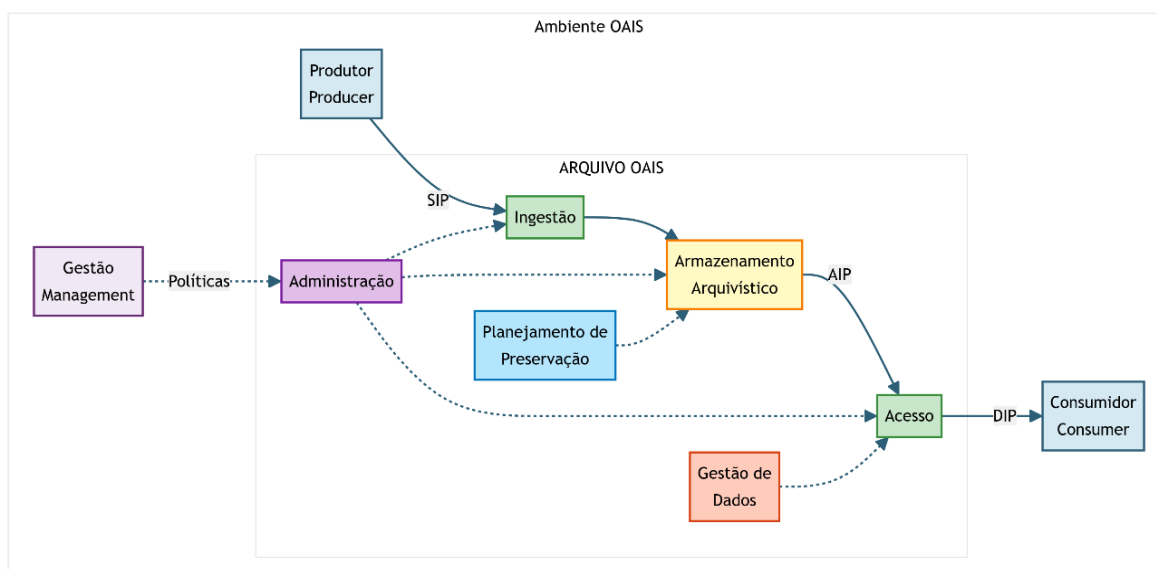
O Modelo de Referência OAIS, formalizado pela norma ISO 14721, constitui a principal estrutura conceitual para a construção e gestão de repositórios digitais de longa duração. Desenvolvido originalmente pelo CCSDS e posteriormente adotado pela ISO, o OAIS fornece uma linguagem comum e um *framework* de alto nível para a preservação digital, sendo amplamente adotado por instituições arquivísticas, agências governamentais e organizações de pesquisa em todo o mundo.

Nesse sentido, essa estrutura conceitual é intencionalmente genérica, não prescrevendo tecnologias ou implementações específicas, tendo como objetivo definir as responsabilidades, as funções e os componentes necessários para que um arquivo possa cumprir sua missão de preservar informações e torná-las acessíveis a uma comunidade designada por um longo período, que pode se estender indefinidamente (International Organization for Standardization, 2012a²¹).

O modelo OAIS opera em um ambiente que envolve a interação entre quatro entidades principais, conforme ilustrado na Figura 2.

²¹ ISO 14721:2012.

Figura 2 - Ambiente OAIS e suas entidades.



Fonte: Elaborado pela autora com base no modelo OAIS.

As entidades que compõem o ambiente OAIS são:

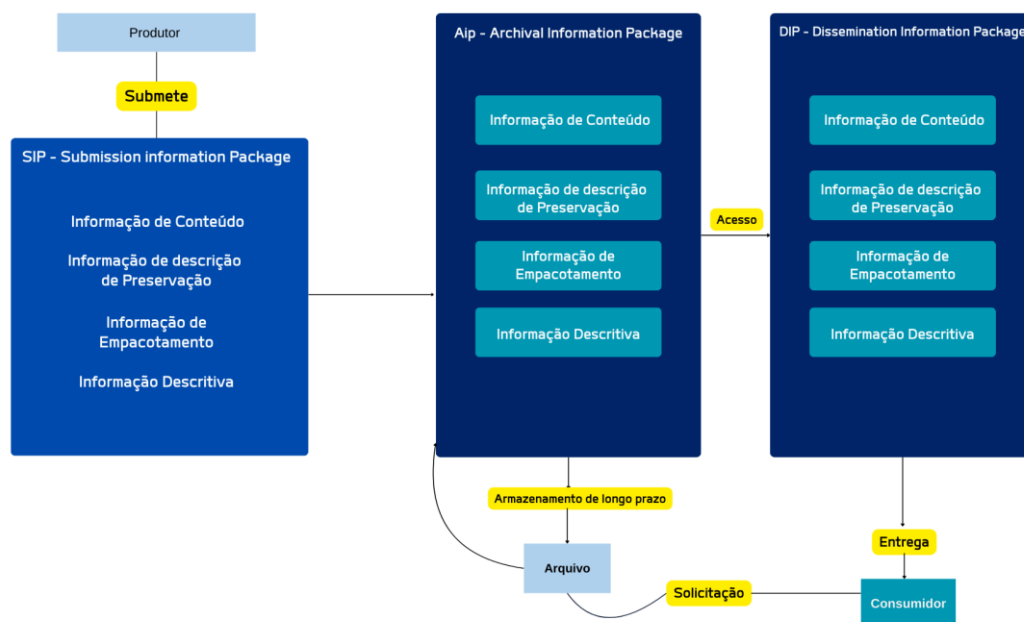
- Produtor (*producer*): a entidade (pessoa ou sistema) que fornece a informação a ser preservada;
- Consumidor (*consumer*): a entidade (pessoa ou sistema) que interage com os serviços do OAIS para encontrar e adquirir informações de interesse. A Comunidade Designada é um subconjunto específico de Consumidores para quem a informação deve ser compreensível;
- Gestão (*management*): a entidade responsável pela formulação de políticas e pela gestão estratégica do OAIS, não se envolvendo nas operações diárias do arquivamento;
- Arquivo (*archive*): a organização que assume a responsabilidade de preservar a informação e torná-la disponível para a Comunidade Designada.

Fonte: (Santos; Flores, 2016; International Organization for Standardization, 2012a).

Ademais, o OAIS define um modelo de informação detalhado para garantir que os objetos digitais permaneçam compreensíveis e utilizáveis ao longo do tempo. O conceito central é o *Information Package* (Pacote de Informação), que

encapsula o conteúdo e os metadados associados. Existem três tipos de pacotes de informação, que transitam em diferentes fases do processo de arquivamento, conforme ilustrado na Figura 3.

Figura 3 - Pacotes de Informação no Modelo OAIS.



Fonte: Elaborado pela autora com base no OAIS (2025).

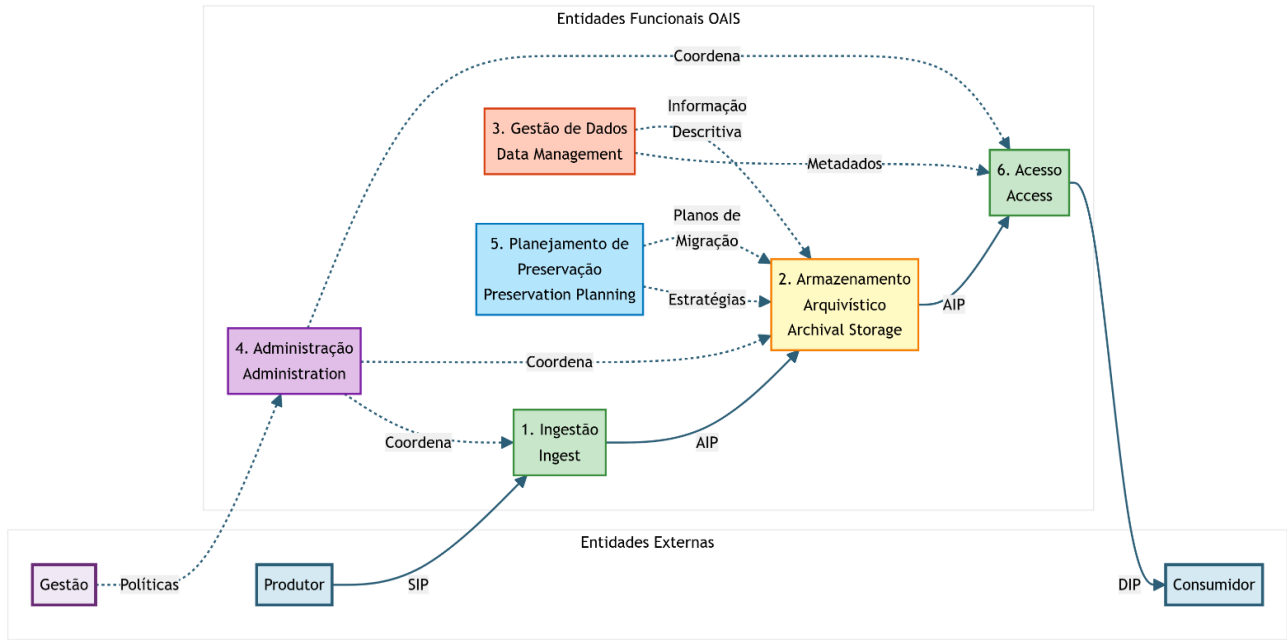
Os três tipos de pacotes são: (1) SIP (Submission Information Package): O Pacote é enviado pelo Produtor ao OAIS o (2) AIP (Archival Information Package): é a versão armazenada pelo OAIS para preservação de longo prazo e o (3) DIP (*Dissemination Information Package*): o pacote disseminação é a versão entregue ao consumidor em resposta a uma solicitação. Um Pacote de Informação é composto por quatro tipos de informação:

1. Informação de conteúdo (*content information*): o objeto de dados em si e a informação de representação necessária para torná-lo compreensível (ex: formatos de arquivo, estrutura de dados, algoritmos);
2. Informação de descrição de preservação (*Preservation Description Information* - PDI): metadados essenciais para a preservação, como proveniência, contexto, fixidez (ex: *checksums*) e direitos;
3. Informação de Empacotamento (*Packaging Information*): informação que une os componentes do pacote;

4. Informação descritiva (*Descriptive Information*): metadados que facilitam a busca e a recuperação do objeto (ex: título, autor, data).
 Fonte: (Santos; Flores, 2016; International Organization for Standardization, 2012a).

O OAIS é composto por seis entidades funcionais que descrevem as atividades necessárias para a operação de um repositório de preservação digital. Além disso, a interação entre essas entidades garante o ciclo de vida completo da informação dentro do arquivo, conforme ilustrado na Figura 4.

Figura 4 - Entidades funcionais do modelo OAIS



Fonte: Elaborado pelo autor com base em OAIS (2025).

O Quadro 2 apresenta uma descrição detalhada de cada entidade funcional.

Quadro 2 - Descrição das entidades funcionais do OAIS

Entidade Funcional	Descrição das Responsabilidades
1. Ingestão (<i>Ingest</i>)	Recebe os SIPs dos produtores, realiza a validação da qualidade, gera um AIP e o submete ao armazenamento arquivístico.

Entidade Funcional	Descrição das Responsabilidades
2. Armazenamento Arquivístico (<i>Archival Storage</i>)	Gerencia o armazenamento de longo prazo dos AIPs, incluindo a substituição de mídias, verificação de integridade e recuperação de AIPs para o Acesso.
3. Gestão de Dados (<i>Data Management</i>)	Mantém o banco de dados com as informações descritivas dos AIPs, gerencia as consultas e atualiza os metadados.
4. Administração (<i>Administration</i>)	Gerencia as operações gerais do arquivo, negocia acordos com os produtores, audita submissões e define políticas e padrões.
5. Planejamento da Preservação (<i>Preservation Planning</i>)	Monitora o ambiente tecnológico e a comunidade designada, desenvolvendo estratégias e planos de preservação para mitigar a obsolescência.
6. Acesso (<i>Access</i>)	Fornecer a interface para os consumidores buscarem e solicitarem informações, gerando e entregando os DIPs a partir dos AIPs correspondentes.

Fonte: Adaptado do OAIS (2025).

Para esta pesquisa, a Entidade funcional de acesso é a entidade que receberá o foco. Neste cenário, o Tainacan não é o ponto de partida, mas, sim, a interface final que entrega a informação ao consumidor. Ele se torna responsável por receber, interpretar e apresentar o Pacote de Informação de Disseminação (DIP), gerado a partir de um Pacote de Informação de Arquivamento (AIP) armazenado em um sistema de preservação. Para que o Tainacan cumpra essa função de maneira aderente ao modelo OAIS, ele deve possuir os seguintes requisitos essenciais de difusão:

Quadro 3 - Requisitos essenciais para que o Tainacan se torne uma Entidade de Acesso com base no OAIS.

Categoria	Requisitos essenciais para que o Tainacan se torne uma Entidade de Acesso com base no OAIS.
Interoperabilidade e Ingestão de DIP	Capacidade de ingerir ou “colher” (<i>harvesting</i>) Pacotes de Informação de Disseminação (DIPs) de um repositório OAIS-compliant, por meio de APIs ou protocolos como OAI-PMH.
Mapeamento de Metadados	Habilidade de interpretar a estrutura do DIP e mapear corretamente os metadados descritivos, técnicos e administrativos para seu próprio modelo de dados, garantindo a exibição correta das informações.

Categoria	Requisitos essenciais para que o Tainacan se torne uma Entidade de Acesso com base no OAIS.
Visualização e Acessibilidade	Fornecer uma interface de usuário robusta e amigável para busca, navegação e visualização dos objetos digitais e seus metadados, cumprindo a função primária de acesso ao acervo preservado.
Gestão de Direitos e Acesso	Capacidade de interpretar e aplicar as políticas de acesso e as informações de direitos autorais contidas no DIP, restringindo ou permitindo o acesso aos objetos digitais, conforme definido pelo sistema de preservação.
Rastreabilidade e Autenticidade	Manter e exibir um vínculo persistente (como um handle ou DOI) que conecte o objeto de disseminação (no Tainacan) ao seu respectivo Pacote de Arquivamento (AIP) no repositório de preservação, assegurando a rastreabilidade e a autenticidade do acervo.

Fonte: Elaborado pela autora (2025).

Dessa forma, esta pesquisa conecta o modelo OAIS à prática das instituições que utilizam o Tainacan, definindo um roteiro de desenvolvimento e configuração que serão apresentados na conclusão deste trabalho (capítulo 6). O diagnóstico desses requisitos oferece uma contribuição direta para a comunidade de preservação digital e para os gestores de acervos que buscam integrar suas plataformas de difusão a ecossistemas de preservação digital aderentes ao padrão OAIS.

2.4 O Modelo Trustworthy Repository Audit and Certification (TRAC)

Se o Modelo OAIS estabelece o *framework* conceitual sobre o que um repositório de preservação digital deve fazer, o documento TRAC emerge como a resposta à questão de como verificar se essas funções estão sendo cumpridas de maneira confiável. Publicado em 2007, o TRAC foi o resultado de um esforço colaborativo entre o *Online Computer Library Center* (OCLC), o *Research Libraries Group* (RLG) e a *National Archives and Records Administration* (NARA), dos Estados Unidos. Seu objetivo principal foi criar um conjunto de critérios auditáveis para avaliar a confiabilidade, o compromisso e a capacidade

de uma instituição em assumir as responsabilidades de preservação digital a longo prazo (TRAC, 2007).

Segundo o Steinhart (2009), o TRAC é uma ferramenta de auditoria e fornece uma *checklist* de métricas que permite a um repositório realizar uma autoavaliação e, subsequentemente, ser auditado por uma entidade externa para obter uma certificação de confiabilidade. Essa verificação externa é fundamental para estabelecer a confiança da comunidade designada e de outras partes interessadas na capacidade do repositório de garantir a integridade e o acesso contínuo ao seu acervo digital (Steinhart, 2009).

Os criadores do modelo OAIS reconheceram que, embora o padrão definisse as responsabilidades de um arquivo, ele não fornecia um mecanismo para verificar se um repositório que se autodeclarava “aderente ao OAIS” de fato cumpria tais responsabilidades (RLG-OCLC WORKING GROUP, 2002). O TRAC foi desenvolvido para preencher essa lacuna, traduzindo os conceitos do OAIS em critérios concretos e mensuráveis (Steinhart, 2009).

O impacto e a relevância do TRAC foram tão significativos que serviram como base para o desenvolvimento de um padrão internacional mais formal. Em 2012, a ISO publicou a norma ISO 16363:2012 – *Space data and information transfer systems — Audit and certification of trustworthy digital repositories* (International Organization for Standardization, 2012b). Essa norma expandiu e formalizou os critérios do TRAC, estabelecendo um padrão global para a auditoria e certificação de repositórios digitais confiáveis (Consultative Committee For Space Data Systems, 2024). Com isso, a ISO 16363 sucedeu ao TRAC como o principal instrumento de certificação, embora os princípios do TRAC permaneçam fundamentais.

O TRAC organiza suas métricas de auditoria em três seções principais, que abrangem todos os aspectos da operação de um repositório digital confiável, conforme apresentado no quadro abaixo.

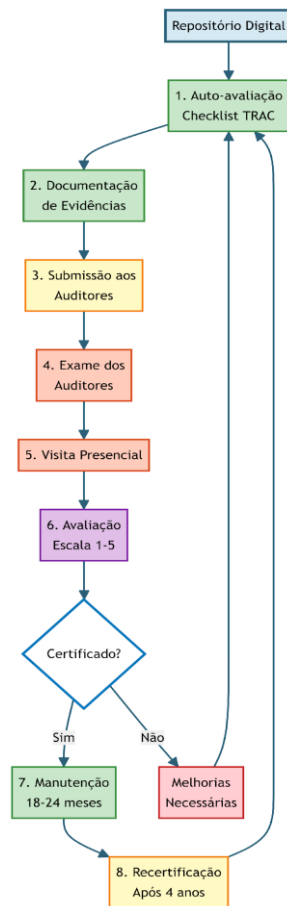
Quadro 4 - Aspectos da operação de um repositório confiável.

Seção	Foco da Avaliação	Exemplos de Critérios
Infraestrutura Organizacional	Viabilidade e sustentabilidade da organização que mantém o repositório	Governança, modelo de negócios, planejamento estratégico, quadro de pessoal, políticas de preservação, ambiente jurídico e ético
Gestão de Objetos Digitais	Funções do ciclo de vida dos objetos digitais dentro do repositório, em alinhamento com as entidades funcionais do OAIS	Aquisição e ingestão de conteúdo (SIP), controle de integridade, gestão de metadados, criação dos pacotes de arquivamento (AIP), estratégias de preservação
Tecnologias, Infraestrutura Técnica e Segurança	Adequação e robustez da infraestrutura tecnológica que suporta o repositório	Arquitetura do sistema, segurança da infraestrutura (contra falhas, ataques e desastres), monitoramento e gestão de componentes de <i>hardware</i> e <i>software</i>

Fonte: Elaborada pela autora com base no TRAC (2025).

O processo de auditoria TRAC é estruturado em etapas sequenciais que garantem uma avaliação rigorosa e transparente da confiabilidade do repositório. A Figura 5 ilustra o fluxo completo do processo de certificação.

Figura 5 - Processo de Auditoria e Certificação TRAC.



Fonte: Elaborada pela autora (2025).

O processo inicia-se com a autoavaliação do repositório, na qual a instituição preenche a *checklist* TRAC, fornecendo evidências documentais e descrevendo suas práticas operacionais para cada métrica. Essas evidências são, então, submetidas aos auditores, que realizam um exame detalhado da documentação. Após essa análise preliminar, os auditores conduzem uma visita presencial ao repositório para verificar *in loco* as informações prestadas e esclarecer eventuais dúvidas (*Consultative Committee For Space Data Systems*, 2024).

Com base nessa avaliação, os auditores atribuem uma nota de 1 a 5 para cada uma das três seções do TRAC. Repositórios que atingem os critérios mínimos recebem a certificação, mas devem passar por uma consulta de manutenção entre 18 e 24 meses para assegurar a continuidade das boas práticas. A certificação tem validade de quatro anos, após os quais o repositório deve submeter-se a uma recertificação completa (*Consultative Committee For Space Data Systems*, 2024).

O conceito de confiabilidade (*trustworthiness*²²) no TRAC fundamenta-se em quatro princípios essenciais que permeiam todos os critérios de avaliação:

- 1 Documentação: todas as afirmações e práticas do repositório devem ser sustentadas por evidências documentais claras e acessíveis;
- 2 Transparência: as políticas, os processos e as decisões do repositório devem ser transparentes para a comunidade designada e para os auditores;
- 3 Adequação: os recursos, as infraestruturas e as práticas devem ser adequados à escala e à natureza do acervo preservado;
- 4 Mensurabilidade: os critérios devem ser objetivos e mensuráveis, permitindo uma avaliação consistente e comparável entre diferentes repositórios.

(Center For Research Libraries, 2007).

O *checklist* do TRAC é composto por 84 critérios, organizados em três seções principais. O Quadro 5 apresenta a estrutura macro, detalhando as subseções e o número de critérios correspondentes dentro de cada uma delas.

Quadro 5 - Estrutura macro do *checklist* TRAC

Seção	Subseção	Nº de Critérios
A. Infraestrutura Organizacional		(28)
	A1. Governança e Viabilidade Organizacional	2
	A2. Estrutura Organizacional e Equipe	3

²² Termo em inglês para “confiabilidade” ou “dignidade de confiança”. No contexto de repositórios digitais, refere-se à demonstração, por meio de evidências auditáveis, de que o repositório é capaz de cumprir sua missão de preservar a longo prazo o acesso a objetos digitais de forma autêntica e fidedigna (International Organization for Standardization, 2012b).

Seção	Subseção	Nº de Critérios
	A3. Responsabilidade Procedimental e Políticas	9
B. Gestão de Objetos Digitais	A4. Sustentabilidade Financeira	5
	A5. Contratos, Licenças e Responsabilidades	5
		(39)
	B1. Ingestão: Aquisição de Conteúdo	8
	B2. Ingestão: Criação do Pacote Arquivável (AIP)	13
	B3. Planejamento de Preservação	4
	B4. Armazenamento e Manutenção de AIPs	5
	B5. Gestão da Informação	4
	B6. Gestão de Acesso	10
C. Tecnologias, Infraestrutura e Segurança		(17)
	C1. Infraestrutura de Sistema	10
	C2. Tecnologias Apropriadas	2
	C3. Segurança	4
Total		84

Fonte: Adaptado pela autora do TRAC.

O TRAC e seu sucessor, a ISO 16363, representam a materialização dos princípios do OAIS em um processo de avaliação. Fornecendo mecanismos

pelos quais um repositório pode não apenas implementar as melhores práticas de preservação digital, mas também demonstrar de forma transparente e verificável seu compromisso com a missão de garantir o acesso futuro à informação digital. Para a presente pesquisa, a compreensão dos critérios do TRAC é fundamental, pois eles estabelecem o padrão de requisitos ao qual o Tainacan deve aspirar ao integrar-se a um ecossistema de preservação digital confiável.

2.5 *Softwares* de aplicação do OAIS: análise comparativa de casos

A aplicação prática do modelo OAIS é viabilizada por um ecossistema de soluções de *software* que traduzem seus conceitos e áreas funcionais em fluxos de trabalho concretos para a preservação digital (cadeia de custódia). Embora o Archivematica seja uma das ferramentas de código aberto mais conhecidas (Flores, 2015), diversas outras plataformas, tanto abertas quanto comerciais, oferecem abordagens distintas para a implementação de repositórios digitais confiáveis baseados no modelo. A análise de soluções como RODA, Preservica, Rosetta e DSpace revela a flexibilidade do modelo e as diferentes estratégias que as instituições podem adotar para garantir a perenidade de seus acervos digitais, cada uma com suas particularidades arquitetônicas, funcionais e de modelo de negócio.

O *software Repository of Authentic Digital Records* (RODA) foi desenvolvido em Portugal pela KEEP SOLUTIONS em colaboração com a Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB). É uma solução robusta para a criação de repositórios digitais confiáveis, com forte ênfase na autenticidade dos registros. Sua arquitetura adere estritamente às unidades funcionais do modelo OAIS, sendo projetado para gerir, preservar e fornecer acesso a diversos tipos de materiais digitais (KEEP SOLUTIONS, 2023).

Além disso, O RODA se destaca por sua abordagem baseada em análise de risco e pelo uso de padrões abertos como Metadata Encoding and Transmission Standard (METS), PREMIS e EAD para garantir a integridade e a proveniência da informação. Disponível em edições como *Community* (código aberto) e *Enterprise* (comercial), o sistema oferece flexibilidade para que diferentes tipos de organizações possam implementar políticas de preservação de longo prazo. Sua capacidade de monitorar continuamente os objetos digitais e alertar sobre riscos de obsolescência o posiciona como uma ferramenta estratégica para arquivos nacionais e grandes corporações que necessitam garantir a perenidade de seus acervos (Ferreira; Ramalho; Farias, 2011).

Contudo, apesar dos pontos positivos de aderência ao OAIS, a implementação do RODA apresenta desafios significativos, os autores Pinheiro, Gloria e Antunes (2018) apontam que a instalação e configuração do sistema são complexas, exigindo um alto nível de conhecimento técnico especializado, o que pode representar uma barreira para instituições com equipes de TI menores ou menos experientes.

Além disso, a interface de administração, embora poderosa, é considerada por alguns usuários como pouco intuitiva, demandando uma curva de aprendizado acentuada. Esses fatores, combinados com a necessidade de uma infraestrutura de *hardware* robusta, podem tornar sua adoção um processo custoso e demorado, especialmente para organizações com recursos limitados (Pinheiro; Glória; Antunes, 2018).

O Preservica é um exemplo de abordagem comercial e proprietária para a preservação digital em conformidade com o OAIS, oferecendo uma solução de ponta a ponta que integra a gestão de ativos digitais com a preservação a longo prazo. Diferentemente de soluções de código aberto, o Preservica é oferecido como um serviço (SaaS - *Software as a Service*), o que atrai instituições que preferem terceirizar a infraestrutura tecnológica e o suporte especializado (PRESERVICA, 2024a). A plataforma é projetada para ser altamente escalável e automatizada, executando um fluxo de trabalho integrado que abrange desde a admissão (ingestão) até a disseminação, aplicando ações de preservação, como a migração de formatos, de forma proativa.

Uma de suas distinções em relação a outros *softwares* é a execução de todos os microsserviços de preservação em uma única etapa durante a ingestão, em contraste com o processo de duas fases, como o do Archivematica (PRESERVICA, 2024b). Essa característica torna o Preservica uma opção competitiva para corporações, agências governamentais e instituições de memória com grandes volumes de dados e requisitos rigorosos de conformidade, mas demanda recorrências financeiras que, muitas vezes, não alcançam a realidade das instituições brasileiras.

O sistema Rosetta, desenvolvido pela empresa *Ex Libris* (atualmente parte do grupo *Clarivate Analytics*), em parceria ativa com diversas bibliotecas nacionais e universidades, incluindo a *British Library*, a *National Library of New Zealand* e a Biblioteca Nacional de Israel, o Rosetta é uma solução de preservação digital de nível empresarial que combina a gestão de ativos digitais com a preservação de longo prazo em uma única plataforma. O sistema foi concebido para gerenciar coleções digitais complexas e em grande escala, desde materiais natos-digitais até os digitalizados, garantindo sua segurança e acessibilidade em conformidade com os mais altos padrões de preservação (*Ex Libris*, 2024).

Um ponto positivo do Rosetta reside em sua abordagem colaborativa e orientada pela comunidade de usuários, que contribui ativamente para a manutenção de recursos essenciais, como uma extensa biblioteca de formatos de arquivo baseada no registro PRONOM, a qual ajuda a identificar e mitigar riscos de obsolescência. Sua arquitetura escalável e flexível permite que as instituições comecem com uma infraestrutura menor e a expandam conforme suas coleções crescem, tornando-o uma escolha estratégica para bibliotecas e arquivos com mandato de preservação perpétua (Vermaaten, 2013).

Embora o Rosetta ofereça compatibilidade com padrões como PREMIS, METS e PRONOM, sua flexibilidade, em termos de interoperabilidade com outros sistemas arquivísticos abertos, é limitada. Essa limitação pode comprometer estratégias integradas de preservação digital distribuída ou soluções híbridas voltadas para redes colaborativas, tal como discutido por Zierau e Schultz (2013). Para contextos nos quais a descentralização, a

interoperabilidade e a progressiva adequação a padrões internacionais são centrais, como proposto nesta tese, a adoção de sistemas proprietários e monolíticos, como o Rosetta, pode restringir a experimentação, a inovação metodológica e a adaptação às realidades locais.

Outro ponto crítico diz respeito à complexidade técnica do Rosetta, que demanda equipes altamente especializadas para sua implementação, parametrização e operação contínua. Isso representa uma barreira relevante para instituições que não contam com setores de tecnologia da informação bem estruturados, comprometendo a sustentabilidade operacional da solução no longo prazo (Becker *et al.*, 2009).

O DSpace, embora seja mais conhecido como um *software de difusão* para dar visibilidade à produção acadêmica, também pode ser configurado para funcionar como um ambiente de preservação digital alinhado ao modelo OAIS. Desenvolvido originalmente pelo MIT e pela HP, o DSpace é uma plataforma de código aberto amplamente adotada por universidades e instituições de pesquisa (Lyrasis, 2024). Sua estrutura permite a gestão de diversos formatos de arquivo e a aplicação de metadados descritivos e de preservação.

A conformidade com o OAIS no DSpace, no entanto, geralmente requer adaptações e a integração com outras ferramentas, uma vez que seu foco principal é o acesso e a disseminação, e não um fluxo de trabalho de preservação ativa tão detalhado quanto o de sistemas como Archivematica ou RODA. Apesar disso, sua flexibilidade e a vasta comunidade de desenvolvedores permitem que ele seja utilizado como a camada de armazenamento e acesso de um sistema de preservação, cumprindo parcialmente as funções de um repositório digital confiável (Biazus; Santos, 2020).

O Archivematica, desenvolvido pela Artefactual Systems, é um sistema de preservação digital de código aberto consolidado no cenário nacional. Estruturado de forma aderente ao modelo OAIS (International Organization for Standardization, 2012a), o Archivematica oferece funcionalidades robustas para a ingestão, verificação de integridade, normalização, controle de metadados de

preservação, além de exportações compatíveis com repositórios de acesso, como o DSpace ²³e o AtoM.

No contexto brasileiro, sua adoção é fortemente incentivada por normativas arquivísticas oficiais. A Portaria nº 47/2020 do Arquivo Nacional, que institui os Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos (E-Arq Brasil), reconhece a conformidade com o modelo OAIS como diretriz essencial para a construção de RDC-Arq.

O mesmo ocorre com a Portaria MEC nº 315/2018, que torna obrigatória, no âmbito do Ministério da Educação, a adoção do RDC-Arq, estrutura normativa dentro da qual o Archivematica se apresenta como ferramenta tecnicamente recomendada para a camada de preservação digital. No Poder Judiciário, a portaria conjunta do Conselho Nacional de Justiça (CNJ) nº 15/2018 também estabelece que os órgãos devem adotar soluções compatíveis com o modelo OAIS, tendo o Archivematica sido frequentemente indicado e incorporado a projetos-piloto e estratégias institucionais.

Sua arquitetura modular e a natureza *open source* ²⁴ do Archivematica oferecem ampla possibilidade de customização, permitindo sua adequação a diferentes contextos institucionais de preservação digital. Isso se deve ao fato de que o sistema é composto por uma série de microsserviços que seguem as entidades funcionais do modelo OAIS, possibilitando que instituições configurem seus próprios *workflows* de ingestão, normalização, verificação de integridade e armazenamento arquivístico (Corrado; Moulaison Sandy, 2015; Artefactual Systems, 2023).

Além disso, por ser desenvolvido sob licença pública, o Archivematica permite controle completo do código-fonte, favorecendo a soberania tecnológica, transparência e autonomia institucional, aspectos frequentemente destacados

²³ O DSpace é uma das plataformas de *software* de código aberto mais difundidas para a criação de repositórios institucionais. Embora seja altamente competente na gestão e disseminação de objetos digitais, sua arquitetura é frequentemente descrita como monolítica, tratando as funções de gestão, preservação e acesso de forma integrada, em contraste com a abordagem desacoplada (preservação + difusão) do ecossistema Archivematica/AtoM (Lyrasis, 2023).

²⁴ Termo em inglês para “código aberto”. Refere-se a um modelo de desenvolvimento de software cuja licença de uso garante aos usuários o direito de acessar, estudar, modificar e redistribuir o código-fonte do programa livremente (Open Source Initiative, [2023]).

como críticos para instituições públicas e arquivísticas que buscam independência de fornecedores comerciais (Corrado; Moulaison Sandy, 2015; Artefactual Systems, 2023)

No entanto, sua implantação exige considerações técnicas importantes. A instalação do Archivematica pressupõe infraestrutura computacional compatível com servidores Linux, domínio sobre configuração de ambientes virtuais, gerenciamento de pacotes e integração com bancos de dados, sendo recomendada para instituições com equipes de tecnologia da informação (TI) capacitadas em ambientes Unix-like²⁵ (Artefactual Systems, 2023). A configuração de *pipelines*²⁶ e o monitoramento dos processos de preservação digital requerem conhecimento técnico intermediário a avançado, além de práticas institucionais maduras de gestão de metadados e governança da informação.

Mesmo diante dessas limitações, o Archivematica permanece como uma das soluções mais amplamente alinhadas aos princípios da preservação digital confiável, por sua conformidade com padrões internacionais, escalabilidade progressiva e compatibilidade com o modelo de RDC-Arq, tal como regulamentado no Brasil pelas Portarias MEC nº 315/2018 e CNJ nº 15/2018 (Brasil, 2018a; Brasil, 2018b).

Por fim, a relevância do Archivematica no cenário internacional de preservação digital é também evidenciada por estudos independentes que avaliam a eficácia e a sustentabilidade de plataformas utilizadas por instituições que possuem acervos informacionais. Destaca-se, nesse sentido, o relatório conduzido pela Ithaka (2020), intitulado *“The Effectiveness and Durability of Digital Preservation and Curation Systems”*, que examinou criticamente os

²⁵ Termo que designa um sistema operacional que se comporta de maneira semelhante ao sistema Unix original, mas cujo código-fonte não deriva diretamente dele. Sistemas *Unix-like* geralmente compartilham a filosofia de design do Unix, como o uso de uma interface de linha de comando, um sistema de arquivos hierárquico e um conjunto de ferramentas modulares que podem ser combinadas (Tanenbaum; Bos, 2016).

²⁶ Em computação, um *pipeline* (duto ou canalização) é um modelo no qual a saída de um processo é conectada diretamente à entrada de um segundo processo, e assim por diante. Isso cria uma cadeia de processamento onde os dados fluem através de uma sequência de módulos, cada um executando uma tarefa específica (Tanenbaum; Bos, 2016).

principais sistemas adotados por universidades, bibliotecas e arquivos em contextos institucionais diversos.

O estudo incluiu o Archivemática como um dos oito sistemas centrais avaliados em termos de aderência a padrões de preservação digital, como OAIS, com capacidade de automação, governança da informação e viabilidade de longo prazo. O relatório enfatiza que, embora o Archivemática apresente desafios relacionados à sua implementação técnica e à necessidade de manutenção ativa, por parte das instituições usuárias, ele se destaca por sua forte base conceitual, transparência operacional e adaptabilidade a diferentes ambientes tecnológicos. Essas são características que o posicionam como uma ferramenta estratégica para instituições que buscam construir RDC-Arq de forma autônoma e sustentável (Ithaka, 2020).

Adicionalmente, o relatório da Ithaka (2020) destaca que o Archivemática é frequentemente escolhido por instituições acadêmicas e arquivísticas que privilegiam soluções abertas, com controle sobre suas infraestruturas, em contraste com modelos comerciais dependentes de fornecedores externos. Isso reforça seu papel dentro de um ecossistema informacional comprometido com os princípios da ciência aberta, da transparência institucional e da soberania informacional, valores centrais às diretrizes arquivísticas brasileiras contemporâneas, como demonstram as Portarias do MEC nº 315/2018, CNJ nº 15/2018 e Arquivo Nacional nº 47/2020.

Diante das possibilidades de aplicação do modelo OAIS analisadas nesta seção, demonstra-se que o Archivemática se apresenta como a ferramenta mais adequada para a implementação no contexto proposto por esta tese. Sua arquitetura modular, baseada em microsserviços, permite a incorporação precisa das seis entidades funcionais do OAIS, promovendo não apenas a conformidade com a ISO 14721:2012, mas também a replicação do checklist do TRAC. Em contraste com soluções comerciais como o Preservica e o Rosetta, o Archivemática apresenta vantagens cruciais para instituições públicas brasileiras: é uma ferramenta de código aberto, de uso gratuito, que respeita os princípios da soberania tecnológica e da interoperabilidade, valores diretamente vinculados às diretrizes normativas nacionais, ampliando sua aplicabilidade em

repositórios digitais confiáveis com vocação para a disseminação pública de acervos, como o Tainacan.

2.6 Archivematica como *software* de aplicação do modelo OAIS: Análise da arquitetura de integração com o *software* de difusão (AtoM)

A preservação digital de longo prazo demanda a implementação de sistemas robustos que não apenas garantam a integridade e a autenticidade dos objetos digitais, mas que também sigam modelos conceituais consolidados na comunidade arquivística. Nesse contexto, o modelo de referência do OAIS, formalizado pela norma ISO 14721:2012, estabelece um *framework* funcional para a gestão de acervos digitais.

O *software* de código aberto Archivematica foi desenvolvido com o propósito de ser uma implementação prática desse modelo, abrangendo suas principais entidades funcionais, como Admissão (*ingest*), Armazenamento Arquivístico (*archival storage*), Administração de Dados (*data management*), Planejamento da Preservação (*preservation planning*) e Acesso (*access*). A sua função primordial é assegurar que os objetos digitais permaneçam autênticos, acessíveis e utilizáveis ao longo do tempo, mitigando riscos, como a obsolescência tecnológica (Artefactual Systems, 2023).

A arquitetura do Archivematica foi projetada para operacionalizar a criação dos pacotes de informação preconizados pelo modelo OAIS. Ao receber um conjunto de objetos digitais, o sistema o encapsula em um SIP, aplicando, em seguida, uma série de microsserviços que executam ações de preservação, como a identificação de formatos, a verificação de vírus e a normalização para formatos mais estáveis.

Como resultado, o sistema gera dois pacotes distintos: o Pacote de Informação de Arquivamento (AIP), que contém o objeto digital original, suas versões normalizadas e um conjunto exaustivo de metadados de preservação (PREMIS), destinado ao armazenamento seguro e perene; e o Pacote de

Informação de Disseminação (DIP), contendo versões de acesso dos objetos e seus metadados descritivos, preparado para a difusão (Duranti, 2009; Guercio, 2017).

Contudo, o Archivematica, enquanto repositório de preservação (RDC-Arq), não foi concebido para ser a interface final de acesso ao usuário. Para completar o ecossistema de gestão arquivística, sua arquitetura prevê a integração com um *software* de difusão. A solução mais proeminente para essa finalidade é o AtoM, também desenvolvido pela Artefactual Systems.

O AtoM é uma aplicação web para descrição arquivística baseada nas normas do Conselho Internacional de Arquivos (ICA), como a ISAD(G), e projetada especificamente para a publicação e consulta online de acervos (Artefactual Systems, 2023). A integração entre as duas plataformas materializa a separação funcional entre o ambiente de preservação e o ambiente de acesso, uma prática recomendada para garantir a segurança da cadeia de custódia e a integridade dos pacotes arquivados (AIPs), que não devem ser expostos ao acesso público direto (Santos; Flores, 2019; ISO 16363, 2012)

A comunicação entre o Archivematica e o AtoM é viabilizada por uma arquitetura de serviços que depende da configuração de APIs e protocolos específicos. O fluxo de integração se inicia no final do processamento no Archivematica, quando, após a geração do DIP, o sistema pode ser instruído a enviá-lo para uma instância do AtoM. Esse envio é tecnicamente realizado por meio de uma API REST²⁷ e, frequentemente, utiliza o protocolo SWORD (*Simple Web-service Offering Repository Deposit*), um padrão para o depósito de pacotes em repositórios (Artefactual Systems, 2023; ISO 16363, 2012).

Para que a comunicação seja bem-sucedida, é necessário que plugins específicos, como o REST API *Plugin* e o SWORD *Plugin*, estejam ativados e configurados no AtoM, e que as credenciais de acesso e o endereço do servidor de destino sejam corretamente inseridos no painel de administração do

²⁷ *Representational State Transfer* (REST) constitui um estilo arquitetural que define princípios para o design de serviços web focados em recursos do sistema, utilizando métodos HTTP explícitos e mantendo a comunicação stateless entre cliente e servidor (Rodriguez, 2008).

Archivematica. Um serviço de *worker*²⁸ no servidor AtoM fica, então, responsável por processar a fila de pacotes recebidos e integrá-los à estrutura de descrição arquivística da plataforma (Artefactual Systems, 2023).

Apesar da robustez dessa arquitetura, a interoperabilidade de metadados apresenta desafios. O envio automatizado de metadados descritivos do Archivematica para o AtoM é otimizado para o padrão Dublin Core (DC). Embora ambos os sistemas suportem padrões de descrição arquivística mais complexos, a sincronização completa e sem perdas de hierarquias e metadados específicos pode exigir ajustes manuais ou desenvolvimentos adicionais (Santos; Flores, 2019; ISO 16363, 2012).

Outro desafio reside na natureza predominantemente unidirecional do fluxo, do Archivematica para o AtoM, o que pode complicar a gestão de acervos que já possuem uma estrutura descritiva preexistente no AtoM e para os quais se deseja criar um vínculo de preservação retroativo. Não obstante, a integração entre Archivematica e AtoM representa um dos ecossistemas de código aberto mais completos para a gestão de acervos digitais, alinhando as funções de preservação e acesso de acordo com as melhores práticas e os modelos conceituais da arquivologia contemporânea.

Para compreender melhor essa integração, pensemos nas aplicações com exemplos lúdicos: o Archivematica é o “cofre de preservação” que conversa com o AtoM a “vitrine de exposição”, pode-se imaginar um sistema de correio especializado e seguro entre dois departamentos institucionais.

1. O REST API Plugin: é o endereço e as regras de correspondência, pode-se interpretar a interface de programação de aplicações (API) como um “cardápio” oferecido por um *software*, listando exatamente o que um outro *software* pode pedir e como deve pedir. O estilo transferência de estado representacional (REST) é um conjunto de regras de arquitetura popular para

²⁸ No contexto de arquitetura de software, o termo *worker* designa um componente responsável pela execução de tarefas, com uso intensivo de recursos, fluxos de trabalho de longa duração ou trabalhos em lote. O *worker* opera de forma assíncrona e desacoplada do *front-end*, comunicando-se por meio de filas de mensagens, o que permite o dimensionamento independente e a separação clara de responsabilidades no sistema (Microsoft, 2026).

criar essas APIs, conhecido por sua simplicidade e eficiência, usando os mesmos verbos do protocolo HTTP²⁹ que se usam para navegar na web (Fielding, 2000).

2. Função no AtoM: ao ativar o REST API Plugin no AtoM, estamos abrindo um “guichê de atendimento” oficial. O AtoM passa a ter um endereço público e regras claras, seria o “cardápio”, citado acima, que permite que sistemas externos, como o Archivemática, possam interagir com ele de forma programática e segura. Permite, por exemplo, que um sistema autorizado crie descrições arquivísticas, anexe objetos digitais a elas, ou leia informações existentes (Artefactual Systems, 2023; Fielding, 2000).

Ou seja, o REST API *Plugin* é como instalar uma “caixa de correio” na porta do AtoM. Essa caixa de correio tem um endereço específico, a URL³⁰ da API, e um manual de instruções, que norteia o envio dos documentos, baseando-se nos protocolos a serem seguidos para que sejam entregues legíveis e traduzíveis pelo destinatário, o AtoM. Sem essa caixa de correio, o Archivemática não teria para onde enviar o pacote (Artefactual Systems, 2023; Fielding, 2000).

3. O SWORD *Plugin*: é “o carteiro especializado”. O *Simple Web-service Offering Repository Deposit* (SWORD) é um protocolo construído sobre a API REST, mas com foco único em depositar conteúdo em repositórios. Sua única função é tornar o processo de envio de pacotes de informação o mais simples e padronizado possível. Foi criado para cenários como o sistema que produz pacotes (Archivemática) e outro que os recebe (AtoM) (Lewis; Castro; Jones, 2012; Fielding, 2000).

Portanto, enquanto a API REST do AtoM abre a porta para várias interações, o SWORD Plugin cria um protocolo de recebimento para os Pacotes de Informação de Disseminação (DIPs). O Archivemática traduz o SWORD,

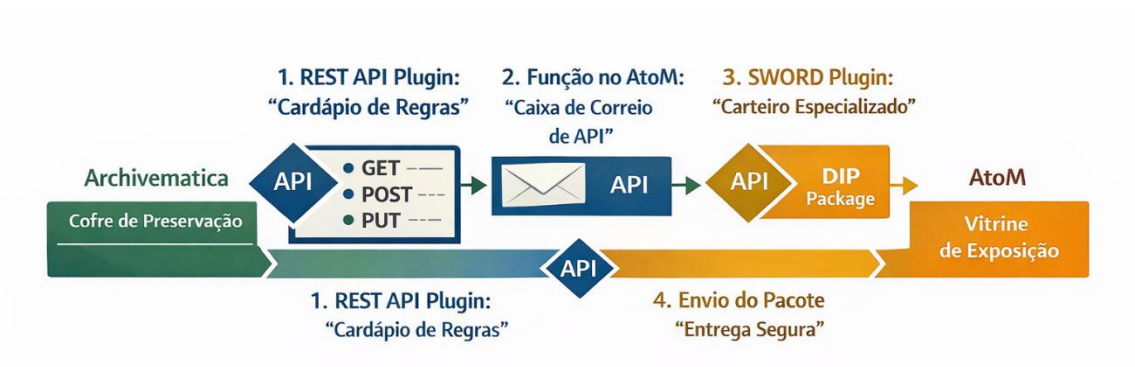
²⁹ O Protocolo de Transferência de Hipertexto (HTTP) é um protocolo de camada de aplicação para sistemas de informação distribuídos, colaborativos e de hipermídia. Ele funciona como um protocolo de requisição-resposta no modelo cliente-servidor, sendo a base para a comunicação de dados na World Wide Web (Fielding *et al.*, 1999).

³⁰ A Uniform Resource Locator (URL) é um subconjunto dos Identificadores Uniformes de Recursos (URI) que especifica onde um recurso está disponível e o mecanismo para recuperá-lo. A sintaxe e a semântica da URL foram formalizadas por Tim Berners-Lee, o inventor da World Wide Web, e seus colaboradores, estabelecendo um padrão para a localização de recursos na Internet (Berners-Lee; Masinter; Mccahill, 1994).

sabendo exatamente como empacotar o DIP e seus metadados para que o AtoM os receba e processe corretamente (Artefactual Systems, 2023; Lewis; Castro; Jones, 2012).

Para elucidar a interação entre os sistemas, ilustrou-se abaixo a sequência de operações e a transferência de informação desde a finalização do processamento no Archivematica até a disponibilização do acervo no AtoM.

Figura 6 - Estrutura de integração entre o Archivematica e o AtoM simplificada.



Fonte: Elaborado pela autora (2025).

Para a completa compreensão dos passos de integração, o quadro abaixo separa e descreve cada elemento:

Quadro 6 - Etapas do fluxo de integração entre Archivematica e AtoM

Etapa	Sistema de Origem	Ação Executada	Descrição da Etapa
1. Geração do Pacote	Archivematica	Geração do DIP	Após o processamento do acervo, o sistema cria o Pacote de Informação de Disseminação (DIP), que contém as versões de acesso dos objetos digitais e os metadados descritivos básicos.
1.2. Empacotamento	Archivematica	Atuação como Cliente SWORD	O sistema empacota o DIP e os metadados em um formato padronizado pelo protocolo SWORD, preparando o "malote" para o envio ao sistema de acesso.
1.3. Envio e Requisição	Archivematica	Envio via HTTP para a API REST	A transação é enviada como uma requisição HTTP (método POST) para o endereço (endpoint) da API REST do AtoM, que está "escutando" por novas submissões.
2. Autenticação	AtoM	Validação da Chave de API	O AtoM recebe a requisição e verifica a chave de API enviada. Essa etapa é crucial para a segurança, garantindo que apenas sistemas autorizados possam depositar conteúdo.
2.1. Recebimento	AtoM	Acionamento do Servidor SWORD	Com a autenticação bem-sucedida, o <i>plugin</i> do servidor SWORD no AtoM é acionado para receber e manusear o pacote de acordo com as regras do protocolo.
2.2. Processamento	AtoM	Desempacotamento pelo <i>Worker</i>	Um processo assíncrono (<i>worker</i>) é iniciado para desempacotar o DIP, extrair os arquivos de objeto digital e o arquivo de metadados (ex: METS.xml).
2.3. Publicação	AtoM	Criação/Atualização da Descrição	O <i>worker</i> processa as informações extraídas para criar uma descrição arquivística no AtoM ou atualizar uma existente, associando os objetos digitais a ela e tornando o acervo visível na interface pública.

Fonte: Elaborado pela autora (2025), com base na documentação técnica do Archivematica e do AtoM (Artefactual, 2012).

O fluxo de dados que caracteriza a integração entre o Archivematica e o AtoM evidencia uma separação clara entre as funções de preservação e acesso. O processo é iniciado no ambiente de preservação (Archivematica) após a conclusão bem-sucedida da geração do Pacote de Informação de Disseminação (DIP). Nesse ponto, o sistema, atuando como um cliente SWORD, empacota os objetos digitais de acesso e seus metadados descritivos. A transação é, então, enviada via protocolo HTTP para o *endpoint* ³¹ da API REST do AtoM, que, por

³¹ Um endpoint de API é uma localização digital, tipicamente uma URL (Uniform Resource Locator), onde uma interface de programação de aplicações (API) recebe requisições para

sua vez, aciona o *plugin* do servidor SWORD. Após a validação da chave de API, que garante a segurança da comunicação, um processo assíncrono (*worker*) no AtoM é encarregado de desempacotar o DIP, processar seu conteúdo e, finalmente, criar ou atualizar a descrição arquivística correspondente, tornando o acervo acessível ao usuário final. Essa arquitetura em etapas garante um desacoplamento entre os sistemas, no qual uma falha no ambiente de acesso não impacta a integridade do repositório de preservação.

A estrita aderência do Archivematica ao modelo de referência OAIS, operando como um RDC-Arq, e sua subsequente comunicação com o AtoM, um sistema especializado na difusão, confirmam a relevância de uma arquitetura que separa claramente as funções de preservação e acesso. Essa separação não é apenas uma recomendação técnica, mas um pilar para a garantia da cadeia de custódia, da autenticidade e da integridade dos objetos digitais a longo prazo (Flores, 2015).

A investigação dos componentes técnicos, como a API REST e o protocolo SWORD, elucidou o fluxo de comunicação, revelando um mecanismo padronizado e automatizado para o arquivamento seguro de informações. Demonstrou-se que, embora a configuração padrão ofereça uma solução de grande valia pela sua simplicidade, ela impõe desafios relacionados à interoperabilidade de metadados descritivos fora do padrão de recebimento. A discussão sobre estratégias alternativas, como o uso de *scripts*³² personalizados ou fluxos de trabalho em múltiplas etapas, evidencia que a implementação de um ecossistema digital completo exige uma análise criteriosa das necessidades de descrição arquivística de cada instituição.

Dessa forma, a análise feita neste capítulo desempenha-se como um alicerce para a argumentação central desta tese, ao estabelecer como um

acessar ou manipular um recurso específico em seu servidor. A interação com o *endpoint* é realizada por meio de métodos do protocolo HTTP, como GET, POST, PUT ou DELETE, que definem a operação a ser executada sobre o recurso (Nosowitz; Goodwin, 2025).

³² *Scripts* correspondem a conjuntos de instruções escritas em linguagens de programação ou de automação, geralmente interpretadas, utilizadas para executar tarefas específicas de forma automatizada, como processamento de dados, manipulação de arquivos, integração entre sistemas ou execução de rotinas técnicas em ambientes computacionais. No contexto de sistemas informacionais e de preservação digital, os *scripts* desempenham papel fundamental na padronização de fluxos de trabalho, na reprodutibilidade de processos e na mitigação de erros humanos (Silberschatz; Galvin; Gagne, 2015)

sistema de preservação digital pode ser efetivamente implementado e integrado, estabelecendo as bases para a próxima etapa da pesquisa: “Aplicação do OAIS a softwares de difusão: Análise Comparativa de métodos de integração.” A compreensão aprofundada dessa arquitetura é, portanto, condição essencial para se avançar na discussão sobre a aplicação prática, os desafios operacionais e o valor estratégico da preservação digital no contexto arquivístico do Tainacan.

2.7 Aplicação do OAIS a *softwares* de difusão: Análise Comparativa de métodos de integração.

A implementação de um ecossistema de preservação digital alinhado ao modelo de referência OAIS transcende a escolha de um *software* de preservação. Um dos aspectos críticos para a funcionalidade do sistema reside na maneira como a entidade funcional de Acesso se materializa e se comunica com as demais funções, especialmente com o Armazenamento Arquivístico (Santos; Flores, 2018).

Nesse sentido, a entidade de Acesso é responsável por receber requisições da Comunidade Designada, processá-las e fornecer os Pacotes de Informação de Disseminação (DIPs) correspondentes (International Organization for Standardization, 2012a). Em arquiteturas modernas, essa função é frequentemente desempenhada por um *software* de difusão especializado, como o AtoM ou o DSpace, que atua como a interface pública do acervo. A integração eficaz entre o repositório de preservação (RDC-Arq) e esse *software* de difusão é, portanto, imperativa.

A análise da literatura técnica e de estudos de caso permite identificar três métodos principais de integração, cada um com distintas características de complexidade, custo e flexibilidade. A escolha entre eles representa uma decisão estratégica que impacta diretamente a gestão do acervo digital (Lewis; Castro; Jones, 2012).

Integração via Protocolo Padronizado (Ex: SWORD): Esse método, como analisado anteriormente no contexto do Archivematica e AtoM, utiliza um protocolo de comunicação desenhado especificamente para o depósito de conteúdo em repositórios. O protocolo SWORD padroniza o formato do pacote a ser transferido e simplifica a transação a uma única requisição, abstraindo a complexidade da interação direta com a API do *software* de difusão (Lewis; Castro; Jones, 2012).

A principal vantagem reside na simplicidade de implementação e na alta interoperabilidade entre sistemas que suportam o mesmo padrão. A configuração é geralmente rápida e a manutenção é facilitada, pois a lógica da integração é mantida pelos desenvolvedores dos próprios *softwares*. A desvantagem é que a padronização inerentemente limita a flexibilidade. Conforme apontado por Santos e Flores (2019), o fluxo de metadados via SWORD é frequentemente restrito a esquemas mais simples, como o Dublin Core, o que pode levar à perda de riqueza descritiva se a instituição utiliza padrões arquivísticos complexos (Santos; Flores, 2019; ISO 16363, 2012).

A Integração via *scripts* personalizados e API Direta, nessa abordagem, a instituição desenvolve um *software* intermediário (*middleware*) ou um conjunto de scripts que orquestram a comunicação entre o repositório de preservação e o *software* de difusão. Esses scripts interagem diretamente com a API REST (ou outra API disponível) de ambos os sistemas para extrair, transformar e carregar os dados, processo conhecido como *Extract, Transform, Load* (ETL).

A partir dessa configuração técnica, como vantagem, destaca-se o oferecimento de máxima flexibilidade. É possível mapear com precisão qualquer esquema de metadados, criar lógicas de negócio complexas e integrar sistemas que não possuem um protocolo comum. Essa abordagem permite a total aderência às especificidades da descrição arquivística da instituição (Lewis; Castro; Jones, 2012). Por outro lado, tendo como desvantagem, destacam-se o custo e a complexidade de implementação. Exige expertise em programação, conhecimento aprofundado das APIs e um esforço contínuo de manutenção, uma vez que qualquer atualização nos sistemas de origem ou destino pode quebrar a integração. A solução torna-se dependente do conhecimento de seus desenvolvedores (Lewis; Castro; Jones, 2012).

A Integração Manual ou Semiautomatizada (baseada em *watch folders*³³), esse método representa a abordagem de menor complexidade técnica, o repositório de preservação é configurado para exportar os DIPs para um diretório de rede específico, - uma "*watch folder*" ou pasta monitorada. O software de difusão, por sua vez, é configurado para monitorar esse diretório e importar automaticamente qualquer novo pacote que ali seja depositado. Esse método é o mais simples e rápido de configurar, exigindo pouca ou nenhuma programação. É uma solução robusta para fluxos de trabalho mais simples e com menor volume de dados (Lewis; Castro; Jones, 2012).

Contudo, sua desvantagem é a capacidade limitada de transferir metadados estruturados, muitas vezes, dependendo apenas do nome do arquivo ou de um arquivo de texto simples dentro do pacote, não havendo um canal de comunicação para retorno de status ou erros, tornando o diagnóstico de falhas um processo manual (Lewis; Castro; Jones, 2012). Segundo a documentação de sistemas como o DSpace, embora o método de *Watch Folder* seja funcional, ele carece da sofisticação das integrações baseadas em API (DSPACE, 2023).

Quadro 7 - Análise comparativa dos métodos de integração

Método de Integração	Complexidade Técnica	Custo de Implementação	Flexibilidade de Metadados	Manutenibilidade
Protocolo Padronizado (SWORD)	Baixa	Baixo	Baixa a Média	Alta
Scripts Personalizados (API Direta)	Alta	Alto	Alta	Baixa
Manual / <i>Watch Folder</i>	Muito Baixa	Muito Baixo	Muito Baixa	Muito Baixa

Fonte: Elaborado pela autora (2025).

³³ Uma *watch folder* (pasta monitorada), também conhecida como *hot folder*, é um diretório em um sistema de arquivos que um software monitora continuamente. Quando um arquivo é depositado na pasta, uma operação pré-configurada, como um fluxo de trabalho, um serviço ou um *script*, é automaticamente iniciada para processar esse arquivo. Esse método constitui uma forma de automação para a integração de tarefas, sendo amplamente utilizado em sistemas de gerenciamento de conteúdo, transcodificação de mídia e fluxos de trabalho de impressão (Adobe, 2025).

A escolha do método de integração deve ser uma decisão informada, que pondere a riqueza dos metadados descritivos da instituição, os recursos técnicos e financeiros disponíveis e a escalabilidade desejada para o futuro. Não existe uma solução universalmente superior. A análise comparativa demonstra que, enquanto protocolos como o SWORD oferecem simplicidade, abordagens baseadas em API direta, embora mais complexas, proporcionam uma flexibilidade indispensável para instituições com modelos de metadados ricos.

Nesse contexto, o *software* brasileiro Tainacan, construído como um *plugin* sobre uma plataforma de API como o *WordPress*, exemplifica a prontidão para integrações avançadas (Tainacan, 2025). Sua arquitetura, nativamente orientada a serviços via API REST, posiciona-o favoravelmente para receber dados de qualquer sistema de preservação compatível com o OAIS, permitindo o desenvolvimento de *scripts* personalizados que podem mapear com precisão hierarquias e metadados complexos, superando as limitações de métodos mais rígidos. A abordagem mais adequada, portanto, é aquela que melhor se alinha à realidade e aos objetivos estratégicos da organização detentora do acervo.

2 METODOLOGIA

Este capítulo detalha os procedimentos metodológicos adotados para responder à pergunta de pesquisa. A investigação caracteriza-se por uma abordagem qualitativa, de natureza aplicada, e utiliza como procedimentos técnicos a pesquisa bibliográfica, a análise documental e o estudo de caso instrumental. A estrutura metodológica foi desenhada para ser sistemática e replicável, permitindo a validação dos resultados e a aplicação do modelo proposto em outros contextos institucionais.

3.1 Natureza e abordagem da pesquisa

A definição clara da natureza e da abordagem metodológica é um pressuposto científico, pois estabelece os “limites e as regras do jogo” da investigação (Marconi; Lakatos, 2017, p. 83). Quanto à sua natureza, ou finalidade, esta pesquisa classifica-se como pesquisa aplicada. Diferentemente da pesquisa pura ou básica, que se dedica à geração de novos conhecimentos e teorias sem uma preocupação imediata com sua utilização prática, a pesquisa aplicada tem como objetivo primordial a solução de problemas concretos e a geração de conhecimentos com aplicação direta em casos reais (Gil, 2008). No campo da Ciência da Informação, a pesquisa aplicada é importante, pois o objeto de estudo, a informação, é frequentemente analisado em seus fluxos, usos e processos dentro de contextos organizacionais e sociais específicos (Le Coadic, 2004).

Le Coadic (2004, p. 5) define a Ciência da Informação como uma disciplina que tem por objeto “as propriedades gerais da informação (natureza, gênese, efeitos) e a elaboração de sistemas de informação [...] para torná-la acessível”. Essa definição intrinsecamente une um polo teórico, o estudo da informação, a um polo prático e tecnológico, a elaboração de sistemas. É precisamente nessa intersecção que a presente pesquisa se situa, caracterizando-se, quanto à sua natureza, como pesquisa aplicada, com foco em um componente tecnológico (Le Coadic, 2004).

Para Le Coadic (2004), a informação registrada funciona como o ponto de partida para o ciclo informacional que envolve produção, coleta, tratamento, organização, armazenamento, recuperação, difusão e uso. Sua pesquisa aborda um ponto de ruptura nesse ciclo: a dissociação entre a difusão, e o armazenamento de longo prazo, que exige uma estratégia de preservação. O problema não é teórico, mas eminentemente prático: como garantir que a informação difundida hoje continue a existir para ser usada amanhã?

A natureza aplicada do estudo se manifesta na busca por uma solução para esse problema concreto. Não se trata de uma reflexão abstrata sobre a

perenidade da informação, mas da “elaboração de sistemas de informação” ou, mais precisamente, de uma arquitetura de integração entre sistemas, para garantir a longevidade do acervo (Le Coadic, 2004). Esta pesquisa visa, portanto, gerar um conhecimento tecnológico e processual diretamente aplicável por instituições que enfrentam esse desafio, alinhando-se à dimensão prática e prescritiva.

A adequação dessa classificação se evidencia ao se analisar a pergunta central do estudo: “Como a adoção de padrões baseados no modelo OAIS pode contribuir para a preservação digital a longo prazo dos acervos museais difundidos por meio do Tainacan?”. A questão não busca apenas compreender teoricamente o modelo OAIS, mas investigar como aplicá-lo para resolver um problema prático e recorrente em instituições museais: a dissociação entre as estratégias de difusão digital e a ausência de uma política de preservação de longo prazo que garanta a perenidade do próprio acervo. O resultado esperado é um modelo de arquitetura e um conjunto de diretrizes práticas que possam ser efetivamente implementados por museus e outras instituições que utilizam o Tainacan, caracterizando o viés aplicativo e intervencionista da investigação.

A abordagem para solucionar o problema proposto é qualitativa, com um viés sistêmico. Le Coadic (2004, p. 89) enfatiza que um sistema de informação não é apenas um amontoado de *hardware* e *software*, mas um “conjunto organizado de recursos [...] para coletar, tratar, armazenar, recuperar e difundir informação”. A compreensão de tal sistema exige uma análise que vá além de métricas quantitativas, focando nas relações entre seus componentes, nos fluxos de informação e na adequação do sistema a seus objetivos.

A presente investigação se alinha à perspectiva de Gil (2008) sobre a relevância da pesquisa qualitativa diante da crescente complexidade e diversificação das esferas sociais. O problema central, que consiste em compreender como o modelo OAIS pode transformar o Tainacan em uma ferramenta de difusão conectada a requisitos de preservação a longo prazo, não se resume a uma questão técnica, mas envolve a análise de processos, práticas e perspectivas dos sujeitos. A pesquisa qualitativa, segundo o autor, é particularmente adequada para estudar o conhecimento e as práticas dos participantes em seus próprios termos. Desse modo, a escolha por essa

abordagem justifica-se por ser a mais apropriada para investigar a lógica subjacente ao fluxo informacional e a viabilidade das conexões propostas, considerando a subjetividade e as interações como parte explícita da construção do conhecimento. Assim, a pesquisa se debruça sobre:

1. A sintaxe da informação: Le Coadic (2004, p. 24) aponta que “a informação é uma sintaxe: um arranjo, uma composição, uma estrutura”. A pesquisa qualitativa permite analisar a estrutura dos pacotes de informação (AIP e DIP), a sintaxe dos padrões de metadados (PREMIS, Dublin Core) e a forma como essa estrutura pode ser transformada na passagem entre o Archivematica e o Tainacan.
2. A mediação tecnológica: A integração proposta é um ato de mediação tecnológica, que visa “facilitar o acesso e o uso da informação” (Le Coadic, 2004, p. 73) não apenas no presente, mas para uma comunidade de usuários no futuro. A análise qualitativa permite interpretar a funcionalidade dos componentes (APIs, protocolos, scripts) como elementos de mediação, avaliando sua eficácia em cumprir o objetivo de garantir um acesso perene.

Ao adotar uma abordagem qualitativa e sistêmica, a pesquisa alinha-se à visão de Le Coadic (2004) e Gil (2008), tratando o ecossistema Archivematica-Tainacan não como duas aplicações a serem conectadas, mas como um sistema de informação integrado com funcionamento em consonância. O objetivo é compreender e modelar as relações internas e os fluxos que o governam, a fim de otimizá-lo para a função de preservação a longo prazo, resolvendo um problema prático para as instituições com acervo informacional que utilizam o Tainacan.

3.2 Etapas de execução da pesquisa

Para responder à pergunta de pesquisa e atingir os objetivos propostos, o percurso metodológico foi organizado em três fases sequenciais e interdependentes. Esse desenho metodológico parte de uma revisão conceitual, avança para a modelagem de uma solução tecnológica e culmina na elaboração

de um modelo de aplicação validado teoricamente. Cada fase, com seus respectivos procedimentos, é detalhada neste capítulo.

A **primeira fase** consistiu na construção do alicerce teórico e técnico da pesquisa. Conforme Le Coadic (2004), a compreensão de um sistema de informação exige o domínio tanto dos conceitos que o regem quanto das tecnologias que o materializam. Assim, essa etapa foi dividida em dois procedimentos complementares:

Procedimento 1 – pesquisa bibliográfica: realizou-se um levantamento sistemático da literatura científica nas áreas de Ciência da Informação, Arquivologia, Museologia e Ciência da Computação. O objetivo dessa revisão é analisar como a adoção de padrões baseados no modelo OAIS pode contribuir para a preservação digital a longo prazo dos acervos museais. Essa etapa deu-se com o objetivo de fornecer uma fundamentação teórica que demonstre a relevância do modelo OAIS no contexto da preservação digital de acervos fora do contexto arquivístico. Além disso, buscou-se esclarecer a relação entre os princípios do OAIS e as práticas de preservação adotadas por instituições que utilizam ferramentas tecnológicas para difundir seus acervos museais.

O escopo desta revisão abrange estudos realizados com foco em iniciativas de preservação digital em museus e instituições detentoras de acervos informacionais que transcendam a forma documental convencionalmente estabelecida pela arquivologia. Foram analisados artigos acadêmicos, relatórios técnicos e estudos de caso que abordaram a implementação de padrões de metadados, estratégias de preservação a longo prazo e a interoperabilidade entre sistemas, baseados no modelo OAIS.

Para garantir uma análise completa do universo da pesquisa, adotou-se um protocolo de revisão sistemática, seguindo diretrizes metodológicas *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA). A revisão sistemática permite sintetizar as evidências científicas e casos correlatos existentes, identificando padrões, desafios e estratégias focados no tema da pesquisa, considerando a aplicabilidade do modelo OAIS

(International Organization for Standardization, 2012a³⁴) como referência metodológica central (Page *et al.*, 2021).

A problematização central da revisão consistiu em diagnosticar o que a literatura trata sobre a aplicação do modelo OAIS na preservação digital de objetos museais.

Para responder a essa questão, foram definidos critérios de inclusão e exclusão. Foram selecionados estudos que tratam da aplicação do OAIS na preservação de objetos museais ou que não fossem tradicionalmente documentos de arquivo considerando abordagens aplicáveis a mídias interativas, audiovisuais, modelagem tridimensional e dados científicos (Arellano, 2008; Gilliland, 2014). Excluíram-se trabalhos que se restringem à preservação de documentos textuais convencionais, abordando exclusivamente formatos clássicos (ex: PDFs).

Os termos de busca foram estruturados para abranger diferentes abordagens da preservação digital, conforme a pergunta de pesquisa: (“digital preservation” OR “preservação digital”) AND (“cultural heritage” OR “objetos culturais”) AND (“OAIS”).

As bases de dados utilizadas para a busca de literatura incluem Scopus, Web of Science, Google Scholar e a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), conforme a recomendação da literatura de Becker (2009). Além dessas fontes, foi utilizada a inteligência artificial SCISPaCE, uma ferramenta especializada em pesquisa acadêmica assistida por IA, que permite analisar grandes volumes de publicações, identificar padrões temáticos e sugerir referências relevantes com base em modelagem semântica. A inclusão da SCISPaCE ampliou a abrangência da revisão, garantindo um levantamento mais eficiente e sistemático dos estudos (Xu; Lin, 2021).

A Scopus e a Web of Science são bases de dados multidisciplinares e indexadas, amplamente utilizadas para revisões sistemáticas devido à sua rigorosa curadoria de publicações científicas. Ambas oferecem filtros avançados para refinamento da pesquisa, permitindo a seleção de artigos de alto impacto e

³⁴ ISO 14721:2012.

relevância na área pesquisada (Becker, 2009). O Google Scholar foi incluído para complementar a busca, pois indexa uma variedade mais ampla de documentos, incluindo *preprints*, teses, dissertações e publicações institucionais que podem não estar disponíveis em bases de dados comerciais (Haddow; Genoni, 2010; Martínez-López *et al.*, 2018). Essa base permite um panorama mais abrangente da produção científica, incluindo estudos emergentes e de acesso aberto.

Além disso, a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) foi selecionada para identificar pesquisas acadêmicas desenvolvidas no Brasil, garantindo a inclusão de estudos nacionais. A plataforma reúne teses e dissertações de diversas universidades brasileiras, muitas das quais trazem contribuições relevantes para o contexto (Santos *et al.*, 2017).

Por fim, a SCISPaCE foi utilizada para otimizar a busca por estudos relevantes. A ferramenta emprega inteligência artificial para mapear tendências, sugerir conexões entre artigos e identificar pesquisas correlatas. Essa abordagem permitiu reduzir vieses de seleção, detectar artigos altamente citados e explorar publicações menos acessíveis por métodos convencionais de busca (Xu; Lin, 2021). Os termos de busca foram estruturados de modo a considerar rigorosamente o tema, garantindo que a recuperação dos trabalhos fosse focada no tema.

Para a seleção dos artigos, adotou-se um processo de triagem em múltiplas etapas. Inicialmente, os trabalhos recuperados foram analisados por meio da leitura de títulos e resumos, verificando se abordavam a aplicação do modelo OAIS na preservação digital de objetos culturais não convencionais ou museológicos. Aqueles que não atenderam a esse critério foram excluídos. Os estudos que passaram por essa primeira fase foram submetidos à leitura integral, na qual foram avaliados quanto à relevância metodológica, aplicabilidade prática e alinhamento com os desafios específicos da preservação digital em contextos museológicos e patrimoniais.

Ademais, os artigos incluídos foram analisados detalhadamente, extraindo-se informações sobre as estratégias de preservação utilizadas, os desafios identificados e as soluções propostas pelos autores. Os trabalhos, que

incluíram artigos, livros e normas técnicas, foram analisados, permitindo uma síntese crítica das abordagens existentes e a identificação de lacunas na literatura, que serviram de base para a formulação de recomendações e diretrizes futuras para a preservação digital de acervos museológicos.

Procedimento 2 - análise documental: procedeu-se à análise de fontes primárias, ou documentos. Esse procedimento focou-se em:

- Normas Técnicas: leitura e interpretação da norma ABNT NBR ISO 14721:2012, que especifica o modelo de referência OAIS;
- Manuais de *Software*: estudo aprofundado da documentação técnica oficial dos *softwares* Tainacan e Archivemática para compreender suas arquiteturas, funcionalidades, modelos de dados e, crucialmente, as especificações de suas APIs REST;
- Padrões de Metadados e Protocolos: análise das especificações dos padrões de metadados PREMIS e *Dublin Core*, e do protocolo de comunicação SWORD, para avaliar suas potencialidades e limitações no contexto da integração.

Portanto, essa fase foi essencial para mapear o território da pesquisa, definindo os contornos do problema e fornecendo o embasamento necessário para a fase de modelagem.

Fase 2: Modelagem da Arquitetura de Integração Sistêmica

Uma vez estabelecida a base conceitual, a segunda fase concentrou-se na elaboração do modelo de aplicação aos sistemas. O objetivo foi projetar uma arquitetura de integração funcional entre o ambiente de preservação Archivemática e o de difusão Tainacan, tratando-os como componentes de um único sistema de informação arquivística e museológica, a fim de garantir a preservação a longo prazo.

Procedimento 2.1 – desenho da arquitetura lógica: utilizando a notação de diagramas de fluxo, foi desenhada a arquitetura lógica do ecossistema. Esse procedimento definiu os papéis de cada *software* em relação

às entidades funcionais do OAIS (Admissão, Armazenamento, Administração de Dados, Planejamento da Preservação e Acesso) e mapeou o percurso do objeto museológico digital através do sistema.

Procedimento 2.2 – análise e seleção do método de integração: com base na análise comparativa desenvolvida no capítulo anterior, justificou-se a escolha do método de integração via *scripts* personalizados e API direta. Essa decisão foi fundamentada na necessidade de flexibilidade para o tratamento de metadados museológicos, que são inerentemente mais complexos e heterogêneos que os esquemas suportados por métodos padronizados como o SWORD.

Procedimento 2.3 – mapeamento de metadados (Crosswalk): elaborou-se um quadro de mapeamento (*crosswalk*) para estabelecer a correspondência entre os metadados de preservação gerados pelo Archivematica (extraídos do padrão PREMIS) e os campos de metadados descritivos e de proveniência no Tainacan. O objetivo foi garantir que informações essenciais para a autenticidade e a cadeia de custódia do objeto digital fossem transferidas e preservadas no ambiente de acesso.

Fase 3: Desenvolvimento e validação do protótipo conceitual

A fase final visou materializar a arquitetura modelada em um protótipo conceitual, que serve como uma prova de conceito e um guia para implementação.

Procedimento 3.1 – especificação do *script* de integração: em vez de desenvolver um *software* funcional, o que estaria fora do escopo de uma pesquisa teórica, elaborou-se um pseudocódigo detalhado para o *script* de integração. Esse pseudocódigo descreve a lógica do processo de ETL (*Extract, Transform, Load*):

Extração (*Extract*): define como o *script* deve monitorar o Archivematica (via API ou sistema de arquivos) para identificar a geração de um novo Pacote de Informação de Arquivamento (AIP) e como extrair dele os objetos digitais e os metadados relevantes.

Transformação (*Transform*): detalha como o *script* deve processar os dados extraídos, aplicando o mapeamento de metadados (*crosswalk*) definido na Fase 2 para converter a informação ao formato esperado pela API do Tainacan.

Carga (*Load*): especifica como o *script* deve interagir com os *endpoints* da API REST do Tainacan para criar um item, popular seus metadados e anexar o objeto digital de acesso.

Procedimento 3.2 – validação teórica do modelo: a validação do protótipo e da arquitetura como um todo foi realizada por meio de uma análise argumentativa. Confrontou-se o modelo proposto com os requisitos mandatórios do modelo OAIS e com os princípios fundamentais da preservação digital (autenticidade, integridade, identidade, proveniência e fixidez). Demonstrou-se, ponto a ponto, como a solução desenhada atende a esses requisitos, respondendo de forma conclusiva à pergunta de pesquisa e confirmando a contribuição do modelo para a preservação a longo prazo dos acervos.

Esse percurso metodológico, partindo do geral (teoria) para o particular (tecnologia), e culminando em um modelo validado, garante a rastreabilidade, a consistência e o rigor científico necessários para a presente investigação.

3.3 Instrumentos de coleta e análise de dados

A execução de cada fase do percurso metodológico exigiu o emprego de instrumentos específicos para a coleta e a análise dos dados. Em uma pesquisa de natureza qualitativa e documental como esta, os dados não são numéricos, mas, sim, o conteúdo textual e estrutural de documentos, normas, manuais e códigos (Le Coadic, 2004). A seleção dos instrumentos foi guiada pela capacidade de adequação para extrair e interpretar a informação necessária em cada etapa.

A coleta de dados, correspondente à Fase 1 da metodologia (Fundamentação teórico-Conceitual), foi realizada por meio de um conjunto de

ferramentas e plataformas digitais que permitiram o acesso ao corpus documental da pesquisa.

1. Bases de dados e repositórios científicos: para a pesquisa bibliográfica, foram utilizados portais de indexação e bases de dados de alta relevância científica, como Scopus, Web of Science, Scielo e Google Scholar. Esses instrumentos foram essenciais para a coleta de artigos, teses e dissertações que compõem o estado da arte sobre preservação digital, OAIS e sistemas de informação para acervos. A utilização de múltiplos portais visou mitigar vieses e garantir uma cobertura abrangente da literatura.
2. Repositórios de documentação técnica (*github* e wikis oficiais): a coleta de dados primários sobre os *softwares* Tainacan e Archivematica foi realizada diretamente em seus repositórios oficiais de documentação. O *Github* e os *wikis* mantidos pelas equipes de desenvolvimento foram os principais instrumentos para acessar manuais de instalação, guias de usuário, e, crucialmente, a documentação detalhada das APIs REST. Esses documentos foram tratados como fontes primárias, fornecendo os dados brutos sobre a arquitetura e as capacidades de cada sistema.
3. Websites de órgãos de normalização: para a coleta de normas e padrões, foram utilizados os websites de órgãos como a Associação Brasileira de Normas Técnicas (ABNT), para acesso à NBR ISO 14721:2012, e da Library of Congress, para acesso aos padrões PREMIS e METS.

A análise dos dados coletados, que perpassa todas as fases da pesquisa, foi realizada com o auxílio de instrumentos que permitem a organização, a interpretação e a modelagem da informação.

1. Fichamento analítico-interpretativo: para a análise da literatura (artigos, livros, teses), o principal instrumento foi o fichamento analítico-interpretativo. Diferente de um simples resumo, esse método envolve a leitura crítica do material, a extração de conceitos-chave, a identificação de argumentos centrais e a anotação de reflexões e conexões com o problema de pesquisa (Marconi; Lakatos, 2017). Esse instrumento foi fundamental para construir o referencial teórico e identificar as lacunas que a pesquisa se propõe a preencher.
2. Análise de Conteúdo: a análise dos documentos técnicos (manuais, normas, especificações de API) foi conduzida sob a ótica da análise de conteúdo, especificamente a análise categorial temática. O procedimento consistiu em:

- 2.1 Pré-análise: leitura flutuante do corpus documental para familiarização.
- 2.2 Exploração do material: codificação do material, identificando unidades de registro (ex: “*endpoint* da API”, “entidade funcional OAIS”, “evento PREMIS”) e agrupando-as em categorias temáticas, como “Requisitos de Autenticação”, “Estrutura de Metadados”, “Funções de Preservação”, etc.
- 2.3 Tratamento dos resultados e interpretação: a partir das categorias, realizou-se a interpretação dos dados, inferindo as capacidades, limitações e a lógica operacional dos sistemas e padrões analisados.
- 2.4 Quadros comparativos e de mapeamento: como instrumento de síntese e visualização da análise, foram utilizados quadros. O Quadro Comparativo foi empregado para contrastar os diferentes métodos de integração (Capítulo 2), enquanto o Quadro de Mapeamento (*Crosswalk*) foi o instrumento central da Fase 2 para correlacionar os campos de metadados entre o Archivematica e o Tainacan. Esses quadros não são apenas formas de apresentação, mas ferramentas analíticas que forçam a sistematização e a tomada de decisão baseada em critérios explícitos.
- 2.5 Diagramas de fluxo de dados: para a modelagem da arquitetura de integração (Fase 2), o principal instrumento foi o diagrama de fluxo de dados, elaborado com a notação padrão de diagramas de processo. Esse instrumento visual permitiu decompor o sistema em seus componentes, mapear as interações e visualizar o percurso da informação, tornando a arquitetura proposta compreensível e passível de análise crítica.

Dessa forma, a utilização conjunta desses instrumentos garantiu que a coleta de dados fosse abrangente e que a análise fosse profunda, sistemática e rastreável, conferindo a robustez necessária ao percurso metodológico desta tese.

3.4 Considerações éticas e uso de ferramentas de Inteligência Artificial (IA)

Reconhece-se que ferramentas baseadas em IA foram empregadas em fases específicas desta pesquisa como instrumentos de auxílio e não como substitutos do raciocínio crítico e da análise do pesquisador. O uso dessas ferramentas foi pautado pelo princípio da autoria e responsabilidade integral,

significando que todo o conteúdo textual, análises e conclusões apresentados nesta tese são de inteira responsabilidade do autor.

Para a execução da revisão bibliográfica sistemática (Fase 1), utilizou-se a plataforma *SciSpace* (anteriormente conhecida como *Typeset*). Essa ferramenta de IA foi empregada para otimizar o processo de busca e triagem inicial da literatura. Suas funcionalidades de busca em linguagem natural e sumarização de artigos foram utilizadas para acelerar a identificação de trabalhos potencialmente relevantes dentro de um vasto universo de publicações. Assim, o uso do *SciSpace* permitiu uma varredura mais ampla e eficiente das bases de dados, aumentando o escopo da revisão. A ferramenta atuou como um assistente de pesquisa na etapa de triagem, mas a decisão final sobre a inclusão ou exclusão de cada artigo, bem como sua leitura aprofundada e análise crítica, foi um processo inteiramente manual e intelectual, realizado pela autora.

É imperativo ressaltar que a ferramenta não foi utilizada para a geração de conteúdo original ou para a formulação das análises e argumentos centrais desta tese. A adoção transparente dessa ferramenta reflete uma prática de pesquisa contemporânea, que integra novas tecnologias de forma ética e responsável para potencializar a capacidade analítica e produtiva do cientista.

4. O TAINACAN COMO OBJETO DE ESTUDO: ESTRUTURA E LIMITES

O Tainacan é uma plataforma de código aberto desenvolvida para a gestão e disseminação de acervos digitais, integrando-se ao *WordPress* para oferecer uma solução flexível e acessível a instituições culturais (Siqueira; Martins, 2014) com a principal função focada em permitir que museus organizem e disponibilizem suas coleções de forma eficaz. Facilitando, portanto, o acesso e a interação com o público (Siqueira; Martins, 2014).

Em uma análise comparativa entre o Tainacan e o Omeka³⁵, Martins, Lemos e Andrade (2021) observam que o Tainacan requer 25% menos esforço tecnológico para implantação e uso, tornando-se uma opção mais viável para instituições com recursos limitados. Além disso, Martins e Martins (2020) discutem os desafios e aprendizados na implementação do Tainacan nos museus do Ibram, ressaltando a importância de estratégias colaborativas e personalizadas para a adoção eficaz da tecnologia.

Ao analisarem essa experiência, os autores destacaram que a transição para plataformas digitais em museus brasileiros não ocorre de forma uniforme, sendo condicionada por fatores, como a infraestrutura tecnológica disponível, a capacitação das equipes técnicas, as particularidades dos acervos e o nível de maturidade digital de cada instituição. Nesse contexto, a aplicação do Tainacan mostrou-se eficaz à medida que permitiu a personalização dos fluxos de trabalho e a adaptação dos metadados às realidades locais, respeitando tanto os princípios museológicos quanto os requisitos da gestão de documentos.

Os autores observaram que a adoção do Tainacan implicou em um processo de aprendizagem organizacional, no qual o conhecimento técnico foi sendo construído coletivamente a partir da troca de experiências entre museus e da mediação realizada por especialistas e desenvolvedores da plataforma. Essa dinâmica gerou um ambiente propício à inovação incremental, no qual os museus foram progressivamente integrando a lógica dos repositórios digitais às suas práticas cotidianas. Logo, a personalização da ferramenta, especialmente no que se refere à estruturação dos metadados e à configuração das taxonomias, foi essencial para garantir que os acervos fossem representados de

³⁵ Omeka é uma plataforma de código aberto voltada à publicação de coleções digitais e exposições online, amplamente utilizada por museus, bibliotecas e arquivos. Desenvolvida pelo *Roy Rosenzweig Center for History and New Media da George Mason University*, a ferramenta combina funcionalidades de gestão de acervos com recursos de exibição pública e suporte à descrição baseada em metadados padronizados, como Dublin Core. Seu modelo modular permite a integração de *plugins* para fins de preservação, interoperabilidade com sistemas de repositórios, e exposição de objetos digitais de forma interativa, sendo particularmente útil em projetos de humanidades digitais e curadoria digital. A facilidade de uso e a possibilidade de customização tornam o Omeka uma solução acessível tanto para instituições pequenas quanto para iniciativas colaborativas de preservação e difusão do patrimônio cultural digital. (Rogers, K.; Ross, 2010)

maneira fiel e interoperável, assegurando também a aderência aos padrões internacionais de descrição, como o Dublin Core³⁶.

Outro aspecto relevante destacado por Martins e Martins (2020) diz respeito à importância das estratégias de engajamento institucional. A implementação bem-sucedida do Tainacan exigiu o envolvimento direto dos gestores, curadores, técnicos e demais profissionais dos museus, reforçando a necessidade de uma governança colaborativa no âmbito da política de preservação digital.

Tal perspectiva está alinhada às discussões mais recentes sobre governança informacional em redes distribuídas, como observado nos modelos híbridos propostos por Zierau e Schultz (2013), que conciliam autonomia institucional e padronização dos processos. Nesse sentido, a experiência do Ibram com o Tainacan serve como um caso exemplar da aplicação de soluções tecnológicas abertas que, ancoradas em metodologias participativas, promovem a sustentabilidade da preservação digital em instituições públicas, ao mesmo tempo em que fortalecem a democratização do acesso aos acervos culturais.

A plataforma já foi adotada por mais de 1.000 (mil) instituições em diferentes regiões do Brasil e do mundo, incluindo museus universitários, secretarias de cultura, bibliotecas públicas, acervos comunitários, arquivos especializados e centros de documentação. Dentre os exemplos notáveis, estão o Museu da República (RJ), o Museu Victor Meirelles (SC), o Arquivo Público do Distrito Federal (DF), o Museu Universitário da Universidade de São Paulo (USP), entre outros. Além disso, o Tainacan tem ganhado projeção internacional, com versões traduzidas para mais de dez idiomas, incluindo espanhol, inglês,

³⁶ O *Dublin Core* é um padrão de metadados desenvolvido pelo *Dublin Core Metadata Initiative* (DCMI), amplamente utilizado para descrever recursos digitais com foco na simplicidade e interoperabilidade. Já o *CIDOC CRM (Conceptual Reference Model)* é uma ontologia formal desenvolvida pelo *International Committee for Documentation* do ICOM (Conselho Internacional de Museus), voltada à representação semântica de informações sobre o patrimônio cultural, promovendo a integração entre sistemas heterogêneos. Para mais informações, ver: (*Dublin Core Metadata Initiative*, 2020)

francês, galego e guarani, possibilitando sua utilização em contextos multilíngues e colaborativos (Tainacan, 2025).

Justifica-se o uso do Tainacan não apenas por sua acessibilidade técnica e financeira, mas também por sua aderência a padrões internacionais de interoperabilidade, como *Dublin Core*, FOAF, SKOS e CIDOC CRM³⁷. A ferramenta permite a modelagem flexível dos metadados, o uso de taxonomias controladas, a exportação em diversos formatos (como CSV, JSON e RDF), e a indexação para mecanismos de busca, promovendo ampla visibilidade dos acervos digitalizados (Martins; Lemos; Andrade, 2021).

Além de ser tecnicamente completo, no que tange à recursos de disseminação de acervos, o Tainacan promove uma abordagem participativa e formativa, ao envolver diretamente as instituições no processo de parametrização da plataforma, incentivando a autonomia e o empoderamento das equipes locais na gestão de seus acervos. Martins *et al.* (2020) ressaltam que a plataforma tem contribuído para o fortalecimento das práticas de documentação, curadoria digital e disseminação do patrimônio cultural, especialmente em museus vinculados ao Ibram, que enfrentam desafios históricos relacionados à padronização da gestão de acervos.

Portanto, o Tainacan representa uma inovação significativa no campo da curadoria digital e da democratização do acesso à informação, ao reunir uma infraestrutura tecnológica alinhada aos princípios da ciência aberta, da inclusão digital e da preservação do patrimônio. Sua adoção crescente por instituições públicas e comunitárias evidencia seu potencial como uma política pública de gestão e difusão de acervos culturais em escala nacional e internacional.

³⁷ O CIDOC CRM (*Conceptual Reference Model*) é uma ontologia formal que constitui uma ferramenta teórica e prática para a integração de informações no campo do patrimônio cultural. Desenvolvido pelo Comitê Internacional de Documentação (CIDOC) do Conselho Internacional de Museus (ICOM), o modelo foi adotado como norma internacional ISO 21127, fornecendo definições e uma estrutura formal para descrever os conceitos e relações implícitos e explícitos utilizados na documentação de patrimônio cultural (Doerr, 2003).

Essas três ontologias são frequentemente utilizadas em conjunto em projetos de dados ligados (*linked data*) e patrimônio cultural digital, sendo FOAF para descrição de pessoas e relações sociais, SKOS para vocabulários controlados e taxonomias, e CIDOC CRM para modelagem conceitual de informações sobre objetos culturais e eventos históricos (Fonte: elaborado pela autora, 2025)

O Tainacan coloca-se como uma significativa ferramenta de política pública voltada à democratização do acesso à informação e à preservação do patrimônio cultural em meio digital. Enquanto política tecnológica aberta e colaborativa, sua concepção e desenvolvimento respondem diretamente às necessidades de instituições públicas brasileiras que, historicamente, enfrentam limitações orçamentárias, técnicas e de pessoal para a adoção de sistemas robustos de gestão de acervos.

Nesse contexto, a proposta do Tainacan é duplamente estratégica: por um lado, fomenta a criação de repositórios digitais interoperáveis e acessíveis, e por outro, fortalece a autonomia institucional e o empoderamento técnico das equipes locais (Martins; Martins, 2020). A escolha por uma plataforma de código aberto permite que o Estado reduza a dependência de soluções proprietárias, contribuindo para a soberania informacional e tecnológica.

Além disso, ao adotar padrões de interoperabilidade internacional e integrar-se ao ecossistema do *WordPress*, o Tainacan oferece uma infraestrutura adaptável, com baixo custo de manutenção e ampla capacidade de difusão dos acervos na internet. Tal arquitetura possibilita sua aplicação em diferentes esferas da administração pública – federal, estadual e municipal – além de atender também a museus universitários e centros culturais comunitários, consolidando-se como uma política transversal e federativa de incentivo à curadoria digital.

Outro aspecto que qualifica o Tainacan como instrumento de política pública é sua integração com as diretrizes do Sistema Brasileiro de Museus e com os princípios da Política Nacional de Informação e Informática em Saúde (PNIIS) e da Política Nacional de Cultura Digital. O suporte institucional contínuo do Ibict e da Universidade Federal de Goiás (UFG), além de sua adoção por organismos vinculados ao Ibram, evidencia o alinhamento entre o desenvolvimento da ferramenta e os marcos regulatórios que orientam a atuação pública na gestão do patrimônio cultural (Martins; Martins, 2020).

Além disso, a aplicação promove valores essenciais ao serviço público, como transparência, inclusão e participação social. Ao permitir que acervos

antes inacessíveis fisicamente sejam amplamente consultados por meio digital, a plataforma amplia o direito à memória e ao patrimônio, em consonância com o princípio constitucional da universalização do acesso à cultura e à informação (art. 5º e 215 da Constituição Federal de 1988). A gestão aberta do projeto, com repositório no GitHub, fóruns de usuários e atualizações contínuas, demonstra seu compromisso com a ciência aberta, à inovação pública e a governança participativa (Martins; Martins, 2020).

O Ibram é uma autarquia federal vinculada ao Ministério da Cultura, criada por meio da Medida Provisória nº 398, de 2009, convertida na Lei nº 11.906/2009, com a finalidade de formular e implementar políticas públicas voltadas à área museológica no Brasil. Suas atribuições englobam a preservação, a valorização e a difusão do patrimônio museológico nacional, incluindo a gestão direta de museus federais e o desenvolvimento de normativas, programas e ações que promovam a articulação e o fortalecimento do setor museológico em âmbito nacional. Nesse contexto institucional, a atuação do Ibram é essencial para a consolidação de políticas estruturantes que assegurem o acesso democrático à cultura.

Entre essas estratégias, destaca-se o projeto “Acervos digitais dos museus do Ibram”, realizado em parceria com a Universidade Federal do Espírito Santo (UFES) e com apoio da Fundação Espírito-santense de Tecnologia. Esse projeto visa à digitalização, documentação, organização e disponibilização dos acervos dos museus sob a guarda do Ibram, utilizando a plataforma Tainacan como repositório digital. A iniciativa alinha-se diretamente à missão do Instituto de democratizar o acesso ao patrimônio cultural musealizado, promover a interoperabilidade entre repositórios e fomentar a adoção de padrões técnicos e normativos, como o Inventário Nacional dos Bens Culturais Musealizados (INBCM), Resolução Normativa nº 5, de 3 de agosto de 2021 do próprio Ibram.

A publicação “Planejamento e política de acervos digitais: desafios e possibilidades” (Silva, 2022) contribui à compreensão desse cenário ao discutir a trajetória de formulação das políticas públicas voltadas à digitalização e difusão dos acervos museológicos no Brasil. A autora evidencia que, desde a criação do Programa acervo em rede pelo Ibram, em 2013, vem se consolidando uma

agenda institucional que reconhece a digitalização como instrumento fundamental de acesso e preservação. Tal programa teve como uma de suas ações centrais o fomento à implementação do Tainacan, *software* desenvolvido pela Universidade Federal de Goiás (UFG), com o objetivo de atender às especificidades da gestão e disseminação de acervos museológicos, consolidando a cooperação interinstitucional entre o Ibram e universidades públicas brasileiras.

A autora recupera documentos fundamentais, como a Carta de Rio Grande (2002), a Política Nacional de Museus (2003) e o Plano Nacional Setorial de Museus (2010), demonstrando como a digitalização e a informatização de acervos já eram demandas reconhecidas desde o início dos anos 2000. No entanto, embora formalmente previstas, essas diretrizes não se materializaram em ações estruturadas e sustentáveis, resultando em iniciativas pontuais e desarticuladas. Diante desse panorama, a autora aponta que o avanço da digitalização no campo museológico depende não apenas de infraestrutura, mas de vontade política, planejamento estratégico e diretrizes normativas consolidadas.

É nesse cenário que se insere o Programa acervo em rede, criado pelo Ibram, em 2013, com o objetivo de democratizar o acesso aos bens culturais musealizados, promover sua documentação digital e instrumentalizar os museus com ferramentas tecnológicas adequadas à gestão e difusão dos acervos. A obra destaca esse programa como um marco institucional importante, pois inaugura uma nova fase no processo de digitalização museológica, conectando o poder público federal a universidades e centros de pesquisa.

Um dos produtos mais notáveis desse esforço foi a parceria estabelecida entre o Ibram e a UFG, que resultou no desenvolvimento da plataforma Tainacan — um *software* livre e de código aberto voltado à organização e à disponibilização *online* de acervos culturais.

A plataforma Tainacan, segundo a autora, surge como uma solução técnica e política: técnica, por oferecer funcionalidades específicas de metadados, interoperabilidade e curadoria digital, e política, por ser fruto de um

projeto acadêmico desenvolvido com recursos públicos e em diálogo com as demandas do campo museológico nacional. A publicação ainda destaca o potencial do Tainacan para padronizar a documentação museológica, por meio da adoção do Inventário Nacional dos Bens Culturais Musealizados (INBCM), que estabelece elementos obrigatórios para descrição dos objetos, conforme a Resolução Normativa nº 5, de 3 de agosto de 2021 do Ibram. Essa padronização é vista como essencial para garantir a interoperabilidade entre os acervos e permitir a construção de uma rede nacional de repositórios culturais.

Outro ponto relevante desenvolvido por Silva (2022) é a constatação de que, apesar dessas iniciativas federais, a maior parte dos museus brasileiros, especialmente os universitários, ainda enfrenta dificuldades estruturais para aderir a projetos de digitalização. A autora destaca dados preliminares de pesquisa coordenada por Martins e Martins (2020), os quais revelam que menos de 30% das coleções universitárias publicam algum conteúdo digital, e apenas 7% utilizam repositórios. Essa realidade é atribuída à ausência de políticas institucionais, à carência de profissionais especializados e à falta de investimento continuado em infraestrutura tecnológica e capacitação.

A experiência da UFRGS com o projeto Gestão de Acervos Museológicos, por sua vez, é apresentada como uma proposta de metodologia replicável, fundamentada no uso do Tainacan, na aplicação de vocabulário controlado e em procedimentos de curadoria e documentação baseados em padrões reconhecidos. O projeto teve como foco inicial o Acervo do Laboratório de Ensino de Física (AMLEF) e a Pinacoteca Barão de Santo Ângelo (PBSA), promovendo a organização, a digitalização e a difusão dos itens dessas coleções com o apoio de docentes e discentes do curso de Museologia.

A autora conclui que os avanços alcançados até agora precisam ser acompanhados de uma política nacional de acervos digitais, integrada aos marcos legais já existentes, que viabilize a articulação entre instituições culturais, universidades e o Estado. A digitalização, nesse sentido, é compreendida não como um fim em si, mas como um instrumento de efetivação do direito de acesso à cultura, à memória e ao patrimônio nacional, conforme prescrito no artigo 215 da Constituição Federal de 1988.

O projeto com a UFES representa não apenas a continuidade dessa política de digitalização, mas um avanço em sua operacionalização, ao propor soluções técnicas e metodológicas voltadas à sistematização da documentação museológica, à capacitação técnica e à estruturação de redes de acervos interoperáveis. A experiência da UFRGS relatada no artigo serve como modelo análogo que reforça a viabilidade dessas iniciativas, sobretudo em contextos institucionais comprometidos com o acesso à informação e com a preservação do patrimônio cultural em consonância com as diretrizes do Ibram.

Portanto, o Tainacan deve ser compreendido não apenas como uma solução técnica, mas como uma política pública estratégica de infraestrutura digital para instituições de memória. Sua concepção centrada na acessibilidade, interoperabilidade e descentralização o posiciona como um elemento-chave para a consolidação de ecossistemas digitais voltados à preservação e à difusão do patrimônio cultural brasileiro.

4.1 Descrição técnica do Tainacan

Um dos componentes tecnológicos mais relevantes do Tainacan é sua API baseada na arquitetura *Representational State Transfer* (REST). Essa API é integral ao funcionamento da plataforma, servindo não apenas como a ponte de comunicação entre o painel administrativo — desenvolvido como uma Aplicação de Página Única (SPA) em Vue.js³⁸ — e o *backend*, mas também como um mecanismo para a interoperabilidade de dados (Tainacan, 2023). A API do Tainacan oferece métodos tanto para leitura (*read*) quanto para escrita (*write*), o que permite que aplicações externas interajam de forma programática

³⁸ Vue.js é um *framework* de código aberto da linguagem JavaScript, focado na construção de interfaces de usuário interativas e aplicações de página única (*Single-Page Applications*). Sua arquitetura baseada em componentes permite o desenvolvimento de interfaces complexas de forma organizada e reativa, onde a visualização dos dados é automaticamente atualizada quando o estado subjacente da aplicação é modificado (You, 2024). No Tainacan, ele é a tecnologia central da interface administrativa.

com o repositório, consultando, criando, atualizando e excluindo coleções, itens e metadados (Tainacan, 2023).

A interoperabilidade, definida como a capacidade de diferentes sistemas de informação se comunicarem e trocarem dados de forma eficaz, é um objetivo central do projeto Tainacan (Martins; Siqueira, 2014). A API RESTful materializa esse objetivo ao expor os dados do acervo em formatos padronizados e amplamente utilizados na web, como JSON (*JavaScript Object Notation*). Essa exposição estruturada dos dados permite que outras instituições, pesquisadores e desenvolvedores possam consumir o acervo para finalidades diversas, como a criação de visualizações de dados personalizadas, a integração com outros repositórios ou a alimentação de portais agregadores de conteúdo cultural (Tainacan, 2023).

Um aspecto fundamental para a interoperabilidade semântica é a capacidade de mapear os esquemas de metadados locais para padrões consolidados. A API do Tainacan, em conjunto com suas funcionalidades de exportação, oferece suporte nativo ao mapeamento para o padrão Dublin Core (DC), um conjunto de elementos de metadados amplamente reconhecido para a descrição de recursos informacionais (Siqueira; Martins; Medeiros, 2021).]

Além do DC, a ferramenta permite a criação de mapeadores para outros padrões, como o INBCM, demonstrando sua flexibilidade para atender a diferentes contextos institucionais e de domínio (Tainacan, 2023). Essa funcionalidade é crucial para garantir que os dados não apenas possam ser trocados, mas também compreendidos por outros sistemas, um desafio persistente na gestão de acervos digitais (Siqueira; Martins; Medeiros, 2021).

Complementando a capacidade de interconexão via API, o Tainacan oferece um conjunto de ferramentas para importação e exportação de dados em massa, um requisito prático para a gestão de grandes volumes de informação (Tainacan, 2023). Essas funcionalidades são essenciais para a migração de acervos de outros sistemas, para a atualização em lote de registros e para a preservação digital dos metadados.

A principal via para a importação de dados é o uso de arquivos no formato *Comma-Separated Values* (CSV). O processo de importação é projetado para

ser flexível, permitindo ao usuário mapear as colunas de sua planilha para os metadados configurados em uma coleção específica do Tainacan. Durante o processo, é possível criar metadados ou novos termos em uma taxonomia (vocabulário controlado) diretamente a partir dos dados da planilha, conferindo agilidade ao processo de catalogação. Além da importação de metadados descritivos, o sistema permite associar os itens aos seus respectivos arquivos (objetos digitais), automatizando a composição do acervo (Tainacan, 2023).

A exportação de dados também é uma funcionalidade central, permitindo que coleções inteiras ou conjuntos de itens filtrados sejam extraídos em formatos, como CSV, XLSX e JSON (Tainacan, 2023). Conforme mencionado, durante a exportação, é possível aplicar mapeadores de metadados, como o Dublin Core, gerando arquivos que já seguem um padrão de descrição internacional, prontos para serem compartilhados ou integrados a outras plataformas (Siqueira; Martins; Medeiros, 2021).

Esse ciclo de exportação e importação viabiliza, por exemplo, a edição de metadados em massa utilizando *softwares* de planilha, onde o usuário pode exportar uma coleção, realizar alterações e, em seguida, importar o arquivo modificado para atualizar os itens no repositório, utilizando um campo especial (*special_item_id*) para garantir a correta identificação dos registros (Luna, 2024; Tainacan, 2024).

Em suma, a combinação de uma API RESTful completa com funcionalidades versáteis de importação e exportação posiciona o Tainacan como uma ferramenta alinhada às melhores práticas de gestão de dados abertos e interoperabilidade, facilitando não apenas a publicação de acervos na web, mas também sua efetiva integração ao ecossistema de dados museais conectados (Tainacan, 2023).

Essa constatação dialoga diretamente com o cerne da presente tese. A arquitetura analisada responde ao primeiro objetivo específico, que visa identificar os componentes que viabilizam a interoperabilidade. A API RESTful, ao expor os dados de forma programática e padronizada, e os exportadores, ao permitirem o mapeamento para esquemas, como o *Dublin Core*, emergem como os pilares técnicos que materializam essa capacidade.

Adicionalmente, a existência de uma API de leitura e escrita, é o que permite que uma instituição integre seu acervo a portais agregadores ou crie visualizações de dados, atendendo ao segundo objetivo específico de avaliar a integração do Tainacan em redes de informação. Da mesma forma, as funcionalidades de importação via CSV são cruciais para o terceiro objetivo específico, pois reduzem a barreira técnica para a digitalização e publicação de acervos, permitindo que instituições com recursos limitados migrem seus catálogos, muitas vezes, mantidos em planilhas simples, para um ambiente *web* robusto e padronizado. Portanto, a análise técnica é a ponte para a comprovação de que a estrutura do Tainacan oferece os meios práticos para a autonomia e a inserção qualificada das instituições que o utilizam a aplicarem requisitos de preservação digital aos seus acervos.

4.2 Análise da estrutura sociotécnica e aplicação dos requisitos TRAC/OAIS ao Tainacan.

A confiabilidade de um repositório digital não reside apenas em sua arquitetura de software, mas em uma complexa interação entre a tecnologia, as pessoas que a mantêm e as políticas institucionais que a governam. Essa interação é conceituada como uma estrutura sociotécnica, um sistema no qual as propriedades sociais e técnicas são interdependentes e se moldam mutuamente (Bijker, 1995). Nesta seção, a análise do Tainacan transcende a sua descrição funcional para avaliar sua estrutura sociotécnica à luz dos critérios estabelecidos pelo documento *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC), referência para a auditoria de repositórios digitais confiáveis baseado no modelo OAIS (RLG-NARA, 2007).

O TRAC organiza seus critérios em três grandes áreas: A) Infraestrutura Organizacional, B) Gestão de Objetos Digitais e C) Tecnologias, Infraestrutura Técnica e Segurança. Embora o TRAC se aplique à instituição que implementa o repositório, é fundamental investigar em que medida o ecossistema do projeto Tainacan, enquanto *software* de código aberto para difusão, mantido por uma

comunidade e com uma governança, provê os subsídios para que uma instituição possa atender a esses requisitos de confiabilidade.

Como descrito no capítulo “2.1 Preservação digital: conceitos fundamentais e desafios”, o ciclo de vida dos documentos se inicia em um ambiente de gestão, perpassando por diversos cenários até a sua disseminação, isto é, nem todos os requisitos do TRAC são referentes ao uso do Tainacan. Nesta seção, serão discutidos os requisitos exclusivos ou relacionados ao *software* de difusão.

A primeira seção do TRAC (A) aborda a viabilidade organizacional, a sustentabilidade financeira e a transparência do repositório. Uma instituição que adota um *software* proprietário depende inteiramente da saúde financeira e das decisões estratégicas de uma única empresa. Em contraste, ao adotar uma solução de código aberto como o Tainacan, a instituição se insere em um ecossistema sociotécnico mais amplo.

A governança do projeto Tainacan, centrada no Laboratório de Inteligência em Redes e Conhecimento e financiada por editais públicos e parcerias institucionais, como a com o Ibram, oferece um modelo de sustentabilidade distribuído (Tainacan, 2024). Isso responde parcialmente a requisitos, como o A2 (Viabilidade e sustentabilidade organizacional), pois o futuro do *software* não depende de um único modelo de negócio comercial, mas do contínuo fomento à pesquisa e extensão em universidades públicas e do interesse da comunidade.

Ademais, a natureza pública e aberta do projeto promove a transparência, um pilar do TRAC. Requisitos, como A5 (Transparência e responsabilização) e A6 (Políticas e procedimentos), são facilitados pela documentação aberta do Tainacan, pelo código-fonte acessível em plataformas, como o *GitHub*, e pelos fóruns de discussão públicos. Uma instituição pode não apenas usar o *software*, mas auditar seu código, compreender suas regras de funcionamento e participar de sua evolução, estabelecendo um nível de transparência que seria impossível com uma solução de proprietária. A comunidade de usuários e desenvolvedores atua, portanto, como uma camada adicional de suporte e vigilância, fortalecendo a infraestrutura organizacional da instituição que adota a ferramenta.

Abaixo, estão listados os requisitos TRAC da sessão A - Infraestrutura Organizacional - bem como a aderência ou a viabilidade que o Tainacan proporciona para o cumprimento desses requisitos para as instituições que o utilizam.

Quadro 8 - Análise de Aderência do Ecossistema Tainacan aos Requisitos da Seção A do TRAC.

ID	Requisito TRAC (Versão Resumida pela autora)	Análise: O Tainacan Atende ou Facilita o Atendimento?
A1	Missão e Escopo: o repositório deve ter uma missão clara e escopo definido para o conteúdo e a comunidade-alvo.	Não atende diretamente, mas viabiliza. O <i>software</i> Tainacan é um sistema de propósito geral. A definição da missão é uma responsabilidade exclusiva da instituição. No entanto, o Tainacan facilita a execução dessa missão ao permitir a criação de múltiplas coleções, cada uma com sua própria política e descrição, e ao oferecer ferramentas para expor essa missão publicamente através de páginas institucionais no <i>WordPress</i> .
A2	Viabilidade e Sustentabilidade Organizacional: o repositório deve ter um plano de negócios ou sustentabilidade.	Não atende diretamente, mas oferece um modelo alternativo. O Tainacan, como <i>software</i> , não tem um plano de negócios para a instituição. Contudo, sua natureza de código aberto e governança pública oferece uma alternativa ao modelo de dependência de um fornecedor comercial. A sustentabilidade da instituição passa a depender menos de licenças de <i>software</i> e mais de sua capacidade de manter a infraestrutura de TI e, opcionalmente, de contratar desenvolvedores para customizações, o que pode ser um modelo mais resiliente para instituições públicas.
A3	Governança e Estrutura Organizacional: o repositório deve ter uma estrutura de governança clara e adequada.	Não se aplica. Este requisito é puramente institucional. A instituição que adota o Tainacan deve definir sua própria governança (comitês, responsáveis, etc.). A estrutura de governança do projeto Tainacan serve como um modelo e um ponto de contato, mas não substitui a governança interna da instituição usuária.
A4	Ambiente Legal e Regulatório: o repositório deve operar dentro de um ambiente legal e regulatório apropriado.	Não atende diretamente, mas viabiliza ferramentas de apoio. O cumprimento de leis (ex: direitos autorais, LGPD) é responsabilidade da instituição. O Tainacan apoia este cumprimento ao possuir um sistema de metadados flexível que pode acomodar campos como "Direitos de Uso", "Licença" e ao permitir a configuração de itens como "privados", restringindo o acesso público a conteúdo sensível ou sob embargo.
A5	Transparência e Responsabilização: o repositório deve ser transparente e responsável perante seus depositantes e usuários.	Sim. A natureza de código aberto do Tainacan é um pilar de transparência. O código-fonte é auditável, a documentação é pública e as discussões de desenvolvimento são abertas (<i>GitHub</i> , Fórum). Para a instituição, isso significa que não há "caixas-pretas". Ela pode justificar suas operações com base em um <i>software</i> cujo funcionamento interno é verificável, aumentando sua própria transparência e responsabilização.

A6	Políticas e Procedimentos: o repositório deve ter políticas e procedimentos documentados.	Não atende diretamente, mas oferece um ponto de partida. O Tainacan não vem com um conjunto de políticas prontas, pois essas são específicas de cada instituição. No entanto, a documentação oficial do Tainacan (<i>wiki</i> , tutoriais) descreve os procedimentos técnicos de uso da ferramenta (como importar, como criar metadados, etc.), servindo como uma base sólida sobre a qual a instituição pode construir seus próprios manuais de políticas e procedimentos de catalogação e gestão.
A7	Divulgação e Acesso a Informações-Chave: o repositório deve se comprometer a divulgar informações sobre seus status e operações.	Sim. O Tainacan, por ser uma extensão do <i>WordPress</i> , permite que a instituição crie páginas e posts para divulgar relatórios, estatísticas de acesso, políticas de preservação e outras informações-chave exigidas pelo TRAC, utilizando a mesma plataforma onde o acervo está publicado.
A8	Sucessão e Continuidade de Acesso: o repositório deve ter um plano para o caso de seu encerramento.	Sim. Este é um dos pontos mais fortes do modelo de código aberto. Um plano de sucessão se torna muito mais viável. Se a instituição que mantém o repositório fechar, os dados não estão presos a um formato proprietário. Eles podem ser exportados em formatos abertos (CSV, JSON), e o próprio <i>software</i> pode ser mantido por outra entidade ou pela comunidade. A ausência de <i>vendor lock-in</i> (dependência de um fornecedor) é um facilitador crucial para um plano de sucessão robusto.

Fonte: Elaborado pela autora com base no Track (2026).

O Tainacan, enquanto *software*, não atende diretamente a nenhum requisito da Seção A do TRAC, pois todos eles se referem a responsabilidades da organização que implementa o repositório.

Contudo, a análise revela que a escolha pelo Tainacan influencia positivamente e facilita o cumprimento da maioria dos requisitos. Sua natureza de código aberto, a governança pública, a documentação acessível e as funcionalidades de importação/exportação em formatos abertos fornecem à instituição um conjunto de ferramentas e um modelo de operação que são intrinsecamente mais transparentes, flexíveis e resilientes do que os oferecidos por sistemas proprietários. O ecossistema Tainacan oferece subsídios fundamentais para os requisitos de Transparência (A5) e Planejamento de Sucessão (A8), que são dois dos pilares para garantir a confiabilidade a longo prazo.

A seção B do TRAC detalha os processos para a gestão do ciclo de vida dos objetos digitais, desde a aquisição até o acesso. É aqui que as funcionalidades técnicas do Tainacan oferecem o suporte mais direto para o

cumprimento dos requisitos (TRAC, 2007). O requisito B2 (Controle do Pacote de Informação de Submissão - SIP), por exemplo, é diretamente atendido pelas funcionalidades de importação do Tainacan. Conforme detalhado na seção 3.3, a importação via CSV permite um controle granular sobre os metadados que são inseridos junto com os objetos, permitindo a criação de um SIP estruturado.

Para o requisito B4 (Criação do Pacote de Informação de Arquivamento - AIP), o Tainacan, por si só, não é uma ferramenta de preservação digital e, portanto, não gera AIPs no sentido estrito do modelo OAIS. Contudo, sua arquitetura é projetada para a interoperabilidade com sistemas que o fazem.

Como demonstrado no Capítulo 4.1 desta tese, a capacidade de exportar metadados e objetos de forma estruturada é o requisito fundamental para que um sistema de preservação, como o Archivematica, possa criar um AIP válido. O Tainacan, portanto, cumpre seu papel na cadeia de custódia da informação ao preparar os dados para o arquivamento.

Os requisitos B5 (Estratégias de preservação) e B6 (Integridade e autenticidade dos objetos digitais) são parcialmente contemplados. O Tainacan garante a integridade dos metadados dentro de seu banco de dados relacional. A integridade do objeto digital em si (o arquivo) é mantida no nível do sistema de arquivos do servidor. Para uma estratégia de preservação robusta, que inclua verificação de *checksums* e monitoramento de obsolescência de formatos, a integração com um sistema como o Archivematica, conforme discutido, torna-se imperativa. O Tainacan funciona como a camada de acesso, enquanto o sistema de preservação atua como a camada de garantia a longo prazo.

O requisito B7 (Controle do Pacote de Informação de Disseminação - DIP) é a especialidade do Tainacan. Toda a sua arquitetura de temas, blocos do *WordPress* e, crucialmente, a API RESTful, são mecanismos para a criação e disseminação de DIPs, permitindo que os dados sejam acessados, consultados e reutilizados pelo público-alvo.

A terceira e última seção do TRAC (C) avalia a adequação e a segurança da infraestrutura técnica que suporta o repositório. Novamente, embora a responsabilidade final pela configuração e manutenção do *hardware* e do

software base seja da instituição, a arquitetura do Tainacan e sua inserção no ecossistema têm implicações indiretas no cumprimento desses requisitos.

O requisito C1 (Identificação e Monitoramento do Sistema) exige que o repositório identifique e monitore continuamente as tecnologias das quais depende. Ao adotar o Tainacan, a instituição assume a responsabilidade de monitorar um conjunto de tecnologias bem definidas: o servidor web, o PHP³⁹, o banco de dados (MySQL/MariaDB) e o próprio *WordPress*, além do Vue.js⁴⁰ no lado do cliente. A grande vantagem é que essa pilha tecnológica (conhecida como “*stack*”) é uma das mais populares e utilizadas na web. Isso resulta na existência de uma vasta gama de ferramentas de monitoramento de código aberto e comerciais, e de uma ampla base de conhecimento na comunidade de administradores de sistemas, o que facilita enormemente o cumprimento desse requisito pela instituição (Song, 2018).

Para o requisito C2 (Mudanças e Atualizações no Sistema), o ecossistema Tainacan oferece um modelo estruturado. O projeto Tainacan possui um ciclo de desenvolvimento público, com versões e notas de lançamento (*changelogs*) disponíveis em seu site e no repositório GitHub. Isso permite que a instituição planeje as atualizações de forma controlada. A responsabilidade de testar e aplicar as atualizações em seu próprio ambiente (primeiro em um ambiente de homologação, depois em produção) é da instituição, mas o processo é facilitado pela previsibilidade e transparência do ciclo de desenvolvimento do Tainacan e do próprio *WordPress*, que possui um robusto sistema de atualizações (Tainacan, 2025; Martins *et al.* 2020)

O requisito C3 (Segurança) aborda a necessidade de proteger o sistema e os dados contra acesso não autorizado, perda ou danos. A segurança é uma responsabilidade compartilhada. A instituição é responsável pela segurança da

³⁹ PHP (acrônimo recursivo para *PHP: Hypertext Preprocessor*) é uma linguagem de *script* de propósito geral, de código aberto, amplamente utilizada para o desenvolvimento web do lado do servidor (*server-side*). Criada originalmente por Rasmus Lerdorf, em 1994, a linguagem permite a criação de conteúdo dinâmico e a interação com bancos de dados, sendo embutida diretamente no código HTML (Lerdorf; Tatro, 2002).

⁴⁰ Vue.js é um *framework* JavaScript de código aberto para a construção de interfaces de usuário, criado por Evan You. Ele é descrito como um “*framework* progressivo”, pois foi projetado para ser adotado de forma incremental, permitindo que seja utilizado tanto para adicionar interatividade a páginas existentes quanto para construir aplicações de página única (*Single-Page Applications*) complexas. Sua arquitetura é baseada em componentes reutilizáveis e utiliza um DOM virtual para otimizar a performance da renderização (VUE.JS, 2024).

infraestrutura (servidor, rede, *firewalls*). O Tainacan, por sua vez, contribui para a segurança no nível da aplicação. Primeiramente, ao ser construído sobre o *WordPress*, ele herda a arquitetura de segurança de um sistema que é constantemente auditado por uma comunidade global de segurança. Tanto os responsáveis pelo *WordPress* quanto os do Tainacan trabalham para corrigir vulnerabilidades e seguir boas práticas de programação, como a sanitização de entradas de dados para prevenir ataques de injeção de SQL⁴¹ e XSS⁴². Em segundo lugar, o Tainacan implementa seu próprio sistema de controle de acesso baseado em papéis e capacidades, permitindo que a instituição defina com precisão quem pode ver, editar, criar ou excluir conteúdo, um aspecto fundamental da segurança da informação. Portanto, o Tainacan atende parcialmente ao requisito, cumprindo sua parte na camada de aplicação (Tainacan, 2023; Martins *et al.* 2020).

Em síntese, o Tainacan insere a instituição em um ecossistema tecnológico bem documentado. Isso não elimina as responsabilidades da instituição em gerenciar sua infraestrutura técnica, mas oferece uma base sólida e um conjunto de ferramentas e conhecimentos comunitários que facilitam o monitoramento, a atualização e a proteção do repositório digital, alinhando-se às boas práticas preconizadas pela Seção C do TRAC. Abaixo, o quadro elucida os requisitos TRAC e a responsabilidade do Tainacan na aplicação.

Quadro 9 - Síntese de Aderência do Ecossistema Tainacan aos Requisitos TRAC.

Seção	ID	Requisito TRAC (versão resumida pela autora)	Responsabilidade do Tainacan acerca dos requisitos TRAC	Justificativa
-------	----	--	--	---------------

⁴¹ SQL (Structured Query Language) é a linguagem padrão para acesso e manipulação de bancos de dados relacionais. Desenvolvida originalmente na IBM por Donald D. Chamberlin e Raymond F. Boyce em 1974, a linguagem permite a definição, consulta e modificação de dados. Foi padronizada pela ANSI, em 1986, e pela ISO, em 1987, tornando-se a base para a maioria dos sistemas de gerenciamento de banco de dados modernos (Chamberlin, 2012).

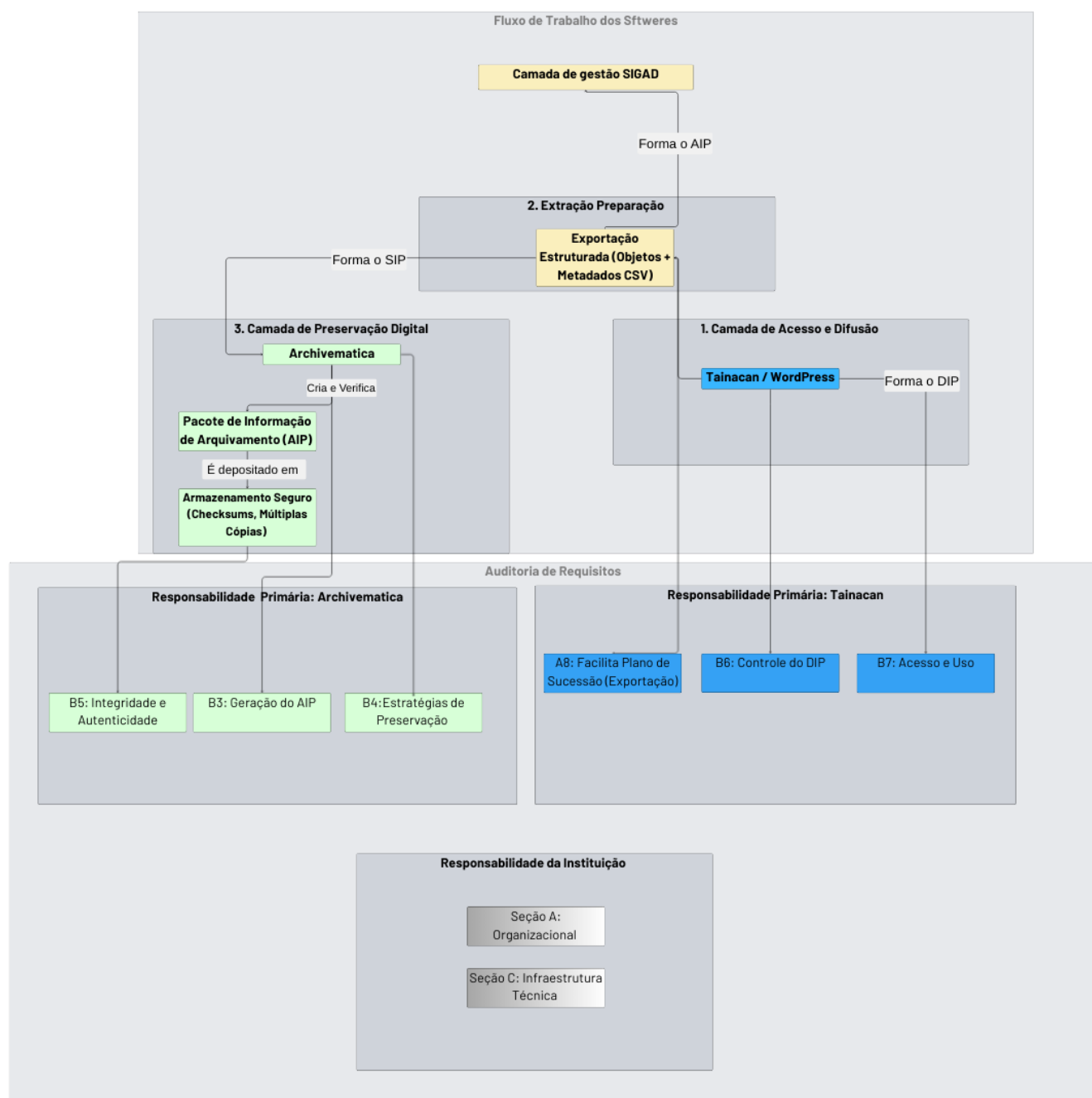
⁴² XSS (Cross-Site Scripting) é uma vulnerabilidade de segurança em aplicações web que permite a um atacante injetar scripts maliciosos (geralmente JavaScript) em páginas visualizadas por outros usuários. Quando a vítima acessa a página comprometida, o script é executado em seu navegador, podendo levar ao roubo de sessões, modificação de conteúdo ou redirecionamento para sites maliciosos. O XSS é uma das falhas de segurança mais prevalentes, sendo consistentemente listado no Top 10 de vulnerabilidades do Open Web Application Security Project (OWASP) (OWASP, 2024).

A	A1	Missão e Escopo	Não	Responsabilidade exclusiva da instituição. O <i>software</i> é uma ferramenta para executar a missão, não para defini-la.
A	A2	Viabilidade e Sustentabilidade Organizacional	Não	Responsabilidade da instituição. O Tainacan oferece um modelo de sustentabilidade alternativo (sem licenças), mas não o plano em si.
A	A3	Governança e Estrutura Organizacional	Não	Responsabilidade exclusiva da instituição.
A	A4	Ambiente Legal e Regulatório	Parcialmente	A instituição é a responsável legal, mas o Tainacan oferece ferramentas (metadados de direitos, acesso restrito) que apoiam o cumprimento.
A	A5	Transparência e Responsabilização	Parcialmente	A instituição é a responsável, mas a natureza de código aberto e documentação pública do Tainacan são facilitadores cruciais para a transparência.
A	A6	Políticas e Procedimentos	Parcialmente	A instituição define as políticas, mas a documentação do Tainacan serve como base para os procedimentos técnicos.
A	A7	Divulgação e Acesso a Informações-Chave	Parcialmente	A instituição deve divulgar, mas o Tainacan, como parte do <i>WordPress</i> , facilita a criação de páginas para essa divulgação.
A	A8	Sucessão e Continuidade de Acesso	Parcialmente	A instituição cria o plano, mas o Tainacan facilita enormemente por usar formatos abertos e não ter dependência de fornecedor (<i>vendor lock-in</i>).
B	B1	Ingestão: Aquisição de Conteúdo	Parcialmente	A instituição assume a custódia, e o Tainacan fornece os mecanismos técnicos (upload, importação) para efetivar a ingestão.
B	B2	Ingestão: Controle do Pacote de Submissão (SIP)	Parcialmente	Controlam os metadados do SIP via importador, mas não realiza verificação de integridade de bits do arquivo (função de um sistema de preservação).
B	B3	Ingestão: Geração do Pacote de Arquivamento (AIP)	Não	Não é sua função. O Tainacan é projetado para <i>enviar</i> os dados para um sistema de preservação (como o Archivematica) que irá gerar o AIP.
B	B4	Estratégias de Preservação	Não	Não executa ações de preservação (ex: migração de formato). Sua função é ser a

				camada de acesso, integrada à camada de preservação.
B	B5	Integridade e Autenticidade dos Objetos Digitais	Parcialmente	Garante a integridade dos metadados em seu banco de dados, mas a verificação contínua da integridade dos arquivos depende de um sistema de preservação.
B	B6	Controle do Pacote de Disseminação (DIP)	Sim	Esta é a sua função principal. A API, os temas e os blocos são todos mecanismos para criar e disseminar DIPs de forma robusta.
B	B7	Acesso e Uso	Sim	Esta é a sua função principal. Oferece busca facetada, filtros, e controle de acesso para que os usuários descubram e utilizem o acervo.
C	C1	Identificação e Monitoramento do Sistema	Parcialmente	A instituição monitora, mas o Tainacan, por ser baseado em PHP/ <i>WordPress</i> , se beneficia de um ecossistema de ferramentas de monitoramento.
C	C2	Mudanças e Atualizações no Sistema	Parcialmente	O projeto Tainacan gerencia as atualizações do <i>software</i> , mas a instituição é responsável por aplicá-las e gerenciar as mudanças em seu ambiente.
C	C3	Segurança	Parcialmente	A segurança geral é responsabilidade da instituição (servidor, rede). O Tainacan contribui com a segurança no nível da aplicação (controle de acesso, sanitização de dados), herdando também as boas práticas do <i>WordPress</i> .

Fonte: Elaborado pela autora com base na documentação TRAC e Tainacan (2025).

Figura 7 - Diagrama de arquitetura dos requisitos TRAC aplicáveis ao Tainacan.



Fonte: Elaborado pela autora (2026).

A Figura 7 ilustra a arquitetura sociotécnica proposta para a implementação de um repositório digital confiável. O fluxo de trabalho, em conformidade com o modelo de referência OAIS, distingue as funções de gestão, preservação e difusão da informação. Nessa arquitetura, o Tainacan representa a Camada de difusão e acesso. De acordo com o modelo OAIS, ele desempenha as funções de entidade de difusão, na qual as informações sobre o acervo (metadados descritivos) são criadas, mantidas e atualizadas. É o ambiente de trabalho do catalogador e do gestor do acervo. Em segundo lugar, e de forma

mais proeminente para o público, utilizam-se como a entidade de Acesso, sendo o principal responsável por gerar e entregar o Pacote de Informação de Disseminação (DIP) ao usuário final (o "Consumidor" do OAIS) (Artefactual, 2025).

A Camada de Preservação Digital tem a sua responsabilidade assumida pelo Archivematica, que representa as entidades funcionais de Ingestão e Gestão de Arquivo do OAIS. Ao receber o SIP, o Archivematica executa seus micro-serviços para validar, normalizar e empacotar a informação, resultando na criação do Pacote de Informação de Arquivamento (AIP).

O AIP, projetado para a longevidade, é então depositado em um Armazenamento Seguro, que corresponde à entidade de Gestão do Armazenamento do OAIS (Artefactual, 2025). Nesse ambiente, a integridade e a autenticidade do objeto digital são garantidas a longo prazo por meio de estratégias, como a verificação de *checksums* e a manutenção de múltiplas cópias e a análise da auditoria de requisitos.

O diagrama distribui as responsabilidades de conformidade com o TRAC, deixando claro que o requisito não provém de um único *software*, mas da arquitetura integrada:

- Responsabilidade primária do Tainacan: cumpre os requisitos ligados à sua vocação principal: o Controle do DIP (B6) e o Acesso e Uso (B7). Adicionalmente, sua capacidade de exportação em formato aberto é um facilitador crucial para o Plano de Sucessão (A8);
- Responsabilidade primária: Archivematica: atende aos requisitos centrais da preservação digital, como a Geração do AIP (B3), a execução de Estratégias de Preservação (B4) e a garantia da Integridade e Autenticidade (B5) a longo prazo;
- Responsabilidade da instituição: por fim, o diagrama corretamente isola as Seções A (Organizacional) e C (Infraestrutura Técnica), como uma responsabilidade transversal da instituição que governa o repositório, embora ambos os sistemas ofereçam ferramentas que suportam e facilitam o cumprimento dessas políticas.

Em suma, o fluxo de trabalho desenhado não apenas é tecnicamente viável, mas também conceitualmente implementa, na prática, a separação de

funções recomendada pelo modelo OAIS, condição essencial para a construção de um repositório digital confiável.

A análise da estrutura sociotécnica do Tainacan, à luz dos modelos TRAC/OAIS, transcende uma avaliação de conformidade de *software*. Ela revela a materialização de princípios teóricos em uma arquitetura tecnológica concreta. A investigação demonstra que a confiabilidade não é um atributo intrínseco de uma única ferramenta, mas uma propriedade emergente de um ecossistema arquitetado, onde as responsabilidades são distribuídas de forma consciente e especializada (Flores, 2015).

A arquitetura proposta, que integra o Tainacan como camada de acesso e o Archivematica como camada de preservação, implementa na prática a separação funcional preconizada pelo modelo OAIS. Essa separação é fundamental, pois reconhece que as atividades de preservação digital e de acesso/difusão possuem naturezas, requisitos e temporalidades distintas.

Conforme argumenta Cook (2001), em sua crítica à era pós-custodial, o foco excessivo no acesso imediato pode negligenciar a responsabilidade arquivística de garantir a autenticidade e a integridade a longo prazo. A arquitetura analisada mitiga esse risco: o Archivematica assume a função de custódia arquivística, garantindo a fixidez e a proveniência do objeto digital (o AIP), enquanto o Tainacan se especializa na função de difusão, oferecendo interfaces ricas e flexíveis para o acesso (o DIP).

A natureza de código aberto do *software* Tainacan dialoga diretamente com a dimensão social e política da Ciência da Informação, especialmente no que tange à democratização do acesso e à autonomia institucional. Durante (2009), ao discutir os desafios da preservação digital, alerta para os riscos da obsolescência tecnológica e da dependência de fornecedores. A adoção de uma solução de código aberto, como demonstrado na análise do requisito TRAC A8 (Plano de Sucessão), é uma estratégia proativa para garantir a soberania da instituição sobre seu próprio acervo. Ela transfere o poder das corporações de *software* para uma comunidade de prática, alinhando-se a uma visão da informação como um bem comum, que deve ser gerido de forma transparente e sustentável, um ponto também defendido por teóricos da informação aberta.

A análise conclui que o Tainacan, quando inserido em uma arquitetura informacional que inclui um sistema de preservação dedicado, não é apenas uma ferramenta para “colocar coleções online”, é, na verdade, um componente estratégico que permite às instituições construir repositórios digitais que são, ao mesmo tempo, dinâmicos em seu acesso e em sua preservação. A confiabilidade, portanto, é alcançada não pela escolha de um *software*, mas pela adoção de uma filosofia de gestão informacional refletida em uma arquitetura sociotécnica coerente e distribuída baseada nos requisitos OAIS.

4.3 Modelagem de metadados e *crosswalk*

A capacidade de descrever, organizar e intercambiar informações de maneira consistente e escalável é um dos pilares fundamentais dos sistemas de informação contemporâneos (Le Coadic, 2004). Nesse contexto, a modelagem de metadados e as estratégias para garantir a interoperabilidade, como os *crosswalks*, assumem um papel central. Este capítulo detalha os fundamentos teóricos e metodológicos que sustentam esses processos.

A modelagem de metadados é o processo de projetar uma estrutura formal para descrever recursos informacionais. Conforme elucidado por Taylor e Joudry (2015), essa estrutura, conhecida como esquema de metadados, funciona como um modelo conceitual que define quais elementos descritivos são necessários, suas definições semânticas e as regras sintáticas para sua representação. A criação de tal modelo é um passo fundamental para garantir a consistência e a interoperabilidade na organização da informação.

A definição clássica de metadados como “dados sobre dados” (DCMI, 2020) é um ponto de partida, mas a literatura oferece uma visão mais funcional. A *National Information Standards Organization* (NISO) define metadados como informações estruturadas que descrevem, explicam, localizam ou, de outra forma, tornam mais fácil recuperar, usar ou gerenciar um recurso informacional (NISO, 2004, p. 1). A explicação do papel multifacetado dos metadados no ciclo

de vida da informação é frequentemente realizada por meio de sua categorização funcional.

A taxonomia mais consolidada na literatura, formalizada pela NISO (2004) e detalhada por autores como Caplan (2003), classifica os metadados em três tipos principais de acordo com sua função: 1- Metadados Descritivos: utilizados para a descoberta e identificação de recursos, respondendo a questões sobre o conteúdo intelectual da obra, como autor, título e assunto; 2- Metadados Estruturais: utilizados para a navegação e apresentação de recursos complexos, descrevendo as relações entre as partes de um objeto informacional, como a sequência de capítulos de um livro digital; 3- Metadados Administrativos: utilizados para o gerenciamento e a preservação de recursos, abrangendo informações sobre direitos de uso, proveniência e dados técnicos para a preservação digital a longo prazo.

A interoperabilidade entre sistemas distintos é possível por meio da adesão a padrões. Dentre os diversos esquemas, o padrão *Dublin Core* (DC) destaca-se por sua universalidade. Concebido para ser simples e flexível, seu conjunto de elementos nucleares (DCMI Metadata Terms) foi projetado para facilitar a interoperabilidade semântica entre domínios distintos, funcionando como uma língua franca para a descrição de recursos na web (Dublin Core Metadata Initiative, 2020). Outros padrões relevantes incluem o MARC 21, tradicionalmente usado em bibliotecas; o *Metadata Object Description Schema* (MODS), que oferece um meio-termo entre a simplicidade do DC e a complexidade do MARC; e o *Learning Object Metadata* (LOM), específico para recursos educacionais.

A heterogeneidade de esquemas de metadados, embora otimizada para comunidades específicas, cria linguagens direcionadas que podem dificultar a integração de dados. Nesse cenário, o *crosswalk* emerge como um mecanismo essencial.

Crosswalk é definido como uma especificação que mapeia os elementos, a semântica e a sintaxe de um esquema de origem para um esquema de destino (Woodley, 2008). Segundo Woodley (2008), a finalidade de um *crosswalk* é permitir que sistemas que utilizam padrões distintos possam trocar e processar

dados de forma eficaz, um pré-requisito para alcançar a interoperabilidade semântica em ambientes de informação distribuídos. Cenários como a agregação de dados de repositórios institucionais ou a migração de sistemas de catalogação dependem da implementação de *crosswalks* planejados.

A construção de um *crosswalk* é um processo metodológico que envolve a análise comparativa dos esquemas, o mapeamento lógico dos elementos (que pode ser um-para-um, um-para-muitos, muitos-para-um ou um-para-nenhum), a especificação de regras para transformação de conteúdo e, finalmente, a implementação técnica e a validação dos resultados (Chan; Zeng, 2006). Esse processo garante que a “tradução” dos metadados ocorra com a mínima perda de significado, preservando o valor informacional do recurso descrito.

Para ilustrar, utilizaremos um objeto do museu da Pinacoteca de São Paulo, catalogado na plataforma Tainacan. O Tainacan, utiliza o esquema *Dublin Core* para seus metadados centrais, frequentemente estendido com metadados personalizados para cada coleção.

Objeto Selecionado: A pintura "Mestiço" (1934), de Candido Portinari.

Link para o objeto: <https://pinacoteca.org.br/acervo/mestico/>

Identificador Tainacan (ID): 231

Abaixo (Quadro 10) estão os metadados extraídos diretamente da página do objeto no acervo da Pinacoteca. A instituição utilizou uma combinação de campos *Dublin Core* padrão e metadados personalizados, específicos para obras de arte.

Quadro 10 - Metadados descritivos (Tainacan - Pinacoteca)

Campo Tainacan (Metadado)	Valor do Campo (Instância)	Mapeamento DC/Custom
Título	Mestiço	dc:title
Autor	Candido Portinari	dc:creator
Data	1934	dc:date
Dimensões	81 x 65 cm	custom:dimensoes
Técnica	Óleo sobre tela	custom:tecnica
Tipo de Objeto	Pintura	dc:type

Número de Registro	1950.1.1	dc:identifier
Histórico da Peça	Pertenceu a Assis Chateaubriand.	dc:provenance
Direitos Autorais	© Candido Portinari	dc:rights

Fonte: Tainacan pinacoteca.org.

Para melhor compreensão, focaremos no arquivo de imagem digital que representa esta obra no acervo. Assumiremos que a Pinacoteca possui um arquivo mestre de alta resolução, que é a fonte para a imagem de acesso (JPEG) exibida no site. O nosso *crosswalk* será para este arquivo mestre. O objetivo é mapear os metadados descritivos e administrativos disponíveis para o dicionário de dados de preservação PREMIS, focando na entidade Objeto (Object) para descrever o arquivo digital mestre.

O quadro, a seguir, detalha o mapeamento proposto para o arquivo de imagem da obra “Mestiço”.

Quadro 11 - Quadro de *crosswalk*: Tainacan (Pinacoteca) para PREMIS (Entidade Objeto).

Unidade semântica PREMIS	Definição PREMIS	Mapeamento/origem (campo Tainacan)	Valor resultante (exemplo)	Tipo e Regras de Mapeamento
objectIdentifier	Identificador único do objeto digital.	dc:identifier (Número de Registro) + sufixo	1950.1.1_m aster	Mapeamento 1:1 com Transformação. Utiliza-se o número de registro, que é um identificador estável, e adiciona-se um sufixo para indicar que se trata do arquivo mestre de preservação.
objectCategory	Categoria do objeto.	dc:type (do objeto físico)	File	Mapeamento por Inferência. Embora dc:type seja "Pintura", o objeto digital é um arquivo. A regra é: "SE o objeto é digital, ENTÃO objectCategory é 'File'".
objectCharacteristics.compositionLevel	Nível de composição do objeto.	<i>Inferido/Padrão</i>	0	Mapeamento por Inferência. O arquivo mestre é um objeto atômico, portanto, seu nível de composição é 0.
objectCharacteristics.formatName	Nome do formato do arquivo.	Não disponível (Assumido como TIFF)	Tagged Image File Format	Lacuna Semântica (1:0). O formato do arquivo não é um metadado descritivo. Deve ser extraído do arquivo via

				ferramenta de caracterização e mapeado para um nome padrão (ex: do registro PRONOM).
objectCharacteristics .fixity.messageDigest Algorithm	Algoritmo de verificação de integridade.	<i>Não disponível</i>	SHA-256	Mapeamento por Geração. Esta informação é gerada por uma ferramenta no momento da ingestão do arquivo no repositório de preservação.
objectCharacteristics .fixity.messageDigest	Valor da verificação de integridade.	<i>Não disponível</i>	a1b2c3d4... (hash de 64 caracteres)	Mapeamento por Geração. O valor do <i>checksum</i> é calculado e registrado.
originalName	Nome original do arquivo.	<i>Não disponível</i>	portinari_mestico_1934.tif	Lacuna Semântica (1:0). Idealmente, o nome do arquivo original seria capturado durante a digitalização. Se não estiver disponível, pode ser deixado em branco ou preenchido com um nome padronizado.
relationship.relationshipType	Tipo de relação.	<i>Inferido</i>	represents	Mapeamento por Inferência. O sistema de preservação estabelece que o objeto digital (o arquivo) "representa" o objeto intelectual/físico (a pintura).
relationship.relatedObjectIdentifier	Identificador do objeto relacionado.	dc:identifier (Número de Registro)	1950.1.1	Mapeamento 1:1. Cria o vínculo explícito entre o arquivo digital e a obra de arte que ele documenta, usando o identificador estável da obra.
linkingEventIdentifier	Identificador do evento relacionado.	Gerado Automaticamente	evt-001-ingest	Mapeamento por Geração. Vincula o objeto ao evento de "ingestão" no repositório, que, por sua vez, conteria informações sobre a data de digitalização e o agente (<i>software</i> /pessoa) responsável.

Fonte: Elaborado pela autora (2026).

Esse exemplo com um objeto real da Pinacoteca reforça a confirmação da lacuna semântica, o caso real confirma que os metadados descritivos, mesmo em um acervo de alta qualidade, são insuficientes para a preservação. Informações cruciais, como formato, tamanho e checksum (*fixity*), não fazem parte do registro catalográfico e são inerentemente técnicas. Isso justifica, a

necessidade de um fluxo de trabalho de preservação digital que vá além da simples catalogação.

O uso do “Número de Registro” (1950.1.1) como base para o identificador do objeto digital (1950.1.1_master) e para o relacionamento (*relatedObjectIdentifier*) é um ponto crucial. Ele demonstra a importância de se ter um identificador persistente e unívoco para a entidade do mundo real, que serve como âncora para todos os seus derivados digitais.

A decisão de mapear *dc:type* (“Pintura”) para *objectCategory* (“File”) não é uma tradução direta, mas uma decisão lógica baseada no contexto. Isso ilustra que um *crosswalk* eficaz não é apenas uma tabela, mas um conjunto de regras de negócio que interpretam os metadados de origem para criar um registro de destino mais rico e preciso.

A conexão com o ecossistema de preservação, demonstrou-se através da inclusão de campos, como “*linkingEventIdentifier*” evidencia que o registro PREMIS de um objeto não é isolado. É parte de uma rede de informações que o conecta a Eventos (cadeia de acontecimentos com o objeto) e Agentes (quem fez o quê), formando uma cadeia de custódia digital completa.

A análise do mapeamento de metadados da obra “Mestiço”, de Candido Portinari, do acervo da Pinacoteca de São Paulo, para o dicionário de dados PREMIS, transcende um exercício técnico. Este estudo de caso concreto fornece uma base empírica para argumentar sobre a necessidade crítica de planejar a interoperabilidade entre sistemas de gestão de acervos, como o Tainacan, e repositórios de preservação digital. Assim, fica demonstrado que, embora os metadados descritivos sejam o ponto de partida indispensável, eles são insuficientes para garantir a perenidade dos objetos digitais.

O *crosswalk* emerge, portanto, não apenas como uma tabela de correspondência, mas como a ponte metodológica que une os universos da descrição e da preservação. Formalizando a extração de valor do registro catalográfico e, ao mesmo tempo, expõe as lacunas que devem ser preenchidas por processos técnicos subsequentes, como a caracterização de arquivos e a geração de metadados de proveniência, consolidando uma cadeia de custódia digital robusta e confiável.

Portanto, o caso apresentado serve como modelo orientador para instituições que utilizam o Tainacan como plataforma de gestão de acervos e almejam a preservação digital em conformidade com um RDC-Arq. Dessa forma, fica evidente que, embora o Tainacan seja um ponto de partida para a difusão, a transição para um ambiente de preservação exige adaptações e implementações complementares. É imperativo que os pacotes de informação sejam enriquecidos com metadados técnicos e de proveniência, garantindo que os objetos digitais, já acessíveis ao público, tenham sua integridade e autenticidade asseguradas a longo prazo.

5. PROPOSIÇÃO DE REQUISITOS DE PRESERVAÇÃO DIGITAL APLICÁVEIS AO TAINACAN

Este capítulo consolida a pesquisa desenvolvida, transitando da análise diagnóstica para a formulação de uma proposta estratégica. Conforme estabelecido, o Tainacan, enquanto *software* de código aberto fomentado pelo Ibram, consolidou-se como uma solução de excelência para a gestão, organização e difusão de acervos museológicos no Brasil. Contudo, a garantia da perenidade desses acervos digitais exige uma infraestrutura e um fluxo de trabalho que transcendem a sua funcionalidade nativa.

A proposta apresentada a seguir parte de uma premissa fundamental, alinhada ao Modelo de Referência OAIS: a separação estrita de responsabilidades. Nessa arquitetura, o Tainacan não é o ponto de origem do fluxo de preservação, mas, sim, o ponto final do fluxo de acesso. Ele atua como a camada de apresentação e disseminação (o “Access” do OAIS), consumindo os resultados de um processo de preservação que ocorre em um sistema especializado. O objetivo é propor um modelo que integre o Tainacan a um RDC-Arq, permitindo que ele se torne uma interface de acesso confiável para objetos digitais preservados.

5.1 Arquitetura conceitual do modelo proposto

A solução proposta baseia-se em uma arquitetura de microsserviços e na separação de responsabilidades (*separation of concerns*), na qual cada sistema desempenha o papel para o qual foi otimizado. O Tainacan permanece como o Sistema de difusão de acervos e a principal interface para a catalogação e o acesso público, enquanto um sistema especializado em preservação digital, como o Archivematica, assume a função de RDC-Arq.

Diagrama da arquitetura conceitual:

1. Tainacan (ambiente de acesso):

Função: Interface para disusão (podendo criar e editar metadados descritivos) e para o público (descoberta e visualização).

Componente Chave: "*Plugin* conector de preservação".

1.1 *Plugin* conector de preservação (A Ponte):

Função: módulo a ser desenvolvido para o Tainacan. Ele atua como orquestrador do fluxo de preservação, comunicando-se com o sistema de preservação via API.

Responsabilidades:

- iniciar o processo de preservação para um item ou coleção;
- empacotar os dados para envio (criação do pacote de informação de submissão - PIS);
- receber o *feedback* do sistema de preservação (*status*, *logs*);
- armazenar o identificador persistente do Pacote de Informação de Arquivamento (PIA) resultante.

2. Archivematica (Ambiente de Preservação - RDC-Arq):

Função: Sistema especializado que executa o fluxo de trabalho de preservação digital.

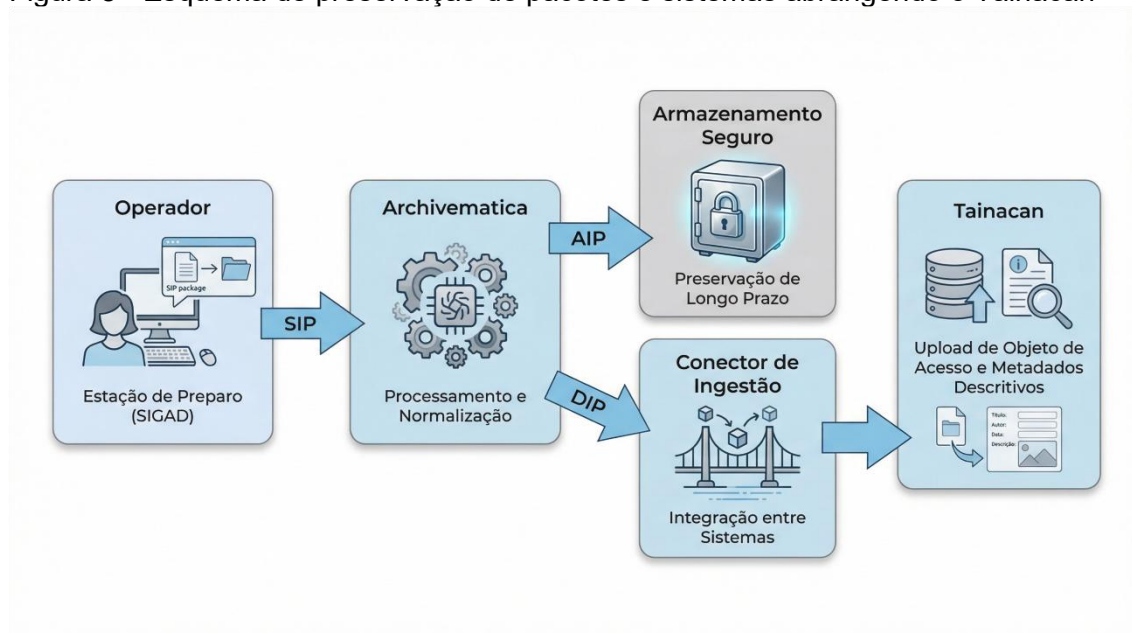
Responsabilidades (Micro-serviços):

- receber o SIP;
- normalizar arquivos para formatos de preservação;
- extrair metadados técnicos (FITS/JHOVE);
- gerar *checksums* (SHA-256);
- criar o registro PREMIS (metadados de preservação);
- gerar o Pacote de Informação de Arquivamento (AIP) e o Pacote de Informação de Disseminação (PID);
- armazenar o AIP em armazenamento seguro e replicado.

3. Armazenamento Seguro (Storage):

Função: Camada de armazenamento de longo prazo, geograficamente distribuída e com verificação de integridade. Pode ser uma solução em nuvem ou *on-premise*⁴³.

Figura 8 - Esquema de preservação de pacotes e sistemas abrangendo o Tainacan



Fonte: Elaborado pela autora com base no OAIS.

⁴³ O termo *on-premise* (ou *on-premises*) refere-se ao modelo de implantação de *software* e infraestrutura de TI no qual os recursos computacionais, como servidores e sistemas de armazenamento, são instalados e operados localmente, dentro das instalações físicas da própria organização. Esse modelo contrasta com a computação em nuvem (*cloud computing*), onde os recursos são hospedados por um provedor externo. A abordagem *on-premise* oferece maior controle sobre os dados e a segurança, mas implica em custos de capital iniciais mais elevados e na responsabilidade total pela manutenção da infraestrutura (Zhang; Nan; Tan, 2020).

5.2 Detalhamento dos componentes e fluxos de preservação

A integração pode ser realizada por duas abordagens principais: uma baseada em *Polling* (verificação periódica) e outra em *Push* (notificação via evento). A própria Artefactual implementou ambas ao longo do tempo na integração do Archivematica ao AtoM.

Segundo Hohpe e Woolf (2004), a abordagem *Polling* (Verificação Agendada) é mais tradicional e desacoplada. Um *script* externo é o responsável por orquestrar a comunicação sendo a referência do modelo: *Scripts* de automação e a API REST do Archivematica. A documentação do *Archivematica Storage Service* é a base para essa interação.

Caminho Computacional para a abordagem *Polling*:

1. Configuração no Archivematica:

1.1 Ao final do fluxo de processamento, o Archivematica é configurado para “Carregar DIP” (*Upload DIP*) para um local específico no seu serviço de armazenamento (Storage Service). Esse local é uma pipeline ⁴⁴ configurada para disponibilizar os DIPs.

1.2 Exemplo de Localização: Um espaço no servidor chamado `/var/archivematica/sharedDirectory/uploadedDIPs/`.

2. O *Script* agendado (o “conector de ingestão”):

A escolha da linguagem *Python* para o desenvolvimento do “Conector de Ingestão” é uma decisão de engenharia fundamentada em princípios de manutenibilidade, eficiência e alinhamento com a comunidade de prática. A própria filosofia de design do *Python*, consolidada no “*Zen of Python*”, prioriza

⁴⁴ No contexto de desenvolvimento de *software*, um *pipeline* de CI/CD (Integração Contínua/Entrega Contínua) é um fluxo de trabalho automatizado que permite a construção, o teste e a implantação de novas versões de um *software* de forma rápida e confiável. O *pipeline* é acionado por alterações no código-fonte e executa uma sequência de etapas, como compilação, testes automatizados e implantação em diferentes ambientes, sendo uma prática fundamental do DevOps (Jani, 2023).

a legibilidade e a simplicidade em detrimento da complexidade (Peterson, 2004). Essa característica é crucial em ambientes institucionais, pois um código claro e legível é mais fácil de manter e evoluir ao longo do tempo, um princípio defendido como essencial para a qualidade de *software* por autores como Martin (2009).

Além disso, o ecossistema de bibliotecas do *Python*, como a *requests* para comunicação HTTP e a *lxml* para análise de XML, acelera o desenvolvimento ao fornecer ferramentas testadas e padronizadas para as tarefas centrais de um fluxo de integração (Van Rossum, 2021). Desse modo, a recomendação não se baseia em preferência, mas na busca por uma solução sustentável e alinhada às boas práticas de engenharia de *software*.

2.1 Tecnologia: um *script* em *Python*, utilizando bibliotecas como *requests* (para chamadas de API), *os* (para manipulação de arquivos) e *json*.

2.2 Agendamento: este *script* é configurado no servidor como uma *cron job*⁴⁵ (no Linux), para ser executado em intervalos regulares (ex: a cada 15 minutos).

2.3 Exemplo de linha no crontab: `*/15 * * * * /usr/bin/python3 /opt/scripts/tainacan_ingest_connector.py`

3. Lógica do Script (tainacan_ingest_connector.py):

Quadro 12 - Lógica do *script* de recebimento e verificação do DIP.

Passo	Ação	Descrição detalhada	Tecnologias/comandos exemplo	Resultado esperado
1	Identificar novos pacotes (<i>polling</i>)	O <i>script</i> consulta a API do <i>Archivematica Storage Service</i> para obter a lista de Pacotes de Disseminação (PID/DIP)	Python: <code>requests.get()</code> API Endpoint: <code>/api/v2/location/<uuid>/entry/</code>	Uma lista de UUIDs de novos pacotes a serem processados.

⁴⁵ Uma tarefa *cron* (*cron job*) é um comando ou *script* agendado para ser executado automaticamente em horários ou intervalos específicos em sistemas operacionais do tipo Unix, como o Linux. O agendamento é gerenciado pelo *daemon cron*, que lê as configurações de um arquivo chamado *crontab* (*cron table*). Esse mecanismo é amplamente utilizado para automatizar tarefas administrativas e de manutenção, como *backups*, limpeza de arquivos temporários e execução de *scripts* de monitoramento (Schwarz, 2000).

		disponíveis. Compara essa lista com um registro local para identificar pacotes ainda não processados.		
2	Obter e preparar o pacote	Para cada novo pacote identificado, o <i>script</i> realiza o <i>download</i> do arquivo compactado (ex: .tar) e o extrai para um diretório de trabalho temporário no servidor.	Python: requests.get(stream=True), tarfile.extractall()	Uma pasta no servidor contendo a estrutura de arquivos e metadados do PID/DIP.
3	Analisar a estrutura do pacote	O <i>script</i> navega pela estrutura de diretórios do PID/DIP extraído para localizar os componentes essenciais: o arquivo de metadados (padrão METS.xml) e a pasta objects/ que contém os arquivos de acesso (ex: imagem.jpg).	Python: os.walk(), pathlib.Path	Caminhos de arquivo exatos para o METS.xml e para o(s) objeto(s) de acesso.
4	Executar o <i>crosswalk</i> de metadados	O script analisa o arquivo METS.xml para extrair os metadados descritivos (ex: dc:title, dc:creator) e o identificador persistente do Pacote de Arquivamento (PIA/AIP). Em seguida, mapeia esses dados para uma estrutura JSON compatível com a API do Tainacan.	Python: xml.etree.ElementTree.parse()	Um objeto JSON contendo os metadados formatados para o Tainacan e o arquivo de acesso pronto para upload.
5	Ingerir o Item no Tainacan	O script interage com a API REST do Tainacan em duas etapas: primeiro, faz o	Python: requests.post() API Endpoints: /wp-json/wp/v2/media, /wp-json/tainacan/v2/coll	Um novo item é criado no Tainacan, com seu arquivo de acesso e todos os

		<i>upload</i> do arquivo de acesso para a biblioteca de mídia, obtendo um ID de anexo; segundo, cria o novo item na coleção, enviando o JSON de metadados e o ID do anexo.	ection/ {id}/items	metadados preenchidos, incluindo o vínculo com o PIA/AIP.
6	Registrar e finalizar (<i>logging & cleanup</i>)	Após a ingestão bem-sucedida, o <i>script</i> registra a operação em um arquivo de log (ex: ingest.log). Em seguida, move o arquivo do PID/DIP processado para um diretório de "concluídos" ou o remove da fila no Archivematica para evitar reproprocessamento.	Python: logging module, shutil.move()	O estado do sistema é atualizado, garantindo que o mesmo pacote não seja processado novamente e que haja um registro auditável da operação.

Fonte: Elaborado pela autora com base nas documentações do archivematica e recomendações dos autores Hohpe e Woolf (2004).

Os autores Hohpe e Woolf (2004) descrevem a abordagem *push* (*webhooks*) mais moderna e eficiente, pois a comunicação é instantânea e iniciada pelo Archivematica. E ocorre por meio da funcionalidade de "DIP Upload" do Archivematica e seus *callbacks*⁴⁶. A documentação sobre *Automating DIP Upload* e o uso de *callbacks* é a referência principal para a sugestão dessa abordagem.

1. Configuração no Archivematica:

Na "Administration Tab"⁴⁷ do Archivematica, na seção "*Pipelines*", é possível configurar um *callback* a ser acionado ao final do processamento.

⁴⁶ Um *callback* (ou função de retorno de chamada) é um padrão de programação no qual uma função é passada como argumento para outra função. A função que recebe o argumento é então responsável por invocar (chamar de volta) o *callback* em um momento apropriado, geralmente após a conclusão de uma operação assíncrona ou em resposta a um evento. Este mecanismo é central para a programação assíncrona em linguagens como JavaScript, permitindo a execução de código de forma não-bloqueante (Gallaba; Mesbah; Beschastnikh, 2015).

⁴⁷ A aba de administração (*administration tab*) é um componente de interface de usuário que serve como ponto de acesso a um conjunto de funcionalidades administrativas de um sistema. Geralmente, ela está contida em um painel de administração maior (conhecido como *dashboard* ou *backend*), que centraliza tarefas como gerenciamento de usuários, configuração de sistema,

O Archivematica pode ser configurado para, após gerar o DIP, enviar uma notificação HTTP POST para uma URL específica. Essa URL será um *endpoint* da aplicação "ponte".

O corpo (*payload*) desta notificação POST contém um JSON com informações sobre o DIP recém-criado, incluindo seu *Universally Unique Identifier* (UUID) ⁴⁸ e o caminho para baixá-lo.

2. O "conector de ingestão" como um serviço web (a aplicação "ponte"):

Tecnologia: em vez de um *script cron*, desenvolve-se uma pequena aplicação web persistente. *Frameworks* como Flask⁴⁹ ou FastAPI ⁵⁰ em *Python*.

Endpoint: a aplicação expõe um *endpoint* de API, por exemplo: POST /api/v1/archivematica-callback. Essa é a URL que se configurou no Archivematica.

3. Para a aplicar a “ponte” entre o Archivematica e o Tainacan:

- 3.1. Passo 1: Receber a notificação (*webhook*): a aplicação fica "ouvindo" no *endpoint* (/api/v1/archivematica-callback). Quando o Archivematica termina um processamento, ele envia a notificação. A aplicação recebe o JSON com os dados do DIP.

moderação de conteúdo e outras operações que não são visíveis para o usuário final. Este padrão de design visa separar as funções de administração das funções de uso geral da aplicação (WORDPRESS, 2024).

⁴⁸ Um UUID (Universally Unique Identifier) é um número de 128 bits utilizado para identificar informações em sistemas computacionais. O seu propósito é permitir a geração de identificadores únicos em sistemas distribuídos, sem a necessidade de uma autoridade central de coordenação, garantindo que a probabilidade de duplicação seja extremamente baixa. O padrão para UUIDs, que define seus diferentes tipos e métodos de geração, foi originalmente especificado na RFC 4122 (Leach; Mealling; Salz, 2005).

⁴⁹ Flask é um microframework para o desenvolvimento de aplicações web em Python, criado por Armin Ronacher. É classificado como "micro" por fornecer um núcleo simples e extensível, sem impor uma estrutura rígida ou ferramentas específicas, como um ORM (Object-Relational Mapper) ou sistema de validação de formulários. Essa filosofia permite grande flexibilidade, sendo ideal para a criação de APIs, microserviços e aplicações de pequeno e médio porte (Aslam; Mohammed; Chandavarkar, 2015).

⁵⁰ FastAPI é um *framework web* para *Python*, moderno e de alta performance, projetado para a construção de APIs. Criado por Sebastián Ramírez, ele se destaca pelo uso de anotações de tipo (*type hints*) do *Python* para validação de dados, serialização e geração automática de documentação interativa (nos padrões OpenAPI e JSON Schema). Sua arquitetura é baseada em *Starlette*, para a funcionalidade web assíncrona, e *Pydantic*, para a validação de dados, o que lhe confere um desempenho comparável ao de aplicações em Node.js e Go (Ramírez, 2018).

3.2. Passo 2: Enfileirar a tarefa (prática recomendada pelos autores Hohpe e Woolf, 2004).

3.2.1. Para não bloquear a resposta ao Archivematica, a aplicação não deve processar o PID/DIP imediatamente. A melhor prática é adicionar a tarefa a uma fila de processamento assíncrono (usando tecnologias como *Celery* com *Redis* ou *RabbitMQ*) (Hohpe; Woolf, 2004).

3.2.2. A aplicação responde imediatamente ao Archivematica, com um HTTP 202 *Accepted* (aceite para processamento) e um *worker* (processo em segundo plano) pega a tarefa da fila para executar os próximos passos.

Segundo Hohpe e Woolf, os passos seguintes dessa abordagem se darão idênticos aos da abordagem anterior. Como descrito abaixo:

4. Etapa de verificação: O *worker* assíncrono executa exatamente a mesma lógica do *script cron*: baixa o DIP a partir das informações recebidas, descompacta, analisa o METS.xml, mapeia os metadados e os ingere no Tainacan via API. A única diferença é que a tarefa foi iniciada por um evento em tempo real, e não por um agendamento.

Quadro 13 - Replicação dos requisitos 3,4,5 e 6 do quadro 12.

3	Analisar a Estrutura do Pacote	O <i>script</i> navega pela estrutura de diretórios do PID/DIP extraído para localizar os componentes essenciais: o arquivo de metadados (padrão METS.xml) e a pasta <i>objects/</i> que contém os arquivos de acesso (ex: imagem.jpg).	Python: os.walk(), pathlib.Path	Caminhos de arquivo exatos para o METS.xml e para o(s) objeto(s) de acesso.
4	Executar o <i>Crosswalk</i> de Metadados	O <i>script</i> analisa o arquivo METS.xml para extrair os metadados descritivos (ex: dc:title, dc:creator) e	Python: xml.etree.ElementTree.parse()	Um objeto JSON contendo os metadados formatados para o

		o identificador persistente do Pacote de Arquivamento (PIA/AIP). Em seguida, mapeia esses dados para uma estrutura JSON compatível com a API do Tainacan.		Tainacan e o arquivo de acesso pronto para <i>upload</i> .
5	Ingerir o Item no Tainacan	O <i>script</i> interage com a API REST do Tainacan em duas etapas: primeiro, faz o <i>upload</i> do arquivo de acesso para a biblioteca de mídia, obtendo um ID de anexo; segundo, cria o novo item na coleção, enviando o JSON de metadados e o ID do anexo.	Python: requests.post() API Endpoints: /wp-json/wp/v2/media, /wpjson/tainacan/v2/collection/{id}/items	Um novo item é criado no Tainacan, com seu arquivo de acesso e todos os metadados preenchidos, incluindo o vínculo com o PIA/AIP.
6	Registrar e Finalizar (<i>Logging & Cleanup</i>)	Após a ingestão bem-sucedida, o <i>script</i> registra a operação em um arquivo de log (ex: ingest.log). Em seguida, move o arquivo do PID/DIP processado para um diretório de "concluídos" ou o remove da fila no Archivematica para evitar reprocessamento.	Python: logging module, shutil.move()	O estado do sistema é atualizado, garantindo que o mesmo pacote não seja processado novamente e que haja um registro auditável da operação.

Fonte: Elaborado pela autora com base nas documentações do archivematica e recomendações dos autores Hohpe e Woolf (2004).

Quadro 14 - Quadro comparativo e recomendação

Característica	Abordagem via Polling (Script Cron)	Abordagem via Push (Webhook)
Eficiência	Menor. Pode haver um atraso (de até 15 min, no exemplo) entre o DIP ficar pronto e ser processado.	Alta. A integração é em tempo real ou quase real.

Complexidade de Implementação	Menor. Um único script Python e uma linha no crontab ⁵¹ .	Maior. Exige uma aplicação web persistente e, idealmente, um sistema de filas (Celery/Redis).
Acoplamento	Baixo. Os sistemas não precisam se conhecer diretamente, apenas compartilhar um local de arquivos.	Médio. O Archivematica precisa conhecer a URL do <i>endpoint</i> da aplicação "ponte".
Robustez	Média. Se o <i>script</i> falhar, só tentará novamente no próximo intervalo.	Alta. Sistemas de filas oferecem mecanismos de retentativa (retry) e monitoramento de falhas.

Fonte: Elaborado pela autora com base nas documentações do archivematica e recomendações dos autores Hohpe e Woolf (2004).

A seleção entre as arquiteturas de *polling* (verificação agendada) e *push* (orientada a eventos via *webhooks*) não é apenas uma escolha técnica, mas uma decisão estratégica que impacta a eficiência, a escalabilidade e a complexidade do ecossistema de preservação digital proposto. A análise de ambas as abordagens, à luz dos padrões de integração de sistemas consolidados por Hohpe e Woolf (2004), revela opções que necessitam de análise institucional e disponibilidade de conhecimento para decidir qual arquitetura implementar.

O modelo de *polling*, no qual o “conector de ingestão” periodicamente verifica o Archivematica em busca de novos pacotes, destaca-se por seu ponto positivo principal: o baixo acoplamento. Conforme argumentam Hohpe e Woolf (2004), o produtor de dados (Archivematica) não necessita de qualquer conhecimento prévio sobre seus consumidores. Ele simplesmente deposita os pacotes em um local acordado, tornando a arquitetura inerentemente mais simples e resiliente a falhas no consumidor. Se o *script* do conector falhar, o Archivematica continua seu trabalho sem interrupções. Essa simplicidade de implementação, que exige apenas um *script* agendado, representa uma vantagem significativa para instituições com recursos de desenvolvimento limitados.

Contudo, o ponto negativo fundamental desta abordagem é a sua ineficiência e latência. O *polling* é, por natureza, um processo que consome

⁵¹ O crontab (abreviação de cron table) é o arquivo de configuração que contém a lista de tarefas a serem executadas pelo daemon cron em sistemas operacionais do tipo Unix. Cada linha no arquivo crontab define uma tarefa agendada, especificando o horário de execução (minuto, hora, dia do mês, mês, dia da semana) e o comando a ser executado. O comando crontab é também o utilitário de linha de comando usado para editar, visualizar e gerenciar esses arquivos de agendamento (Schwarz, 2000).

recursos de rede e Unidade Central de Processamento (CPU) a cada verificação, mesmo quando não há novos dados a serem processados. Hohpe e Woolf (2004) descrevem esse cenário como um “ciclo de espera ocupada” (*busy-waiting cycle*), que pode ser ineficiente. Além disso, ele introduz um atraso no fluxo de trabalho, determinada pelo intervalo do agendamento. Um pacote pode ficar pronto para disseminação, mas só será publicado no Tainacan minutos ou horas depois, na próxima execução do script (Hohpe; Woolf, 2004).

A arquitetura orientada a eventos, implementada via Webhooks, representa um paradigma mais moderno e eficiente. Seu ponto positivo mais evidente é a comunicação em tempo real. Ao adotar um padrão de “publish-subscribe”, onde o Archivematica (publicador) notifica ativamente o conector (subscritor), o processamento é iniciado no momento exato em que o evento ocorre. Isso elimina a latência e o desperdício de recursos do *polling*, resultando em um sistema altamente responsivo e escalável, alinhado com as melhores práticas descritas por Hohpe e Woolf (2004) para sistemas distribuídos de alto desempenho.

O ponto negativo, entretanto, reside na sua maior complexidade de implementação e acoplamento. O produtor (Archivematica) agora precisa conhecer o endereço do consumidor (a URL do *webhook*), criando um acoplamento, ainda que simples, entre os sistemas. Mais significativamente, a implementação exige a criação de um serviço web persistente (uma API) e, idealmente, um sistema de filas para processamento assíncrono, como Celery e Redis. Essa infraestrutura, embora muito mais robusta e resiliente a picos de carga, representa uma barreira técnica e de manutenção mais elevada para a instituição.

Dessarte, a escolha entre *polling* e *push* é um balanço entre o pragmatismo imediato e a visão de longo prazo. A abordagem de *polling* oferece um caminho de implementação mais rápido e simples, adequado para ambientes com baixo volume de processamento. No entanto, para uma solução de nível institucional, escalável e robusta, a arquitetura orientada a eventos (*push*) é superior. Alinhada aos princípios de eficiência e reatividade defendidos por Hohpe e Woolf (2004), ela constitui a base para um ecossistema de preservação digital integrado e de alto desempenho, sendo, portanto, a recomendação final

deste trabalho para uma implementação de referência pelo Ibram e pela comunidade do Tainacan.

5.3 Implicações e recomendações

A presente tese demonstrou a viabilidade e a necessidade de integrar o Tainacan, *software* de gestão de acervos fomentado pelo Ibram, a um ecossistema de preservação digital robusto, baseado em padrões internacionais, como o Modelo de Referência OAIS. A arquitetura proposta, que posiciona o Tainacan como a camada de acesso e disseminação de um RDC-Arq, não é apenas um exercício técnico, ela carrega implicações estratégicas para a política de gestão de acervos digitais liderada pelo Ibram. Esta seção articula essas implicações e formaliza um conjunto de recomendações para todas as instituições que utilizam o Tainacan como difusor de informações.

A transição de uma ferramenta de gestão para um componente de um ecossistema de preservação significa uma mudança de paradigma, na qual o Tainacan evolui de um *software* primariamente focado na difusão para uma peça pertencente a um conjunto de soluções que asseguram a perenidade dos acervos digitais. Essa evolução agrega valor à aplicação, alinhando-a aos princípios de gestão do ciclo de vida completo da informação digital, um conceito central para a garantia de acesso contínuo ao conhecimento (Le Coadic, 2004).

Ao facilitar a digitalização, as instituições assumem uma responsabilidade fiduciária sobre os ativos digitais gerados. A ausência de uma estratégia de preservação anexa, conforme alertam os modelos de maturidade, como o *Digital Preservation Capability Maturity Model* (DPCMM), dos riscos de perda e obsolescência, tornando a integração com um RDC-Arq um imperativo para o cumprimento desta responsabilidade. Nesse cenário, a unidade gestora do Tainacan tem a oportunidade de exercer liderança na padronização de uma infraestrutura nacional para a preservação de acervos museológicos, um papel de coordenação central, que é fundamental para o desenvolvimento de ecossistemas de dados coesos e interoperáveis (Flores, 2012).

Com base nessas implicações, a primeira recomendação é que, as unidades gestoras do Tainacan fomentem o desenvolvimento e a manutenção do “conector de ingestão” para o Tainacan. A sustentabilidade de *softwares* de código aberto nas instituições depende de um roteiro de desenvolvimento claro e do suporte institucional, evitando que se tornem soluções sem um orientador responsável. Ao incorporar o conector ao *roadmap* oficial do Tainacan e ao orquestrar seu desenvolvimento comunitário, o Ibram assegura a perenidade e a qualidade da solução, reduzindo a barreira de adoção para toda a rede que necessita.

Em segundo lugar, recomenda-se que as unidades gestoras do Tainacan definam e documentem uma Arquitetura de Referência Nacional. A publicação de manuais de boas práticas e notas técnicas que formalizem o fluxo “preservação primeiro” é uma ação de governança da informação essencial. Tais documentos funcionam como instrumentos de normalização que promovem a interoperabilidade semântica e técnica entre as instituições, um pré-requisito para a criação de redes federadas de acervos digitais (Santos; Flores, 2018).

A terceira recomendação propõe que as unidades gestoras do Tainacan avaliem a criação de uma infraestrutura de “Preservação como Serviço” (PaaS). A literatura sobre economia da preservação digital demonstra que os custos de implementação e manutenção de um RDC-Arq podem ser proibitivos para instituições de pequeno e médio porte (Lazorchak, 2011). Um modelo de serviço centralizado, gerenciado pelo Ibram ou por um consórcio, representa uma solução economicamente viável, que democratiza o acesso à preservação baseada em modelos consolidados, mitigando a desigualdade tecnológica entre as instituições e potencializando a criação de uma rede de preservação distribuída e resiliente.

Subsequentemente, é imprescindível que se desenvolva um Programa Nacional de Capacitação em Preservação Digital. A eficácia de qualquer infraestrutura tecnológica está intrinsecamente ligada à competência dos profissionais que a operam. A capacitação contínua é um dos pilares para a certificação de repositórios digitais confiáveis, como o TRAC (ISO 16363), e deve abranger não apenas o uso de ferramentas, mas os princípios teóricos do ciclo de vida da informação digital (Conselho Nacional de Arquivos, 2012). Ao integrar

a preservação digital de forma transversal em seus programas educacionais, o Ibram capacita os profissionais de museus a se tornarem agentes ativos na salvaguarda do patrimônio digital.

Portanto, recomenda-se o estabelecimento de um Comitê Consultivo de Preservação Digital. A governança de uma estratégia nacional de preservação é um processo dinâmico que exige reavaliação constante diante das rápidas mudanças tecnológicas. A criação de um comitê de especialistas, alinhado aos modelos de governança participativa, assegura que as políticas e os padrões técnicos permaneçam relevantes e eficazes, fornecendo ao Ibram os subsídios técnico e estratégico necessários para a tomada de decisões informadas e para a evolução contínua do ecossistema de preservação (Lee; Tibbo, 2011).

6 CONCLUSÃO

A presente tese partiu da constatação de uma lacuna crítica no ecossistema de gestão de acervos digitais no Brasil: a dissociação entre as plataformas de difusão, representadas pelo *software* Tainacan, e a necessidade imperativa da preservação digital a longo prazo. Investigou-se como essa ponte poderia ser construída, não apenas do ponto de vista técnico, mas também conceitual e estratégico. Este capítulo consolida os achados desta pesquisa, articula suas contribuições para a teoria e para a prática, reconhece suas limitações e propõe caminhos para investigações futuras.

A tese demonstrou que o Tainacan, em sua concepção atual, é uma ferramenta otimizada para as funções de gestão de metadados descritivos e disseminação de acervos, correspondendo à entidade funcional “Access” do Modelo de Referência OAIS (ISO 14721). Contudo, ele não abarca nativamente as complexas responsabilidades das entidades “Ingestão” (*Ingest*), “Gestão de Arquivo” (*Archival Storage*) e “Gestão de Dados” (*Data Management*), que são o cerne de um RDC-Arq. A análise de caso, ao simular um *crosswalk* de metadados do Tainacan para o padrão de preservação PREMIS, revelou empiricamente esta lacuna: informações vitais para a perenidade, como dados

de fixidez, proveniência de eventos de preservação e caracterização técnica detalhada, são inexistentes no modelo de dados descritivo.

Concluiu-se, portanto, que a estratégia mais robusta não é tentar transformar o Tainacan em um sistema de preservação monolítico, mas, sim, integrá-lo a uma arquitetura de referência baseada na separação de responsabilidades. O fluxo de trabalho "Preservação Primeiro" (*Preservation First*), no qual os objetos digitais são primeiro processados em um RDC-Arq (como o Archivematica) e, como resultado, publicados no Tainacan através de um Pacote de Informação de Disseminação (DIP), emergiu como a solução arquiteturalmente mais sólida e sustentável. A viabilidade técnica dessa integração foi detalhada através da especificação de dois modelos de conectores — um baseado em *polling* e outro em *webhooks* —, ambos fundamentados nos padrões de integração de sistemas (Hohpe; Woolf, 2004).

A principal contribuição teórica desta tese reside na contextualização e aplicação de um modelo conceitual OAIS-TRAC para o domínio específico dos museus brasileiros que utilizam o Tainacan. Embora o modelo OAIS e os critérios de confiabilidade (como o TRAC/ISO 16363) sejam padrões internacionais genéricos, este trabalho os traduziu para a realidade concreta da comunidade nacional.

Ao posicionar o Tainacan explicitamente como a entidade "Acesso" e ao detalhar sua necessária articulação com as demais entidades funcionais do OAIS, a tese oferece um modelo mental e um vocabulário teórico que permitem aos profissionais de museus e ao Ibram discutirem e planejarem a preservação digital de forma estruturada e alinhada aos padrões globais. Assim, desmistifica-se a ideia de que uma única ferramenta pode resolver todos os problemas do ciclo de vida digital, propondo, em vez disso, um modelo de ecossistema interoperável.

No campo prático, a tese oferece duas contribuições diretas. A primeira é a especificação técnica detalhada dos requisitos para o "Conector de Ingestão", um novo componente de *software* para o ecossistema Tainacan. A lógica do *script*, os *endpoints* de API a serem consumidos e a estrutura de dados a ser manipulada foram delineados, fornecendo um roteiro claro que pode servir de

base para o desenvolvimento efetivo do *plugin* pela comunidade ou pelo Ibram. A segunda contribuição prática é o conjunto de recomendações estratégicas direcionadas ao Ibram.

Essas recomendações (fomento ao desenvolvimento, definição de uma arquitetura de referência, avaliação de um modelo PaaS e capacitação) traduzem os achados da pesquisa em um plano de ação institucional, com potencial para impactar positivamente centenas de museus no país, transformando a política de gestão de acervos em uma política de salvaguarda do patrimônio digital.

A principal limitação deste trabalho é de natureza empírica: a arquitetura e o "Conector de Ingestão" foram propostos e especificados em nível teórico e técnico, mas não foram implementados e testados em um ambiente de produção real. Um protótipo funcional não foi desenvolvido, de modo que a avaliação de desempenho, a complexidade real da implementação e os desafios de usabilidade em um museu-piloto não puderam ser aferidos. Ademais, a análise de custos para a implementação da infraestrutura de RDC-Arq e do modelo PaaS foi abordada apenas em nível conceitual, carecendo de um estudo de viabilidade econômica aprofundado.

As limitações identificadas abrem caminhos para uma agenda de pesquisas futuras. Recomenda-se, primeiramente, o desenvolvimento de um protótipo funcional do "Conector de Ingestão" e sua implementação em um projeto-piloto com um ou mais museus que utilizam o Tainacan. Tal estudo permitiria validar e refinar os requisitos técnicos, além de conduzir uma análise de usabilidade com os profissionais da instituição.

Em segundo lugar, sugere-se uma pesquisa dedicada à análise de custo-benefício e aos modelos de sustentabilidade econômica para a oferta de "Preservação como Serviço" (PaaS) pelos órgãos que adotarem. Nesse sentido, a investigação poderia comparar diferentes provedores de infraestrutura em nuvem e modelos de subscrição.

Por fim, uma terceira via de investigação poderia explorar a aplicação de tecnologias de inteligência artificial e aprendizado de máquina para aprimorar o fluxo de trabalho, como a categorização automática de objetos e a extração de

metadados a partir do conteúdo visual dos arquivos, otimizando ainda mais o processo de ingestão no RDC-Arq. Tais pesquisas dariam continuidade ao trabalho aqui iniciado, aproximando ainda mais a comunidade museológica brasileira de uma solução completa, sustentável e arquivística para a preservação do patrimônio digital.

REFERÊNCIAS

- ADOBE. Watched folder in AEM forms. **Adobe experience manager 6.5**, [s.l.], 23 nov. 2025. Disponível em: <https://experienceleague.adobe.com/en/docs/experience-manager-65/content/forms/publish-process-aem-forms/watched-folder-in-aem-forms> Acesso em: 8 jan. 2026.
- ALENCAR, Thayane Moraes; MARTINS, Dalton Lopes. Desafios e estratégias para a preservação digital de objetos culturais: uma revisão sistemática. **Informação & Informação**, [S. l.], v. 30, n. 4, p. 111–128, 2025. DOI: 10.5433/1981-8920.2025v30n4p111. DOI: <https://doi.org/10.5433/1981-8920.2025v30n4p111> Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/52597> Acesso em: 14 jan. 2026.
- ARTEFACTUAL SYSTEMS. **Archivematica documentation**. Vancouver: Artefactual Systems, 2023. Disponível em: <https://www.archivematica.org/en/docs/> Acesso em: 10 dez. 2025.
- ARTEFACTUAL SYSTEMS. **Archivematica**: open-source digital preservation system. [S. l.], [2023]. Disponível em: <https://www.archivematica.org/> Acesso em: 15 dez. 2025.
- ARTEFACTUAL SYSTEMS. **Technical architecture**. [S. l.], [2023]. (Archivematica Documentation, 1.10.2). Disponível em: <https://www.archivematica.org/en/docs/archivematica-1.10/getting-started/overview/technical/> Acesso em: 15 dez. 2025.
- ARTEFACTUAL SYSTEMS. **Using AtoM 2.x with Archivematica**. [S. l.], [2023]. (Archivematica Documentation, 1.17.1). Disponível em: <https://www.archivematica.org/en/docs/archivematica-1.17/admin-manual/installation-setup/integrations/atom-setup/> Acesso em: 15 dez. 2025.
- ARELLANO, Márdero A. **Critérios para a preservação digital da informação científica**. 2008. 215 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2008. Disponível em: <https://repositorio.unb.br/handle/10482/3439> Acesso em: 15 dez. 2025.
- ASLAM, Farheen A.; MOHAMMED, Haleema Noor; CHANDAVARKAR, B. R. Efficient way of web development using python and flask. **International journal of advanced research in computer science**, [s. l.], v. 6, n. 2, p. 54-57, mar./abr. 2015. Disponível em: <https://doi.org/10.26483/ijarcs.v6i2.2434> Acesso em: 14 jan. 2026.

BECKER, Christoph. *et al.* A framework and testbed for identifying and evaluating viable preservation strategies. *In*: INTERNATIONAL CONFERENCE ON DIGITAL PRESERVATION, 6., 2009, San Francisco. **Proceedings** San Francisco: California Digital Library, Disponível em: https://www.researchgate.net/publication/221323948_A_Generic_Framework_for_the_Description_and_Evaluation_of_Preservation_Plans, 2009. p. 1-8. Acesso em: 8 jan. 2026.

BERNERS-LEE, T.; MASINTER, L.; MCCAILL, M. **RFC 1738**: Uniform Resource Locators (URL). [S. l.]: IETF, dez. 1994. Disponível em: <https://www.rfc-editor.org/rfc/rfc1738.html>. Acesso em: 8 jan. 2026.

BELLOTTI, Heloísa Liberalli. **Arquivística**: objetos, princípios e rumos. São Paulo: Associação de Arquivistas de São Paulo, 2002.

BIAZUS, Glaucia; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. O uso do DSpace para a preservação digital: uma análise da literatura. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 25, p. 1-20, 2020.

BIJKER, Wiebe E. **Of bicycles, bakelites, and bulbs**: toward a theory of sociotechnical change. Cambridge, MA: MIT Press, 1995.

BRASIL. **Decreto n. 10.278 de 18 de março de 2020**. Regulamenta o disposto no inciso X do caput do art. 3º da Lei nº 13.874, de 20 de setembro de 2019, e no art. 2º-A da Lei nº 12.682, de 9 de julho de 2012, para estabelecer técnica e os requisitos para a digitalização de documentos públicos ou privados, a fim de que os documentos digitalizados produzam os mesmos efeitos legais dos documentos originais. 2021. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/D10278.htm Acesso em: 23 ago. 2022.

BRASIL. **Lei n. 8.159, de 8 de janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8159.htm Acesso em: 26 set. 2024.

BRASIL. **Lei n. 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991 e dá outras providências. Diário Oficial da União: seção 1, Brasília, DF, p. 1, 18 nov. 2011. Edição extra.

BRASIL. **Lei n. 13.869, de 5 de setembro de 2019**. Altera o Decreto-Lei nº 3.689, de 3 de outubro de 1941 (Código de Processo Penal), para incluir o art. 158-A e seguintes, que tratam da cadeia de custódia. Diário Oficial da União: seção 1, Brasília, DF, p. 4, 6 set. 2019.

CANADA. Canadian Heritage Information Network. **Digital Preservation Recommendations for Small Museums**. Ottawa: Government of Canada, 2024. Disponível em: <https://www.canada.ca/en/heritage-information-network/services/digital-preservation/recommendations-small-museums.html>. Acesso em: 19 fev. 2025.

CAPLAN, Priscilla. **Understanding metadata**. Chicago: American Library Association, 2003.

CENTER FOR RESEARCH LIBRARIES. **Trustworthy Repositories Audit & Certification (TRAC)**: criteria and checklist. Chicago, IL: CRL, 2007. Disponível em: <https://www.crl.edu/sites/default/files/2025-01/trac.pdf> Acesso em: 15 dez. 2025.

CHAMBERLIN, Donald D. Early history of SQL. **IEEE Annals of the History of Computing**, [s. l.], v. 34, n. 4, p. 78-82, out./dez. 2012. DOI: 10.1109/MAHC.2012.61.

CHAN, Lois Mai; ZENG, Marcia Lei. Metadata Interoperability and Standardization: a study of methodology, part i: achieving interoperability at the schema level. **D-Lib Magazine**, v. 12, n. 6, jun. 2006. Disponível em: <http://www.dlib.org/dlib/june06/chan/06chan.html> Acesso em: 5 jan. 2026.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **Resolução n. 43, de 4 de setembro de 2015**. [S. l.]: Ministério da Justiça, 4 set. 2015. Disponível em: http://conarq.arquivonacional.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf Acesso em: 11 dez. 2025.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **Resolução n. 51, de 25 de agosto de 2023**. Dispõe sobre as diretrizes para a implementação de repositórios Arquivísticos Digitais Confiáveis - versão 2. Diário Oficial da União, seção 1, n. 91, Brasília, DF, p. 89, 2023. Disponível em: <https://www.gov.br/conarq/pt-br/legislacao-arquivistica/resolucoes-do-conarq/resolucao-conarq-no-51-de-25-de-agosto-de-2023> Acesso em: 10 abr. 2025.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **MoReq-Jus**: modelo de requisitos para sistemas informatizados de gestão de processos e documentos do poder judiciário brasileiro. Rio de Janeiro: Arquivo Nacional, 2012.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Audit and certification of trustworthy digital repositories**: CCSDS 652.0-M-2. Washington, DC: CCSDS, 2024. (Magenta book). Disponível em: <https://ccsds.org/Pubs/652x0m2.pdf> Acesso em: 15 dez. 2025.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Reference model for an Open Archival Information System (OAIS)**: CCSDS 650.0-M-2.

Washington, DC: CCSDS, 2012. (Magenta book). Disponível em: <https://public.ccsds.org/Pubs/650x0m2.pdf> Acesso em: 15 dez. 2025.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS). **Reference model for an Open Archival Information System (OAIS)**: CCSDS 650.0-M-3. Washington, DC: CCSDS, 2024. Disponível em: <https://ccsds.org/Pubs/650x0m3.pdf> Acesso em: 14 dez. 2025.

CONWAY, Paul. Preservation in the age of google: digitization, digital preservation, and dilemmas. **The Library Quarterly**, v. 80, n. 1, p. 61-79, jan. 2010.

CORRADO, Edward M.; MOULAISON SANDY, Heather. **Digital preservation for libraries, archives, and museums, rowman and littlefield publishers**. 2. ed. Lanham: Rowman & Littlefield Publishers, 2017. 402 p.

COOK, Terry. Archival science and postmodernism: new formulations for old concepts. **Archival Science**, v. 1, n. 1, p. 3-24, 2001.

DATE, C. J. **Introdução a sistemas de bancos de dados**. 8. ed. Rio de Janeiro: Elsevier, 2004.

DSPACE. **DSpace Documentation**. [S.l.]: Lyris, 2023. Disponível em: <https://wiki.lyris.org/display/DSDOC/> Acesso em: 26 dez. 2025.

DOERR, Martin. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. **AI Magazine**, [s. l.], v. 24, n. 3, p. 75-92, 2003. Disponível em: <https://aaai.org/ojs/index.php/aimagazine/article/view/1720>. Acesso em: 8 jan. 2026.

DUBLIN CORE METADATA INITIATIVE (DCMI). **DCMI Metadata Terms**. Dublin Core, 2020. Disponível em: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms> Acesso em: 5 jan. 2026.

DUBLIN CORE METADATA INITIATIVE (DCMI). **DCMI Metadata Terms**. Dublin Core, 2020. Disponível em: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms> Acesso em: 5 jan. 2026.

DURANTI, Luciana. Fidedignidade e autenticidade: o caso dos documentos arquivísticos eletrônicos. **Estudos históricos**, Rio de Janeiro, v. 8, n. 16, p. 213-224, 1995.

DURANTI, Luciana. From digital diplomatics to digital records forensics. **Archivaria**, v. 68, p. 39-66, 2009.

DURANTI, Luciana. **The InterPARES 2 Project (2002-2007)**: an overview. **Archivaria**: [s. l.], 2007. p. 113-121.

EX LIBRIS. **Rosetta**: digital asset management and preservation. Jerusalém: Ex Libris, 2024. Disponível em: <https://exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation> Acesso em: 10 dez. 2025.

FERREIRA, Miguel; RAMALHO, José Carlos; FARIAS, Luís. RODA: a digital repository for authentic digital objects. In: INTERNATIONAL CONFERENCE ON DIGITAL LIBRARIES & THE FUTURE OF KNOWLEDGE, 13., 2011, Nova Deli. **Proceedings...** Nova Deli: TERI, 2011. p. 1-10.

FIELDING, R. *et al.* RFC 2616: Hypertext Transfer Protocol - HTTP/1.1. [S. l.]: IETF, jun. 1999. Disponível em: <https://www.rfc-editor.org/rfc/rfc2616.html>. Acesso em: 8 jan. 2026.

FIELDING, Roy Thomas. **Architectural styles and the design of network-based software architectures**. 2000. 180 f. Tese (Doutorado em Ciência da Computação e Informação) – University of California, Irvine, 2000. Disponível em: https://roy.gbiv.com/pubs/dissertation/fielding_dissertation.pdf. Acesso em: 8 jan. 2026.

FLORES, Daniel. Políticas de preservação digital para documentos arquivísticos. **Perspectivas em Ciência da Informação**, v. 20, n. 4, p. 197-217, out./dez. 2015.

FLORES, Daniel; ROCCO, Brenda Couto de Brito; SANTOS, Henrique Machado dos. Cadeia de custódia para documentos arquivísticos digitais. **Acervo**, Rio de Janeiro, v. 29, n. 2, p. 117–132, jul./dez. 2016. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/717> Acesso em: 10 dez. 2025.

HARVEY, Ross; WEATHERBURN, Jaye. **Preserving digital materials**. 3rd ed. Lanham: Rowman & Littlefield, 2019.

GALLABA, Keheliya; MESBAH, Ali; BESCHASTNIKH, Ivan. Don't call us, we'll call you: Characterizing callbacks in JavaScript. In: ACM/IEEE INTERNATIONAL CONFERENCE ON AUTOMATED SOFTWARE ENGINEERING, 30., 2015, Lincoln. **Proceedings...** Lincoln: IEEE, 2015. p. 247-258. DOI: 10.1109/ASE.2015.18.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GILLILAND, A. J. Setting the stage: defining archival description and its components. **Archival science**, v. 14, n. 1, p. 3-15, 2014.

HEDSTROM, Margaret. **Digital preservation**: a time for action. Ann Arbor: School of Information, University of Michigan, 2000. Disponível em: <https://www.clir.org/pubs/reports/pub89/archival/> Acesso em: 10 dez. 2025.

HOHPE, Gregor; WOOLF, Bobby. **Enterprise integration patterns**: designing, building, and deploying messaging solutions. Boston: Addison-Wesley, 2004.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 14721:2012**: space data and information transfer systems - Open archival information system (OAIS) - reference model. Geneva: ISO, 2012a.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 16363:2012**: space data and information transfer systems — audit and certification of trustworthy digital repositories. Geneva: ISO, 2012b.

ITHAKA S+R. The effectiveness and durability of digital preservation and curation systems. New York: Ithaka S+R, 2020. Disponível em: <https://sr.ithaka.org/publications/the-effectiveness-and-durability-of-digital-preservation-and-curation-systems/> Acesso em: 10 dez. 2025.

JANI, Yash. Implementing continuous integration and continuous deployment (CI/CD) in modern software development. **International Journal of Science and Research**, [s. l.], v. 12, n. 5, p. 1456-1462, maio 2023.

JONES, Maggie; BEAGRIE, Neil. **Preservation management of digital materials**: a handbook. London: The British Library, 2001.

KEEP SOLUTIONS. RODA: Repository of Authentic Digital Records. Braga: KEEP SOLUTIONS, 2023. Disponível em: <https://www.keep.pt/roda/> Acesso em: 10 dez. 2025.

LAZORCHAK, B. The Economics of Digital Preservation. *In*: HARVEY, R. (Ed.). Digital Curation: A curation-education-praxis-nexus. Chicago: ALA, 2011. p. 123-140.

LE COADIC, Yves-François. **A Ciência da Informação**. 2. ed. Brasília: Briquet de Lemos, 2004.

LEE, Christopher A.; TIBBO, Helen R. where's the “archivist” in “digital curation?": exploring the possibilities for archival augmentation of scientific data. **Archival Science**, v. 11, n. 3-4, p. 245-268, 2011.

LEACH, P.; MEALLING, M.; SALZ, R. **RFC 4122**: A Universally Unique Identifier (UUID) URN Namespace. [S. l.]: IETF, jul. 2005. Disponível em: <https://datatracker.ietf.org/doc/html/rfc4122>. Acesso em: 8 jan. 2026.

LERDORF, Rasmus; TATROE, Kevin. **Programming PHP**. 1. ed. Sebastopol, CA: O'Reilly Media, 2002.

LEWIS, S.; CASTRO, P.; JONES, R. SWORD: Facilitating deposit to repositories. **Information services & use**, v. 32, n. 3-4, p. 171-181, 2012.

LOPES, Marilu; GABRIEL, Mariana M.; BARETA, Gisele M. S. Cadeia de custódia: uma abordagem preliminar. **Revista acadêmica: ciências agrárias e**

ambientais, Curitiba, v. 12, n. 1, p. 51-56, jan./mar. 2014. Disponível em: <https://revistas.ufpr.br/academica/article/download/9022/6315/26746> Acesso em: 10 dez. 2025.

1. LUNA, M. Dúvidas de padrões atendidos. Tainacan, jul. 2024. Disponível em: <https://tainacan.discourse.group/t/duvidas-de-padroes-atendidos/1997> Acesso em: 10 dez. 2025.

LOR, Peter; REIS, Alcenir Soares dos. Preservação digital distribuída: uma revisão de literatura. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 25, p. 1-21, 2020.

LYRASIS. **DSpace**. Atlanta: LYRASIS, 2024. Disponível em: <https://dspace.lyrasis.org/> Acesso em: 10 dez. 2025.

LYRASIS. **DSpace documentation**. [S.l.]: Lyrasis, 2023. Disponível em: <https://wiki.lyrasis.org/display/DSDOC/> Acesso em: 26 dez. 2025.

MARCONI, M. de A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 8. ed. São Paulo: Atlas, 2017.

MARTINS, Dalton; SIQUEIRA, Joyce. **Manual repositório Tainacan**: 2014. [S. l.]: L3P/UFG, Disponível em: <https://tainacan.org/wp-content/uploads/2014/05/Manual-Repositorio.pdf> Acesso em: 2 jan. 2026.

MARTINS, Luciana Conrado; MARTINS, Dalton Lopes. Experimentações sociotécnicas para organização e difusão de coleções digitais universitárias: o caso do projeto Tainacan. **Revista CPC**, v. 15, n. 30esp, p. 34-61, 2020.

MARTINS, Dalton Lopes. *et al.* Repositório digital com o software livre Tainacan: revisão da ferramenta e exemplo de implantação na área cultural com a revista filme cultura. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 18., 2017. **Anais... XVIII Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação**, 2017.

MARTINS, Dalton Lopes; LEMOS, Daniela Lucas da Silva; ANDRADE, Morgana Carneiro de. Tainacan e Omeka: proposta de análise comparativa de softwares para gestão de coleções digitais a partir do esforço tecnológico para uso e implantação. **Informação & Informação**, [S. l.], v. 26, n. 2, p. 569–595, 2021. DOI: 10.5433/1981-8920.2021v26n2p569. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/41208> Acesso em: 15 jan. 2026.

MICROSOFT. Estilo de arquitetura *web-queue-worker*. Azure Architecture Center, 6 jan. 2026. Disponível em: <https://learn.microsoft.com/pt-br/azure/architecture/guide/architecture-styles/web-queue-worker>. Acesso em: 8 jan. 2026.

MUMMA, C. C. Building On OAIS: Processing born-digital records using archivematica. **Archiving Conference**, Springfield, v. 2012, n. 1, p. 14-19, 2012. DOI: 10.2352/issn.2168-3204.2012.9.1.art00014.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). **Understanding metadata**. Bethesda, MD: NISO Press, 2004. Disponível em: <https://www.niso.org/publications/understanding-metadata-2004> Acesso em: 5 jan. 2026.

NOGUEIRA, Fabiano Mikalauskas de Souza; AMORIM, Arivaldo Leão de. Reconstrução digital: a preservação em meio digital da herança cultural. **Revista FÓRUM PATRIMÔNIO: Ambiente Construído e Patrimônio Sustentável**, [S. l.], v. 11, n. 1, 2020. Disponível em: <https://periodicos.ufmg.br/index.php/forumpatrimo/article/view/34009>. Acesso em: 15 jan. 2026.

NOSOWITZ, Dan; GOODWIN, Michael. What Is an API endpoint? **IBM**. Disponível em: <https://www.ibm.com/think/topics/api-endpoint>. Acesso em: 8 jan. 2026.

OPEN SOURCE INITIATIVE. The open source definition. Disponível em: <https://opensource.org/osd> Acesso em: 26 dez. 2025.

OWASP. Cross Site Scripting (XSS). OWASP Community [2024]. Disponível em: <https://owasp.org/www-community/attacks/xss/>. Acesso em: 8 jan. 2026.

PAGE, Matthew J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, [S. l.], v. 372, n. 71, p. 1-9, 2021. DOI: 10.1136/bmj.n71.

PETERSON, Tim. **PEP 20: the zen of python**. Python software foundation, 2004. Disponível em: <https://peps.python.org/pep-0020/> Acesso em: 5 jan. 2026.

PINHEIRO, Ana; GLÓRIA, Pedro; ANTUNES, Gabriel. Desafios na implementação de um repositório digital confiável: o caso do RODA. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 13., 2018, Lisboa. **Atas...** Lisboa: BAD, 2018. p. 1-12.

PREMIS EDITORIAL COMMITTEE. **PREMIS data dictionary for preservation metadata**: version 3.0. Washington, D.C.: Library of Congress, 2015. Disponível em: <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf> Acesso em: 26 dez. 2025.

PRESERVICA. **Digital preservation software**. Abingdon: Preservica, 2024a. Disponível em: <https://preservica.com/>. Acesso em: 10 dez. 2025.

PRESERVICA. **Archivematica vs Preservica**. Abingdon: Preservica, 2024b. Disponível em: <https://preservica.com/resources/white-papers/archivematica-vs-preservica> Acesso em: 10 dez. 2025.

RAMÍREZ, Sebastián. **FastAPI**: modern, fast (high-performance), web framework for building APIs with python. [S. l.], [2018-]. Disponível em: <https://fastapi.tiangolo.com/>. Acesso em: 8 jan. 2026.

RESEARCH LIBRARIES GROUP; ONLINE COMPUTER LIBRARY CENTER. Trusted digital repositories: attributes and responsibilities. Mountain View, CA: RLG, 2002. Disponível em: <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf> Acesso em: 10 dez. 2025.

RESEARCH LIBRARIES GROUP (RLG); COMPUTER LIBRARY CENTER (OCLC). Working group on digital archive attributes. **Trusted digital repositories**: attributes and responsibilities. Mountain View, CA: RLG, 2002. Disponível em: <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf> Acesso em: 15 dez. 2025.

RESEARCH LIBRARIES GROUP (RLG); COMPUTER LIBRARY CENTER (OCLC). **Trusted Digital Repositories**: attributes and responsibilities. Mountain View, CA: RLG, 2002. Disponível em: <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf> Acesso em: 10 dez. 2025.

RIBEIRO, Célia; LUCCHESI, Cláudia. A preservação digital distribuída e a definição de um modelo de governança. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 18., 2017, Marília. Anais... Marília: UNESP, 2017.

RLG-NARA. **Trustworthy repositories audit & certification**: criteria and checklist. Chicago: The Center for Research Libraries, 2007. Disponível em: https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf Acesso em: 2 jan. 2026.

RODRIGUEZ, Alex. **RESTful Web services**: the basics. IBM developerWorks, 06 nov. 2008. Disponível em: <https://cs.calvin.edu/courses/cs/262/kvlinden/references/rodriguez-restfulWS.pdf>. Acesso em: 08 jan. 2026.

ROGERS, K.; ROSS, S. *Omeka and the Digital Humanities*. **Journal of Digital Humanities**, v. 1, n. 1, p. 23-34, 2010.

SANTOS, Henrique Machado dos; FLORES, Daniel. Políticas de preservação digital para documentos arquivísticos. **Perspectivas em Ciência da Informação**, v. 20, n. 4, p. 197–217, out./dez. 2015. Disponível em: <https://www.scielo.br/j/pci/a/s5JqMw9Mkk3VVgKpCK7GRZC/?lang=pt> Acesso em: 10 dez. 2025.

SANTOS, V. M.; FLORES, D. A implantação de um repositório digital confiável: o caso do Arquivo Público do Estado do Rio Grande do Sul. **Revista do Arquivo Público do Estado do Rio Grande do Sul**, n. 1, p. 1-20, 2018.

SANTOS, Maria Célia Teixeira Moura; LOUREIRO, Maria Lúcia de Niemeyer Matheus. O incêndio do Museu Nacional: patrimônio, memória e responsabilidade do Estado. **Revista Museologia & Interdisciplinaridade**, Brasília, v. 7, n. 14, p. 9-24, 2018.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. Florianópolis: UFSC, 2005.

SILVA, Ana Celina Figueira. Planejamento e política de acervos digitais: desafios e possibilidades. **Revista Mouseion**, n. 42, 2022.

SIQUEIRA, J.; MARTINS, D. L.; MEDEIROS, V. N. Mapeamento de metadados no Tainacan: nova funcionalidade para os museus digitais vinculados ao Instituto Brasileiro de Museus. **Advanced notes in information science**, v. 2, p. 184-194, 2021.

SILBERSCHATZ, Abraham; GALVIN, Peter Baer; GAGNE, Greg. **Sistemas operacionais: conceitos e aplicações**. 9. ed. São Paulo: Pearson, 2015.

SONG, Peijian *et al.* The ecosystem of software platform: a study of asymmetric cross-side network effects and platform governance. **MIS Quarterly**, v. 42, n. 1, p. 1-25, 2018.

STALLINGS, W.; BROWN, L. **Segurança de computadores: princípios e práticas**. 4. ed. São Paulo: Pearson Education do Brasil, 2017.

STEINHART, Gail. Establishing trust in a chain of preservation: the TRAC checklist applied to a data staging repository (DataStaR). **D-Lib Magazine**, [s. l.], v. 15, n. 9/10, 2009. DOI: 10.1045/september2009-steinhart. Disponível em: <https://www.dlib.org/dlib/september09/steinhart/09steinhart.html> Acesso em: 15 dez. 2025.

SCHWARZ, Michael A. Linux job scheduling. **Linux Journal**, [s. l.], v. 2000, n. 70es, article 32, fev. 2000. Disponível em: <https://dl.acm.org/doi/fullHtml/10.5555/350314.350332>. Acesso em: 8 jan. 2026.

TAINACAN. **Tainacan Beta 0.21.0**: arrumando a casa com PHP 8 & Vue3. [S. l.], 22 abr. 2024. Disponível em: <https://tainacan.org/blog/pt/tainacan-beta-0-21-0-arrumando-a-casa-com-php-8-vue3/> Acesso em: 2 jan. 2026.

TAINACAN. **Documentação oficial**. Brasília: UnB/L3P, 2023. Disponível em: <https://wiki.tainacan.org/> Acesso em: 2 jan. 2026.

TAINACAN. **Fórum Tainacan**. Atualização de itens pela importação. Dúvidas, 2021. Disponível em: <https://tainacan.discourse.group/t/atualizacao-de-itens-pela-importacao/597> Acesso em: 2 jan. 2026.

TAINACAN. **O Projeto**. [S. l.], [2024?]. Disponível em: <https://tainacan.org/o-projeto/> Acesso em: 2 jan. 2026.

TAINACAN. **Tainacan**: plugin do WordPress. WordPress.org Brasil, [2014]. Disponível em: <https://br.wordpress.org/plugins/tainacan/> Acesso em: 2 jan. 2026.

TAINACAN. **Tainacan**: crie repositórios e coleções digitais no WordPress. Brasília: Tainacan, 2025. Disponível em: <https://tainacan.org/> Acesso em: 10 dez. 2025.

TAINACAN. **Tutorial**: WP-API Tainacan. [2014]. Disponível em: <https://www.scribd.com/doc/283570997/Tutorial-WP-API-Tainacan> Acesso em: 2 jan. 2026.

TANENBAUM, A. S.; BOS, H. **Sistemas operacionais modernos**. 4. ed. São Paulo: Pearson Education do Brasil, 2016.

TANENBAUM, A. S.; WETHERALL, D. **Redes de computadores**. 5. ed. São Paulo: Pearson Education do Brasil, 2011.

TAYLOR, Arlene G.; JOUDRY, Daniel N. **The organization of information**. 4th ed. Santa Barbara: Libraries Unlimited, 2015.

TESSER, Gelson João. Principais linhas epistemológicas contemporâneas. **Educar em revista**, Curitiba, n. 10, p. 91-98, 1994. Disponível em: <https://revistas.ufpr.br/educar/article/view/2241> Acesso em: 8 dez. 2024.

VERMAATEN, San. Implementing Rosetta at the National Library of the Netherlands. *In*: INTERNATIONAL CONFERENCE OF THE ALLIANCE OF DIGITAL HUMANITIES ORGANIZATIONS, 2013, Lincoln. Conference Abstracts. **Proceedings...** Lincoln: University of Nebraska-Lincoln, 2013. p. 396-398.

VUE.JS. **The Progressive JavaScript Framework**. [S. l.], 2024. Disponível em: <https://vuejs.org/>. Acesso em: 8 jan. 2026.

WOODLEY, Mary S. Crosswalks, metadata harvesting, and interoperability: using the OAI-PMH to improve access to digital collections. **Library Hi Tech**, v. 26, n. 4, p. 536-553, 2008.

WORDPRESS.ORG. **Administration Screens**. [S. l.], [2024?]. Disponível em: <https://wordpress.org/documentation/article/administration-screens/>. Acesso em: 8 jan. 2026.

YOU, Evan. **Vue.js**. [S. l.: s. n., 2024]. Disponível em: <https://vuejs.org/> Acesso em: 2 jan. 2026.

XU, Jing; LIN, Hong. Knowledge domain and emerging trends in digital preservation: a bibliometric review. **The Electronic Library**, v. 39, n. 5, p. 713-733, 2021. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/EL-12-2020-0352/full/html>. Acesso em: 10 mar. 2026.

ZIERAU, Eld; SCHULTZ, Ulla B. A-posteriori-interoperability between preservation repositories. *In*: INTERNATIONAL CONFERENCE ON THEORY AND PRACTICE OF DIGITAL LIBRARIES, 17., 2013, Valletta. **Proceedings. Acrddo em http://purl.pt/24107/1/iPres2013_PDF/iPres2013-Proceedings.pdf** Valletta: Springer, 2013. p. 399-402.

ZHANG, Zhiwei; NAN, Guofang; TAN, Yong. Cloud services vs. on-premises software: competition under security risk and product customization. **Information Systems Research**, [s. l.], v. 31, n. 2, p. 848-864, jun. 2020. DOI: 10.1287/isre.2019.0919.

APÊNDICE 1

Os resultados obtidos a partir da busca estruturada nas bases de dados passaram por um processo de triagem conforme os critérios de inclusão e exclusão previamente definidos. Nesta etapa, procedeu-se à identificação e remoção de registros redundantes, resultando na exclusão de 143 duplicatas recuperadas em mais de uma base consultada. Tal procedimento assegurou a consistência do corpus de análise, evitando a contagem repetida de trabalhos e garantindo que apenas referências únicas, pertinentes ao escopo da pesquisa, fossem mantidas para avaliação integral. Os resultados triados são os seguintes:

Título	Autor(es)	Ano	Selecionado	Base de Dados
Digital Preservation Recommendations for Small Museums	Canadian Heritage Information Network (CHIN)	2024	Sim	Google Scholar; SciSpace

Digital Preservation for Libraries, Archives, and Museums	Corrado, Edward M.; Moulaison Sandy, Heather	2015	Sim	Scopus;Web of Science; Google Scholar; SciSpace
Creating a Framework for Applying OAIS to Distributed Digital Preservation	Zierau, Eld; Schultz, Mark	2013	Sim	Scopus; Web of Science; Google Scholar; SciSpace
Digital preservation: preservation planning in the OAIS model	Strodl, Stephan	2008	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Estado da arte da preservação digital e as temáticas pesquisadas por pesquisadores brasileiros	Neves, João Paulo Pastana; Grácio, José Carlos Abbud	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Digital cultural heritage: from OAIS until the personalised augmented experiences: perspectives	Marcos, Adérito	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital Preservation of Cultural Heritage for Small Institutions	Däßler, Rolf; Preuß, Ulf	2020	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus; Web of Science; Google Scholar; SciSpace
Digital Preservation Archives – A New Future Architecture for Long-term Interoperability	Kearney, Michael W.; Giaretta, David; Garrett, John; Hughes, Steven	2018	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus; Web of Science; Google Scholar; SciSpace

Digitization and Preservation of Cultural Heritage Products	Belhi, Abdelhak; Foufou, Sebti; Bouras, Abdelaziz; Sadka, Abdul H.	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Digitization and preservation of cultural heritage: The CEPROQHA approach	Belhi, Abdelhak; Bouras, Abdelaziz; Foufou, Sebti	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
DIGITAL ARCHIVES AND PRESERVATION IN CLOUD FOR THE INTANGIBLE CULTURAL HERITAGE	黃淑淑	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Digital Preservation challenges, requirements, strategies and scientific output	Formenton, Danilo; Gracioso, Luciana de Souza	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital preservation preparedness of the National Library of Namibia for digital preservation of documentary heritage: a case study	Uutoni, Wilhelm Elinashe	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Digital preservation: terminology, techniques, testing and trust	Giaretta, David	2010	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Auditoria de repositórios arquivísticos digitais confiáveis: uma análise das normas ISO 14721 e ISO 16363	Santos, Henrique Machado dos	2018	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Context and its role in the digital preservation of cultural objects	Beaudoin, Joan E.	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
From Oais to DPS to NDHA	Knight, Steve	2009	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

Convergence of digital preservation practices with digital repositories management practices for the Research Output and Cultural heritage	CONSTANTINESCU, Nicolaie; REPANOVICI, Angela	2021	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
A Risk Register for Digital Curation aligned with the OAIS reference model	Bairrao, Ana Raquel Correia	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Digitalização, preservação digital e acesso em instituições de memória: contributos para o projeto Museu Digital da U. PORTO	Rua, João Pedro Almeida	2016	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital Preservation (challenges, preparedness and reaction)	Perossini, F.; Boi, S.; Capua, M. C.	2020	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
and Information Interoperability: A Digital Preservation Perspective	Kearney, Mike; Giaretta, David; Garrett, John; Hughes, Steven	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Advanced Digital Preservation	Wilson, Andrew	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital Preservation	Nadal, Jacob	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Cost model for digital preservation: Cost of digital migration	Kejser, Ulla Bøgvad; Nielsen, Anders Bo; Thirifays, Alex	2011	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

Finding access and digital preservation solutions for a digitized oral history project: A case study	Matusiak, Krystyna K.; Tyler, Allison; Newton, Catherine; Polepeddi, Padma	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Media Archives and Digital Preservation: Overcoming Cultural Barriers	Delaney, Beth; De Jong, Annemieke	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
3D Data Long-Term Preservation in Cultural Heritage	Amico, Nicola; Felicetti, Achille	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

DIGITAL PRESERVATION OF UKRAINIAN AUDIOVISUAL HERITAGE DURING WARTIME: CHALLENGES AND INSTITUTIONAL PRACTICES	Kovalenko, Yuliia	2025	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital Preservation: How to Be Trustworthy	Bountouri, Lina; Gratz, Patrick; Sanmartin, Fulgencio	2018	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace

Curadoria e preservação digital: um estudo sobre o Portal Brasileira Fotográfica	Assis, Isabella Carolina Pongelupe	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
--	--	------	---	--

Encoding power: The scripting of archival structures in digital spaces using the Open Archival Information System (OAIS) reference model	Bettivia, Rhiannon Stephanie	2016	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
--	---------------------------------	------	---	--

Introduction to the Session: Conceptual Framework and Chain of Custody for Sustaining the Digital Preservation	Guercio, Maria	2010	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital curation e long-term digital preservation nelle biblioteche.	Pasqui, Valdo	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

Implementing digital
preservation in
repositories: knowledge
and

practices/Implementacao
da preservacao digital
em repositórios:
conhecimento e
práticas/Implementacion
de la preservacion digital
en repositórios
conocimientos y
práticas.

Pavão, Caterina
Groposo;
Caregnato, Sonia
Elisa; da Rocha,
Rafael Port

2016

Não - Trata
de algum
buscador,
mas não
diretamente
sobre o
assunto de
preservação
digital
ligado
diretamente
ao tema.

Scopus;Web of
Science;
Google
Scholar;
SciSpace

A digital archives
framework for the
preservation of cultural
artifacts with
technological
components

Boutard,
Guillaume;
Guastavino,
Catherine; Turner,
James

2013

Não - Trata
de algum
buscador,
mas não
diretamente
sobre o
assunto de
preservação
digital
ligado
diretamente
ao tema.

Scopus;;
Google
Scholar;
SciSpace

for Digital Preservation	Verheul, Ingeborg	2006	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
Digital records pathways: topics in digital preservation	Rogers, Corinne	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

From theory to action: Good enough digital preservation for under- resourced cultural heritage institutions	Schumacher, Jaime; Thomas, Lynne M.; VandeCreek, Drew; Erdman, Stacey; Hancks, Jeff; Haykal, Aaisha; Miner, Meg; Prud'homme, Patrice-Andre; Spalenka, Danielle	2014	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Digital preservation: Overview of current developments	Hodge, Gail M.	2002	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

Digital Cultural Heritage Roadmap for Preservation	Fresa, Antonella	2014	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Identificação de padrões de metadados para preservação digital	Danilo, Formenton	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace
Digital preservation for libraries, archives, and museums	Corrado, Edward M.; Moulaison-Sandy, Heather	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;Web of Science; Google Scholar; SciSpace

DCH-RP and PREFORMA: Two Case Studies on the Digital Preservation of Cultural Heritage	Fresa, Antonella; Prandoni, Claudio	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Scopus;; Google Scholar; SciSpace
El OAIS en la preservación digital de archivos sonoros	Rodríguez Reséndiz, Perla Olivia	2016	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; SciSpace
Mapping significance properties in OAIS: A case study with video games extended poster abstract	Bettivia, Rhiannon S	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; SciSpace

A preservação digital em longo prazo amparada por planos de ações uma revisão sistemática de literatura	Souza, Luciana Gonçalves Silva; Aganette, Elisângela Cristina	2020	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; SciSpace
Avoiding technological quicksand: coming to grips with the preservation of digital information in Nigeria	Alegbeleye, G. O.	2009	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; SciSpace
A brief history of digital preservation	Baucom, Erin	2019	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; SciSpace

'Grey crossroads' in cultural heritage preservation and resource management.	De Biagi, Luisa; Puccinelli, Roberto	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
---	---	-------------	--	---

Main Architectures Used in ETL as a Tool, Necessary for the Integration of Data in Large Volumes-a Task in the Field of Digital Preservation of Cultural Heritage	Angelov, Kamen	2016	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
--	-----------------------	-------------	--	---

Aspectos de preservação digital para fotografias forenses digitais	Silva, Ezequiel Sousa	2020	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
A nine dimensional framework for digital cultural heritage organizational sustainability: A content analysis of the LIS literature (2000–2015)	Eschenfelder, Kristin R.; Shankar, Kalpana; Williams, Rachel D.; Salo, Dorothea; Zhang, Mei; Langham, Allison	2019	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
A foundation for automatic digital preservation	Ferreira, Miguel; Baptista, Ana Alice; Ramalho, José Carlos	2006	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar

Arquivamento da web e preservação digital	Rockembach, Moisés; Pavão, Caterina Groposo	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
Digital preservation: interoperability ad	Dobreva, Milena; Ruusalepp, Raivo	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; BDTD
Preservação do patrimônio documental arquivístico em ambiente digital	Santos, Henrique Machado dos; Flores, Daniel	2017	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; BDTD

Preservation and Digital Repositories: Connections, Possibilities, and Needs	Corujo, Luís; Revez, Jorge; da Silva, Carlos Guardado; de Macedo, LS Ascensão	2025	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
Preserving the cultural heritageAn investigation into the feasibility of the OAIS model for application in small organisations	Tedd, Lucy A.; Ellis, David; Spence, Jaqueline	2006	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
Um perfil de padrões de metadados de preservação para documentos arquivísticos audiovisuais	Machado, João Guilherme Nogueira	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; BDTD

PRESERVAÇÃO DIGITAL DA WEB	TERRADA, GABRIELA AYRES FERREIRA		Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
Método de Elaboração de um Plano de Preservação Digital	Krause, Carla	2019	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; BDTD
Repositório arquivístico digital confiável para o patrimônio documental oriundo do processo judicial eletrônico	Rodrigues, Mateus de Moura	2015	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar; BDTD

REVISION OF DIFFERENT IMPLEMENTATIONS FOR DIGITAL PRESERVATION: TOWARDS A METHODOLOGICAL PROPOSAL FOR PRESERVING AND AUDITING IR RELIABILITY	LA PRESERVACIÓN, UNA PROPUESTA METODOLÓGICA PARA; DE RI, CONFIABILIDAD; DE PRESERVAÇÃO, PARA UMA PROPOSTA METODOLÓGICA; DE RI, E. AUDITORIA DE CONFIANÇA	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
--	--	---	--------------------------------

Preserving Digital Cultural Heritage: A Call for Participatory Models	Chou, Rose	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Web of Science; Google Scholar
---	------------	------	---	--------------------------------

Relational databases digital preservation	Freitas, Ricardo André Pereira	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
The power of imaginary users: Designated communities in the OAIS reference model	Bettivia, Rhiannon S.	2016	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Preserving digital heritage: assessing the compliance of digital repositories in South Africa with OAIS and TDR standards	Balogun, Tolulope	2025	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Uma outra face dos metadados: informações para a gestão da preservação digital	Sayão, Luís Fernando	2010	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar; BDTD
The Application of OAIS Model as a Framework for Digital Preservation of Indigenous Knowledge Systems: The Roles of Educational Managers	Okenjom, Godian Patrick; Asuquo, Michael Ekpenyong	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Uso do archivematica no Brasil	Machado, João Guilherme Nogueira; Arellano, Miguel Angel Márdero	2022	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
The use of metadata systems for the preservation of digital records in cultural heritage institutions	Masenya, Tlou Maggie	2021	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Metadata standards in web archiving technological resources for ensuring the digital preservation of archived websites/Padroes de metadados no arquivamento da web: recursos tecnologicos para a garantia da preservacao digital de websites arquivados.

Formenton, Danilo; de Souza Gracioso, Luciana

2022

Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.

Google Scholar

Rethinking digital preservation: definitions, models, and requirements

Wilson, Thomas C.

2017

Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.

Scopus;Web of Science; Google Scholar

New Standards for Archive and Information Interoperability: A Digital Preservation Perspective	Kearney, Mike; Giaretta, David; Garrett, John; Hughes, Steven	2025	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	--	------	--	----------------

Plano de Preservação Digital da VideoSaúde: estratégias para a gestão de documentos audiovisuais VideoSaúde Digital Preservation Plan: strategies for the management of audiovisual documents Plan de Preservación Digital VideoSaúde: estrategias para la gestión de	Buarque, Marco Dreer; Machado, João Guilherme Nogueira; Pontes, Eliane Batista		Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	--	--	--	----------------

OAIS compliant preservation workflows in an AV archive	De Jong, Annemieke; Delaney, Beth; Steinmeier, Daniel	2013	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
--	--	------	--	----------------

Repositorios digitales: aplicación del modelo OAIS y los esquemas de metadatos a la conservación del patrimonio documental archivístico	Raventós, Pepita; de Archivo, Servicio	2010	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	--	------	--	----------------

Sustainable digital preservation-PPREDIS as a tool to ensure the protection and accessibility of documentary cultural heritage.	Daza, Nelson Javier Pulido; Ríos, Mayra Alejandra Bustos; Cabrera, Alba Catalina; Carranza, Eiser Elena Rengifo Reyes		Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
The Hipátia Model: Paths toward the Brazilian Archival Digital Preservation Era	Braga, Tiago; Alves, Larissa	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Towards Sustainable Knowledge: Digital Preservation in the Era of Open Science	Dos Santos, Andrey Anderson; Dos Santos Bahia, Eliana Maria; Bianchi, Isaias	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

The role of metadata in the longevity of cultural heritage resources	Dobreva, Milena; Ikononov, Nikola	2009	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Towards preserving cultural heritage of Finland	Helin, Heikki; Koivunen, Kimmo; Lehtonen, Juha; Lehtonen, Kuisma	2012	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Preservation for diverse users: digital preservation and the "Designated Community" at the Ontario Jewish Archives	Moles, Nathan	2022	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

PRESERVAÇÃO DIGITAL: tendências atuais dos conceitos e técnicas	Cunha, Jacqueline de Araújo; Lima, Marcos Galindo	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	---	------	--	----------------

Stability of cultural information in unstable environments: Data management for digital preservation of immovable cultural heritage against natural hazards	Fisher, Michael T.	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
--	--------------------	------	--	----------------

The case for yet another digital preservation evaluation tool	Gillesse, Robèrt; Vanstappen, Henk	2014	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Plano de Preservação Digital da VideoSaúde: estratégias para a gestão de documentos audiovisuais	Buarque, Marco Dreer; Machado, João Guilherme Nogueira; Pontes, Eliane Batista	2020	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
The CASPAR approach to digital preservation	Giaretta, David	2007	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Modelo de WorkFlow como processo de simbiose entre a preservação digital do patrimônio cultural documental e a lei geral de proteção de dados-LGPD	de Carvalho, Vagner da Silva; Pedrazzi, Fernanda Kieling	2022	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar (6 em diferentes links)
Strategies for digital preservation	Holdsworth, David	2006	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar (2 vezes em diferentes links)

Why Digital Preservation Should Be Treated as Public Policy	Silva, Wellington Da	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar (3 vezes em diferentes links)
---	----------------------	------	---	--

Padrões de metadados no arquivamento da web: recursos tecnológicos para a garantia da preservação digital de websites arquivados	Formenton, Danilo; Gracioso, Luciana de Souza	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
--	---	------	---	----------------

Online cultural heritage as a social machine: a socio-technical approach to digital infrastructure and ecosystems	Pereda, Javier; Willcox, Pip; Candela, Gustavo; Sanchez, Alexander; Murrieta-Flores, Patricia A.	2025	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	--	------	---	----------------

Preserving Digital Data without an Archive: Illuminating the path towards digital preservation through knowledge of essential requirements and strategies	Trognitz, Martina; Baró, Sabina Batlle; Matskevich, Sveta; Péter, Réka Katalin; Szabó, Alexandra Diána; Gál, Márton; de Almeida, Vera Moitinho		Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
---	--	--	---	----------------

Towards effective management and preservation of digital cultural heritage resources: an exploration of contextual factors in Ghana	Boamah, Eric	2014	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Preservação digital: tendências atuais dos conceitos e técnicas	de Araújo Cunha, Jacqueline; Lima, Marcos Galindo	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
What Small Projects Producing Digital Resources Need to Know about Digital Preservation?	Ikonomov, Nikola; Dobрева, Milena	2009	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Meta-informação de preservação digital: o caso de estudo da gestão de cursos no arquivo da FEUP	do Rego, Bruno José Freitas	2013	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Re-envisioning access for the digital preservation community: challenges, opportunities and recommendations	Talboom, Leontien	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Preservação da memória à luz da sustentabilidade digital: estudo de caso na Universidade Federal de Sergipe	Costa, Luani Messias da	2024	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar

Tools for Preservation and Use of Complex and Diverse Digital Resources	Giaretta, David	2009	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar
Metadata standards in web archiving technological resources for ensuring the digital preservation of archived websites	Formenton, Danilo; Gracioso, Luciana de Souza	2023	Não - Trata de algum buscador, mas não diretamente sobre o assunto de preservação digital ligado diretamente ao tema.	Google Scholar